The AAPS Journal, Vol. 16, No. 5, September 2014 (© 2014)
DOI: 10.1208/s12248-014-9633-4

White Paper

Theme: Best Practices for Bioanalytical Methods: Recommendations from the Global Bioanalysis Consortium
Guest Editors: Binodh DeSilva and Philip Timmerman

Small Molecule Specific Run Acceptance, Specific Assay Operation, and Chromatographic Run Quality Assessment: Recommendation for Best Practices and Harmonization from the Global Bioanalysis Consortium Harmonization Teams

Eric J. Woolf,[1] Stuart McDougall,[2] Douglas M. Fast,[3,29,30] Maristela Andraus,[4] Matthew Barfield,[5] Michael Blackburn,[2] Ben Gordon,[6] David Hoffman,[7] Noriko Inoue,[8] Gabriel Marcelin-Jimenez,[9] Amy Flynn,[10] Richard LeLacheur,[11] Scott Reuschel,[12] Ravisankar Santhanam,[13] Patrick Bennett,[14] Barbara Duncan,[15] Roger Hayes,[16] Berthold Lausecker,[17] Abhishek Sharma,[18] Kazutaka Togashi,[19] Ravi Kumar Trivedi,[20] Miguel Vago,[21] Stephen White,[5] Hollie Barton,[14] John A. Dunn,[22] Raymond H. Farmen,[23] Katja Heinig,[24] Christopher Holliman,[25] Junji Komaba,[26] Maria Francesca Riccio,[27] and Elizabeth Thomas[28]

Received 30 January 2014; accepted 30 May 2014; published online 25 June 2014

Abstract. Consensus practices and regulatory guidance for liquid chromatography-mass spectrometry/mass spectrometry (LC-MS/MS) assays of small molecules are more aligned globally than for any of the other bioanalytical techniques addressed by the Global Bioanalysis Consortium. The three Global Bioanalysis Consortium Harmonization Teams provide recommendations and best practices for areas not yet addressed fully by guidances and consensus for small molecule bioanalysis. Recommendations from all three teams are combined in this report for chromatographic run quality, validation, and sample analysis run acceptance.

KEY WORDS: bioanalytical assay validation; LC-MS/MS; sample analysis; small molecule.

INTRODUCTION

This summary is based on the discussion among scientific representatives from Mexico, Brazil, Argentina, Japan, India, the UK, Germany, Switzerland, and the USA serving on the Global Bioanalytical Consortium S1, S2, and S3 harmonization teams (HT). These are recommended best practices for specific laboratory operations associated with run acceptance, assay operation, and chromatographic run quality assessment not addressed clearly or not addressed at all in current regulatory guidances or guidelines. In some cases, the team assessing an activity intentionally chose to recommend practices based on good scientific principles that are different from what current regulatory guidances may

[1] Merck Research Laboratories, WP 75B-300, West Point, Pennsylvania 19486, USA
[2] Covance Laboratories Ltd, Willowburn Avenue, Alnwick, Northumberland NE66 2JH, UK
[3] Covance Laboratories Inc, Madison, Wisconsin, USA
[4] ChromAnalysis, São Paulo, Brazil
[5] GlaxoSmithKline, Ware, UK
[6] Servier (I.R.I.S), Suresnes-Cedex, France
[7] Sanofi, Bridgewater, New Jersey, USA
[8] JCL Bioassay, Osaka, Japan
[9] Global Bioanalytical Consulting SC, Mexico City, Mexico
[10] QPS LLC, Newark, Delaware, USA
[11] Agilux Laboratories, Worcester, Massachusetts, USA
[12] Tandem Laboratories, Salt Lake City, Utah, USA
[13] GVK Biosciences Private Limited, Ahmedabad, India
[14] PPD, Richmond, Virginia, USA
[15] Duncan Bioanalytical Consultancy LLC, Ledyard, Connecticut, USA
[16] MPI Research, Mattawan, Michigan, USA
[17] CRS Mannheim, Mannheim, Germany
[18] Advinus Therapeutics Ltd, Bangalore, India
[19] Sumika Chemical Analysis Service Ltd, Osaka, Japan
[20] Syngene International, Bangalore, India
[21] Laboratorios Raffo SA, Buenos Aires, Argentina
[22] GlaxoSmithKline, Research Triangle Park, North Carolina, USA
[23] Celerion, Lincoln, Nebraska, USA
[24] F. Hoffman-La Roche AG, Basle, Switzerland
[25] Pfizer Inc, Groton, Connecticut, USA
[26] Ono Pharmaceutical Co., Ltd, Tsukuba, Ibaraki, Japan
[27] CAEP Centro Avançado de Estudos e Pesquisas LTDA, Campinas, São Paulo, Brazil
[28] Alderly Analytical Ltd, Cheshire, UK
[29] Covance Laboratories Inc., MC-11, 3301 Kinsman Blvd., Madison, Wisconsin 53704, USA
[30] To whom correspondence should be addressed. (e-mail: douglas.fast@covance.com)

 1550-7416/14/0500-0885/0 © 2014 American Association of Pharmaceutical Scientists

include. For the large majority of practices that have been discussed and reviewed and for which consensus has been reached over the interval since the first Workshop on Analytical Methods Validation: Bioavailability, Bioequivalence, and Pharmacokinetic Studies in December, 1990 (Crystal City-I), the teams agreed that international harmonization and consistent practices dictate that no changes should be recommended.

The recommendations are applicable to a wide variety of molecules, including but not limited to, pharmaceuticals, peptides, proteins, and biomarkers that elicit a quantitative chromatographic instrument response. The discussion focused predominately on high-performance liquid chromatography-mass spectrometry/mass spectrometry (LC-MS/MS) techniques, but the general principles should be translatable to other forms of chromatography and detection.

Our recommendations are divided into three sections, chromatographic run quality, validation, and sample analysis run acceptance, and we will address topics in that sequence.

CHROMATOGRAPHIC QUALITY

The basis of any robust and reliable bioanalytical assay is to ensure selectivity during extraction, chromatographic separation, and detection.

Chromatographic Selectivity

Although there have been significant advances in chromatography over the past decade, the basic principles for a good chromatographic separation are still valid. As such, it is recommended that the following should be considered during method establishment based on good scientific practice:

- Ensure adequate chromatographic retention, such that a capacity factor (k' for isocratic elution or k* for gradient elution) of at least 2 is obtained (isocratic: $2<k'<10$, gradient: $2<k^*<10$). Maintaining a high-capacity factor ensures that the analyte of interest is adequately separated from interferences both seen and unseen and reduces the potential impact of ion suppression/enhancement, especially in incurred samples.
- Ensure adequate chromatographic separation such that the resolution (Rs) of two quantifiable peaks is at least 2.0 for a robust assay. Maintaining baseline resolution overcomes the potential for peak area integration errors while allowing for some deterioration in chromatographic performance throughout a run. Alternatively, quantitation using peak height may be considered and may be more effective with forms of protein LC-MS/MS analysis using target peptides and enzymatic digestion due to complicated baseline interferences.
- Maintaining adequate chromatographic efficiency (N) such that the theoretical plate number should be over 2,000 where applicable.
- Ideally, chromatographic peaks should be symmetrical, but asymmetry of an analyte peak is acceptable provided that the degree of asymmetry observed in the incurred samples is reflected in the calibration and quality control (QC) samples.

Once an assay is developed and validated, it is essential that the chromatographic selectivity is monitored, especially in study samples, to assess impact from metabolites, exogenous xenobiotics, or potential interferences from dose formulation.

Chromatographic Sensitivity

The sensitivity of any bioanalytical assay is critical to its quantitative performance but for chromatographic assays, this is related directly to the signal (generated from analyte) *versus* the noise (generated from matrix or instrument). Both the FDA and EMA bioanalytical guidances (1, 2) indicate that the chromatographic response at the lower limit of quantitation (LLOQ) should be at least five times the response compared to the blank response which is often (incorrectly) interpreted to be a measure of the assay signal to noise (S/N). However, in certain circumstances, this gives rise to some ambiguity, since S/N is related to both electronic noise (instrument) and chemical noise (matrix) and is often calculated from data acquired in a single chromatographic run. This is subtly different from a discrete peak in blank samples that exhibits the same retention time of the analyte (interference).

When assays are developed, it is often difficult to accurately and reliably assess the S/N and it is recommended that S/N is used as a tool for bioanalytical method development but that reaching S/N of over five is not what the EMA and FDA Guidances recommend (1, 2). Rather, they expect a comparison of separate injections of blanks and LLOQ calibration standards as one of the assessments of the response at the LLOQ. Ultimately, the definitive evidence that the required assay performance has been achieved is the precision and accuracy calculated at the LLOQ during validation.

Chromatographic Data Acquisition

There are many parameters that affect chromatographic performance but often little is understood with regard to the acquisition of the chromatographic signal itself. Data sampling frequency can have a significant effect on the accuracy and precision of the assay (3–6). If the sampling frequency is too low, peak detail will be lost and measurement accuracy reduced. If the sampling frequency is too high, unwanted baseline noise peaks will be detected and measured. It is therefore recommended that a peak is defined by at least 10 points, optimally 15 to 20 (5). However, the data-sampling frequency defined in validation must be applied to production (incurred study samples) and any changes in the data sampling frequency would constitute a change to the validated method and would require a documented assessment on assay performance.

Instrumental measurement, as used in chromatographic analysis, consists of two components: the signal that contains information about the analyte (signal) and the background (noise) signal. Noise can be further divided into long-term

noise derived from external source variables (temperature, humidity, stationary phase bleed, baseline drift, etc.) and short-term high-frequency noise derived from each component of the analytical instrument (detector noise, power source, pressure pulsations, etc.). In LC-MS/MS analysis, vendor-specific hardware (filtering) and software (smoothing) methods are commonly employed to reduce noise and improve the S/N in instrumental analysis. When filtering and/or smoothing (such as Savitzky-Golay, moving average, etc.) are used, it is recommended that what was defined and used in validation must be applied to the measurement of study samples. Consequently, any changes in filtering or smoothing methods occurring after an assay has been validated should be assessed with respect to assay performance before use.

Chromatographic Integration and Reintegration

The chromatographic performance of a bioanalytical assay, in terms of its sensitivity and selectivity, is critical in reaching the required accuracy, precision, and robustness for long-term use. However, it is also very important to note that good peak integration is also a result of good chromatography. Integration is of use in controlled circumstances when chromatography is good but it cannot improve bad chromatography; only the analyst can do that (1).

Peak integration is one area of chromatographic analysis that has been subject to the most regulatory scrutiny as it has a potential for misuse and fraud (7). To fully understand peak integration, it is important to understand that this is only one facet of the chromatographic process. Much has been discussed with regard to having stricter control and potentially "locking" the chromatographic integration parameters at assay validation under the misconception that this will somehow improve quality. However, it must be accepted that chromatographic analysis is a dynamic process and that changes in peak integration during the lifetime of an assay will be required. LC-MS/MS (and hence peak integration) can be affected by changes in both instrument response (manufacturing differences, ageing, sensitivity, etc.) and chromatographic performance (changes in peak shape, noise, etc.) and as such, it is scientifically invalid to assume that the chromatographic integration parameters must also be fixed. It is therefore recommended that the integration parameters used in assay validation be considered as an initial guide for any subsequent run integration, but ultimately that the integration used adequately characterizes the peak of interest and is consistently applied to calibration, QC, and incurred samples within a run.

The process of chromatographic integration can be further subdivided into the "initial integration" and "reintegration," which are quite distinct processes and need clear definition and control. It is our recommendation that initial integration is defined as the process by which the area (or height) of a chromatographic peak is adequately defined by trained personnel using the most appropriate parameters prior to regression. This definition includes the following:

- Automated integration (integration parameters affecting all chromatograms and can be adjusted on a run by run basis).
- Modified automated integration (integration parameters affecting a specific chromatogram and adjusted on a chromatogram-by-chromatogram basis).
- Manual integration (hand-drawn integration performed by a trained analyst for a specific chromatogram on a chromatogram-by-chromatogram basis).

In practice, there is little real distinction between "modified automated integration" and "manual integration" as both require human involvement and adjustment on a "chromatogram-by-chromatogram" basis. The emphasis for this process is on "trained personnel" and "prior to regression" such that the analyst can adequately integrate the chromatographic peak yet is essentially blind to the final regressed concentration. It is recommended that "automated," "modified automated," and "manually" integrated chromatograms be retained for inspection, that the method and reason for adjusting the integration parameters be clearly documented, and that the whole process be defined under a standard operating procedure (SOP). The user should not rely solely on the audit trail to capture this information unless it can be proven that the audit trail contains all changes necessary for reconstruction.

With regard to "reintegration," it is proposed that this is defined as the process that occurs post regression, i.e., after the back-calculated concentration, acceptance of calibration and QC samples, and acceptance or rejection of the whole run has been made. Reintegration should only occur after rejection of the initial integration following documented peer review by trained personnel. Reintegrated data as well as the initial integrated data should be retained and the process defined by SOP. Where appropriate and practical, it is recommended that process control, such as system access rights and audit trails, be implemented to ensure transparency and validity. While we recognize the potential for misconduct (peak skimming or peak enhancing), a clear definition, clarity of process and separation of integration from regression will improve confidence in the overall data quality.

Provided that chromatographic peaks are adequately and consistently integrated, the ultimate control on a run-by-run basis is made by the calibration, QC, and internal standard acceptance criteria.

Chromatographic Data Review and Retention

To ensure enduring quality and confidence of bioanalytical data, it is recommended that post regression review of every run should be performed by an individual separate from the analyst that performed the integration and regression. The review should include an audit of all chromatograms for consistency of both chromatography and integration. The analyst performing integration must have documented evidence in both the general principles of chromatographic integration and the specifics of the software system used, whereas the reviewer must have documented evidence and experience of chromatographic data audit. Both the initial and reintegrated chromatograms (as defined earlier), together with associated integration parameters, should be retained and be available for regulatory review if requested. Only the final and valid integrated chromatogram is used to generate an analyte concentration. Generating and

reporting the analyte concentration data from earlier preliminary chromatograms is not scientifically valid and may not be practical or possible in many cases.

VALIDATION

Calibration Curves/Calibration Standards

Most criteria for calibration curves are already well accepted and consensus in both practice and guidance has been achieved. The LLOQ of an assay is defined as the lowest concentration of analyte in a sample which can be quantitated reliability (e.g., with acceptable accuracy and precision). Acceptable accuracy and precision at the LLOQ is 20%.

The HT recommends that the LLOQ of an assay should be set equal to or less than 5% of the anticipated pharmacokinetic profile maximum concentration where feasible. The LLOQ should enable adequate characterization of the pharmacokinetics of the analyte.

The LLOQ is generally considered to be equal to the lowest calibration standard. There is no need to specify a limit of detection (LOD) in addition to the LLOQ as concentration data below the LLOQ are not reportable in analyzed incurred samples.

Each run (validation or sample analysis) from which quantitative results are to be obtained should include at least a single set of calibration standards. Depending on the assay, duplicate calibration standards placed at the beginning and end of each run may be preferred. Generally, at least six calibration levels (including LLOQ and upper limit of quantitation (ULOQ)) are required.

When calibration standards are run in duplicate, both results for a calibration standard at a specific concentration may be omitted from the regression except for those placed at the LLOQ or the ULOQ in the core accuracy and precision runs. Truncation of the calibration range is not permitted in core accuracy and precision runs but truncation is permitted in subsequent validation or sample analysis runs provided the low, middle, and high QCs are still within the calibration range.

Calibration standards should ideally be prepared in the same matrix as that of the intended study samples; the use of surrogate matrices for calibration standards should be scientifically justified. The standards are prepared by spiking the blank matrix with known concentrations of the analyte. The standards are generally spiked freshly when an assay is first initiated. Once long-term stability data on the analyte in matrix are generated, one can consider preparing standards in bulk, aliquoting them in daily use portions and then freezing the aliquots. The use of frozen calibration standards may minimize inter-run variation. The solvent content of standards is not critical as long as the quality control samples accurately back calculate, but practical considerations of protein-solvent interactions may limit the amount of solvent present.

Regression Models

The recommendation of the HT was largely in accordance with the EMA, FDA, and ANVISA guidance documents (1, 2, 8). The specific procedure used to determine a regression model should be described in a SOP (preferably) or similar controlling document. This document should describe what regression models are permitted, what regression model is used by default, and the procedure and criteria for evaluation and selection of an alternate regression model where appropriate for a given application.

With respect to the assignment of a default regression model, discussion of determining the "simplest model" was limited to linear *vs.* nonlinear regression models. Linear is considered the simplest and as such may be the default regression model (does not require justification). Weighting was considered within the scope of linear regression and was not considered a separate issue. Therefore, it was considered acceptable that a default regression model that is $1/x$- or $1/x^2$-weighted linear regression may be designated in the SOP or other controlling document. When a regression model is determined in validation, it should not be altered in subsequent sample analysis.

QC Placement and Preparation

Criteria for placement and use of QC samples (in validation) are already well accepted, and consensus in both practice and guidance has been achieved in most details. The HT does recommend that the middle QC concentration should be set according to a SOP. Options for placement include near 50% of the high QC or at or near the geometric mean of the low- and high-QC samples. Placing the middle QC at or near the geometric mean may give better positioning in log concentration space. When no *a priori* range expectation exists for sample analysis, it may be prudent (but is not required) to place two QC concentrations between the low- and high-QC levels to increase the possibility that the sample analysis range will be represented by QC concentrations with few additional assessments. This option should be defined by SOP if it is used.

Quality control samples prepared at the LLOQ are used during validation to assess accuracy and precision at the LLOQ. These samples may be freshly prepared on the day of analysis or stored frozen provided that supporting stability data is available. QCs at the LLOQ are not run during sample analysis.

It is a best practice with respect to increasing assay confidence prior to use to determine during validation intra- and inter-day precision at the LLOQ using the LLOQ QCs. A ULOQ QC is not recommended for inclusion in the validation.

QC samples should be prepared to mimic the study samples as closely as possible. In this regard, the organic solvent content of QCs should be minimized; these samples should contain no more than 5% solvent. QC samples should mimic any additives (anticoagulant, stabilizer and/or modifier) that study samples may contain. They should be stored under the same conditions as study samples.

Internal Standards

Internal standards are generally incorporated into assays to compensate for sources of systemic variation (recovery variation, ionization effects, etc.). Suitability of a particular internal standard must be demonstrated as part of the method

validation process. Stable isotope-labeled internal standards are generally preferred for mass spectrometry-based assays. For such internal standards, suitability is demonstrated by proving a lack of isotope exchange. In addition, the standard zero, containing internal standard only, should show a lack of significant response (less than 20% of that of the LLOQ) in the analyte channel. As long as such suitability is demonstrated, there are no purity criteria for internal standards and certificates of analysis are not required for them.

The concentration of internal standard employed in an assay should be determined during assay development. Generally, it is recommended to use a concentration that results in a response equivalent to 30–50% of that of the ULOQ; however, the concentration may be adjusted lower to eliminate detection of interfering impurities or raised to eliminate adsorption issues.

From a best practice standpoint, it is recommend to add the internal standard separately to samples and allow equilibration, but it is recognized that if the internal standard is added along with precipitation solvent and validation criteria are met, there is no issue.

Generation of stability data for internal standard solutions is considered to be a best practice but is not an absolute requirement for assay validation.

In agreement with the EMA guidance (1), it is recommended that a SOP describing internal standard response be in place prior to both validation and study sample analysis, but this does not need to have restrictive mathematical criteria (e.g., mean ±50% or similar). Scientific judgment should take precedence and should be able to discriminate between missing internal standard and double spikes as common events. The impact of internal standard drift within a run, internal standard changes in specific subjects, or other internal standard response events cannot be mandated by a simple mathematical rule.

Dilution Integrity

Dilution integrity refers to the ability to dilute with control matrix prior to the analysis of a sample containing the target analyte at a concentration higher than that of the ULOQ into the concentration range of the assay and obtain an accurate estimate of the concentration prior to dilution.

Dilution integrity should be assessed during assay validation by preparing and analyzing replicate samples prepared at a concentration equal to or greater than the highest concentration expected in study samples. These samples are then frozen, thawed, diluted into the range of the standard curve with control matrix, and analyzed. The mean analyzed value should be within 15% of nominal, and precision of the replicates should be equal to or better than 15% coefficient of variation (CV). Validation of a single dilution concentration at a single dilution factor (e.g., ×10) at or above the highest concentration expected in study samples is sufficient.

The concentration at which dilution integrity is assessed represents the highest reportable concentration of the assay. If samples are found with concentrations that exceed the dilution integrity limit, dilution integrity should be reassessed using more concentrated fortified samples. Solubility as well as stability issues should be considered in all dilution integrity assessments.

Multiple Analyte Analysis Considerations

For initial validation of an assay with multiple analytes, it is recommended that at least one of the three accuracy and precision runs be acceptable for all analytes in the same run; however, exceptions to this recommendation may be allowed based on the number of analytes in the method and/or the degree of complexity required to analyze all analytes (e.g., separate elution steps, separate LC-MS injection parameters, etc.) from a single-sample aliquot.

Acceptance criteria should be ±15% (±20% at the LLOQ) for the analyte (s) of interest intended for "regulated" bioanalysis; however, acceptance criteria for some analytes may be expanded depending on the nature of the analysis (i.e., fit for purpose) and should be based upon the performance of the analytes during validation.

Selectivity

Selectivity, that is, the ability of the analytical method to be able to differentiate the analyte (s) of interest as well as internal standard from endogenous components should be proved using at least six individual sources of the appropriate blank matrix. The lots of matrix should be individually analyzed and evaluated for interference. If one of the lots fails, the number of lots should be increased to 10, and at least 90% of the lots need to meet acceptance criteria. Lack of interference is demonstrated by the blank lot generating a response of less than 20% of that of the LLOQ standard and less than 5% of that of the internal standard response.

One lot of hemolyzed matrix should be included in the six lots assessed for selectivity. This is a requirement for all species. Hemolyzed matrix may be prepared by adding 2% lysed whole blood to plasma. Multiple degrees of hemolysis generally do not need to be assessed; the 2% whole blood in plasma is felt to represent a worst case scenario.

For assays intended to support clinical studies, one of the lots should be representative of a hyperlipidemic sample. Assessment of the impact of hyperlipidemic plasma on assays intended to support preclinical studies is not required unless samples from special populations of animals, known to be hyperlipidemic, are intended to be analyzed.

Generally, stripped matrices are to be avoided for bioanalytical work. These may be an appropriate surrogate matrix for calibration curves, but QC samples should be prepared in nonstripped matrix.

Matrix blanks should be included in production analyses and will be used to fortify calibration standards and quality controls as well as act as blank/zero samples. If needed, pooled predose samples may be used as matrix blank samples if special circumstances require this approach (e.g., endogenous analytes, rare or limited matrices). During validation analyses, matrix blanks will be used to assess carryover, recovery, and matrix effects/factor.

Salt/Counter Ion Changes of the Analyte and Anticoagulant

When changing the anticoagulant (e.g., EDTA to Heparin, EDTA to Fluoride/Oxalate), these matrices should be regarded as different matrices and a full revalidation of the method is required including all matrix stability assessments but there is no need to reestablish stability for unchanged stock solutions.

When changing only the salt or counter ion in matrices with the same validated anticoagulant (e.g., NaHeparin to LiHeparin, K_3EDTA to K_2EDTA or Na_2EDTA), these matrices should be regarded as equal matrices and no additional validation of the method is required.

This recommendation is based upon published research which indicates that a change in anticoagulant has no impact on bioanalytical LC-MS/MS assays (9, 10). In addition, the European Bioanalysis Forum (EBF) has also recommended that "blood related matrices containing the same anticoagulants, but different counter ions, should still be regarded as equal matrices, and that no partial validation needs to be performed when changing counter ion only." (10, 11).

Increasing anticoagulant concentrations beyond that validated also does not require additional work. Decreasing anticoagulant concentration may require some stability assessments, especially if the anticoagulant serves to stabilize the analyte.

Generally, no additional experiments are required if the salt form/counter ion of the analyte changes in the course of a development program. This recommendation assumes that the solubility of the new form is equal to or greater than that of the original; compound characterization data should be leveraged to evaluate solubility differences.

Furthermore, it is recognized that the salt/counter ion of reference standards do not need to match those of administered analyte; analyte concentrations should be reported as free acid/base.

Sample and Run Reinjection

When a batch is interrupted due to any identifiable or unknown reasons, the decision whether to reinject a partial or full batch must be made. To make that decision, results of reinjection reproducibility/autosampler stability evaluations conducted during the validation are used. Reinjection of a full analytical run or of individual calibration standards or QC samples simply because the calibrators or QCs failed without an identifiable analytical cause is not acceptable.

During the reinjection reproducibility experiment, freshly prepared calibrators should not be added for the second injection of a whole batch; originally injected calibrators should be used. The reinjected QC samples are compared separately to both the original calibration curve and to the reinjected calibration curve.

Matrix Effects

The assessment of matrix effects is required for all MS-based assays regardless of whether or not a stable labeled internal standard is utilized. Causes and reduction of matrix effects should be studied during method development and processes to reduce/minimize impact should be developed.

Assessment of matrix effects is performed by determining the variability of the analyte/internal standard ratio across multiple lots of matrix at fixed concentrations and is defined in the EMA Guideline (1) as the "matrix factor." This experiment can be performed using either QCs or calibrators. In order to be considered acceptable, the variability should be less than or equal to 15% CV. The assessment should be performed at a minimum of two concentrations, one near the low end of the standard curve range and the other near the high end.

For assays intended to support human clinical studies, the assessment should be performed using six lots of matrix, including two from male donors, two from female donors, one lipemic lot, and one hemolyzed lot. If an assay is to support clinical studies in only one gender of subjects, then the six lots of matrix should include lots from the selected gender.

The need to include lots from male and female animals in the assessment of assays intended to support preclinical work should be dependent on the gender of the animals that are to be utilized in the studies. The impact of hemolysis on preclinical assays should also be assessed.

Matrix effects due to dosing vehicles, metabolites, and co-medications may manifest themselves through changes in internal standard responses during sample analysis. Such changes should be monitored and addressed on the basis of scientific judgment.

Recovery

Recovery of the extraction process associated with an assay is viewed, from a scientific point, as important to assess during assay development. In this regard, relative recovery is defined as the comparison of analyte responses from extracted samples with those of matrix blanks spiked post extraction. Absolute recovery is defined as a comparison of extracts to neat samples. Absolute recovery is impacted by ionization and/or matrix effects as well as by the extractability of the analyte from matrix. While a formal criterion with respect to minimal acceptable relative recovery was not felt to be necessary, it was agreed that relative recovery should be consistent over the dynamic range of the assay and in multiple lots of matrix. Inconsistent absolute recovery over the assay range should be investigated; inconsistent recovery may be acceptable if it can be adequately modeled through the choice of the regression model. Precision of relative recovery rather than absolute recovery is put forth as the appropriate acceptance criterion to be assessed during assay validation; a precision less than or equal to 15% CV at each assessed concentration should be obtained.

Acceptable assay precision and accuracy in multiple lots of matrix over the dynamic range eliminate the need for a separate recovery assessment if online sample preparation is employed.

Co-medication and Metabolite Interference Assessment

The impact of known or anticipated co-medications and metabolites on assay selectivity should ideally be assessed before the analysis of samples containing the co-medications. In a patient population, it is generally

difficult to perform an *a priori* assessment of co-medications as these may vary greatly across subjects. In this case, predose samples (which are samples containing the co-medications obtained prior to the administration of study medication) should be assessed to ensure lack of interference if appropriate samples are available. Additionally, internal standard responses in post dose samples should be compared to those of standards in order to demonstrate that the co-medications do not affect the internal standard. There are a variety of other techniques that may be employed to attempt to mimic incurred samples containing co-medications and to assess assay performance in the presence of other administered drugs.

Failed Run Reporting—Validation

A table of all analytical runs together with analysis dates and status (passed or failed) should be included in the validation report; the reason for failure should be included for failed runs.

Multiple consecutive failed runs should have assignable analytical cause to continue validation. Assignable cause related to methodology triggers returning to method development. Consecutive failed runs during validation may also indicate the need to return to method development.

If multiple assessments are conducted in a validation run, not all need to be acceptable for the run to be accepted. For example, the failure of dilution QC assessment samples in a run would not invalidate any other assessments which did meet criteria in that run.

SAMPLE ANALYSIS

System Equilibration

The topic of system equilibration has, in recent years, generated much discussion within the bioanalytical community. This discussion was primarily a result of inspectional observations. As a result of these inspections, it has become evident that regulatory agencies have clear expectations with respect to system equilibration activities.

System equilibration should never be conducted using unanalyzed study samples or a sample (LLOQ standard, low QC, etc.) that could possibly be substituted in an official run. Rather, pooled samples, generally from previous runs, or purpose-generated samples should be used; if purpose-generated samples are employed, they should be clearly identified and separated from any study samples, calibrators, or QC samples used in the official run.

Furthermore, it is recommended that objective criteria be used to assess whether or not a system has been adequately equilibrated. An example of a suitable criterion is an assessment of the precision of analyte and internal standard responses of sequential injections of an equilibration sample. Target precision should be specified in the assay protocol.

System equilibration is not viewed as an absolute requirement for an assay. However, if equilibration injections are employed, they must be documented in the study file.

System Suitability

System suitability testing should be designed to assess key performance aspects of the assay (sensitivity, resolution, etc.). The testing is generally conducted prior to a series of analytical runs of the same method. Testing methodology is assay dependent and should be defined during assay development. Purpose-generated samples independent from unknown study samples should be used for the testing. These samples are also independent from system equilibration samples. The test should be defined in the assay protocol along with acceptance criteria. System suitability is not generally a run acceptance criterion; however, appropriate documentation should be maintained if a run is initiated following a failed system suitability test. Finally, it is recommended that system suitability testing occur before resuming injections after an analytical interruption of a run or a batch reinjection following an instrument failure.

Individual Run Acceptance

For difficult assays or new technologies, wider acceptance criteria may be set with advanced planning, provided that an *a priori* statement of the revised criteria is appropriately documented. It is recommended that the acceptance limits be established taking into account data accumulated during validation activities. If all replicates of the LLOQ, ULOQ, or both calibrators at any other concentration fail to meet acceptance criteria and are excluded from the run regression, the criteria requiring at least 75% passing calibrators with a minimum of six concentrations must be observed. In this case, low-QC and high-QC sample concentrations must be between the revised LLOQ and/or ULOQ concentrations.

The LLOQ and ULOQ of an assay should be based on the nominal concentrations of the lowest and highest standards. These limits are not dependent upon the back calculated concentrations for the standards in sample analysis runs.

Internal Standard Assessment During Sample Analysis

Assessing internal standard response during sample analysis may point to issues with a given sample or run. Assessment should be conducted in accordance with predefined criteria as discussed earlier and must be conducted in a consistent manner for all runs of study samples analyzed by the same analytical method.

Dilution Verification Samples

Dilution QC samples should be included in sample analysis runs only if diluted study samples are also included in the run. Dilution QC samples are used only to evaluate the acceptability of samples diluted into the calibration range and are not used to accept or reject undiluted calibration standards, QCs, or undiluted sample results.

If more than one dilution factor is used for sample analysis in a single sample analysis run, dilution verification at the maximum dilution factor used within the run is

minimally sufficient to assess the acceptance of all dilution factors within the run.

These dilution QC samples should back calculate with a concentration within 15% of nominal. If the dilution QC samples fail, only results from diluted samples in the run need be rejected assuming that the undiluted QCs meet acceptance criteria.

Lowering the sample volume is not an acceptable substitute for dilutions; the volume of matrix processed for samples, QCs, and calibration standards should be identical to the volume used during the assay validation.

Carryover

Carryover refers to an instrumental response in excess of that directly associated with the analyte content of the injected sample caused by residual analyte in the injection path. Carryover is generally variable during a run in that it is dependent upon previous samples. Carryover is distinct from contamination. Contamination is an excess instrumental response caused by other factors such as sample collection and processing. Contamination may be in the sample when it is received at the bioanalytical laboratory.

It is required that sources of carryover be understood and carryover be minimized during assay development. Minimization of carryover may require an adjustment of assay range; that is, the ULOQ of the assay may need to be lowered or the LLOQ raised.

Traditionally, carryover is assessed by injecting a double-blank sample after a high standard. Carryover is considered acceptable if the analyte response in the blank is less than 20% of that of the lower limit of quantitation standard. This assessment approach may be problematic if the S/N of the LLOQ is at or below 5. In such a case, a carryover response of 20% of the LLOQ would be nondetectable as it could not be distinguished from background. For cases of low S/N, an alternative approach to carryover assessment should be employed. One alternative approach involves injecting a LLOQ standard followed by a high standard, followed by a second LLOQ standard. Carryover is considered acceptable if the responses of the two LLOQ standards differ by less than 20%.

Carryover should be assessed during each run. It is recommended that the assessment should minimally be performed twice, once at the beginning and once at the end of the run. The double assessment is recommended to detect carryover that may develop during the course of a run, for example, due to the depletion of needle wash solvent during the run. Failure of the carryover assessment does not lead to an automatic rejection of the run but rather an impact assessment should be performed. As long as samples containing high-analyte concentrations were not injected immediately before those containing low-analyte concentrations, the impact of carryover on analyte quantitation may be minimal. The impact of carryover may be minimized by injecting standards in order from low to high followed by samples in sequence of anticipated increasing concentration. Randomizing the sequence of samples in a run is not recommended as carryover impact may be significant if a highly concentrated sample is injected immediately preceding a low-concentration sample.

Implications of Contamination

A blank sample and a blank sample with internal standard (IS) are run in each batch of samples during method validation as well as in sample analysis. If the response at the retention time of analyte is more than 20% of the LLOQ response and the response at the retention time of the internal standard is more than 5% of IS response in the blank sample with IS, then the batch should be investigated (except when endogenous analytes are measured). Investigation procedures should be predefined in a SOP. Special bioanalytical procedures may be required if the investigation determines a laboratory contribution to blank contamination.

Implications of Anomalous Sample Results

When an anomalous assay result is observed, an investigation should be conducted. This investigation is required only for study sample analysis runs where the entire run is considered analytically valid and only the concentration value of a single (or a few) sample(s) is/are in question. Possible examples of anomalous results may include but are not limited to a single incurred sample reproducibility failure as predefined by SOP, positive predose or control concentrations, differences in IS response, chromatographic anomalies, or other major systemic problems. An investigation and remediation system may be a useful tool in managing investigations but is not required. A SOP should describe evaluation, investigation, and documentation of any anomalous results for bioanalytical studies.

Sample and Run Reinjection

If a batch halts during sample analysis, three possible actions can then be taken when analyzing unknown samples:

1. If in validation the complete batch reinjection meets acceptance criteria using the reinjected calibration curve, an entire batch may be reinjected during sample analysis after batch interruption. There is no need to reextract the samples.
2. If in validation the reinjected QCs meet acceptance criteria against the original calibration curve, a partial batch of incurred samples and QCs may be reinjected during sample analysis and quantitated against the originally injected calibration curve in that batch. There is no need to reinject the calibration standards.
3. If in validation the reinjected QCs and/or reinjected standards fail to meet acceptance criteria, no reinjections may be performed and complete reextraction and reanalysis is required during sample analysis.

Isolated samples should not be reinjected alone. Rather, they should be reinjected with standards and QCs as a new run.

In the event of instrument failure, the resulting repairs should be documented along with the steps taken to verify that the system was functional. QC samples injected with the reinjected samples are used to support the validity of data.

Failed Run Reporting—Sample Analysis

Sequential failed runs during sample analysis should be a trigger for an investigation during the analysis. A high (>33%) percentage of failed runs overall in a study should trigger investigation and acceptance of study data should depend upon the result of the investigation.

The study report should contain a listing of failed analytical runs by date. A reason for run failure should be included in the table.

Runs may be accepted if LLOQ or ULOQ standards fail as long as the QCs pass and are bracketed by standards. If the assay range is truncated by the removal of LLOQ and/or ULOQ standards, samples with reported concentrations below the lowest or above the highest acceptable standard should be flagged for repeat analysis. Extrapolation below the lowest or highest standards is not permitted.

Chromatographic failure of individual samples can be considered outliers and be flagged for repeat analysis. Chromatographic failure of QCs or standards is not an acceptable reason for excluding these samples from the calculation of general run acceptance criteria.

Audit Trail

The regulations regarding audit trailed activities as they relate to the predicate rules applicable to chromatographic run quality have been clearly defined by the World Health Organization (WHO). As per the WHO, an audit trail is a "Paper or electronic trail that gives a step by step documented history of a transaction. It enables an examiner to trace the data source document. The presence of a reliable and easy to follow audit trail in an indicator of good internal controls instituted by a firm, and forms the basis of objectivity." (12). Audit trails are not routinely examined; they are generally only reviewed during an investigation or inspection. This review may also be considered as part of an automated quality process. Regardless of their review frequency, it is recommended that audit trail clarity is improved by using predefined responses either as defined in SOP or via specific training.

REFERENCES

1. European Medicines Agency. Guideline on bioanalytical method validation 2011. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2011/08/WC500109686.pdf. Accessed 26 Jan 2014.
2. US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Veterinary Medicine (CVM), May 2001. http://www.fda.gov/downloads/Drugs/Guidances/ucm070107.pdf . Accessed 29 Jan 2014.
3. Dyson N. Chromatographic integration methods. 2nd ed. London: RSC Chromatography Monographs; 1998. doi:10.1039/9781847550514. 978-0-85404-510-5.
4. Meyer VR. Error sources in the determination of chromatographic peak ratios. In: Advances in chromatography. New York: Marcel Dekker; 1995. p. 383.
5. Hill HM, Smeraglia J, Brodie RR, Smith GT. Where do you draw the line? Points to consider when implementing the FDA method validation guidance on quantifying chromatographic peaks. Chromatographia. 2002;55(Suppl):S79–81.
6. Snyder LR, Kirkland JJ, Dolan JW. Introduction to modern liquid chromatography. 3rd ed. Hoboken: Wiley; 2010. doi:10.1002/9780470508183. 978-0-470-16754-0.
7. Garofolo F, Rocci ML, Dumont I, Martinez S, Lowes S, Woolf E, *et al*. 2011 White paper on recent issues in bioanalysis and regulatory findings from audits and inspections. Bioanalysis. 2011;3(18):2081–96.
8. Ministério da Saúde, Agência Nacional de Vigilância Sanitária, Brazil. RESOLUÇÃO - RDC Nº 27, DE 17 DE MAIO DE 2012. http://bvsms.saude.gov.br/bvs/saudelegis/anvisa/2012/rdc0027_17_05_2012.html . Accessed 29 Jan 2014.
9. Bergeron M, Bergeron A, Furtado M, Garofolo F. Impact of plasma and whole-blood anticoagulant counter ion choice on drug stability and matrix effects during bioanalysis. Bioanalysis. 2009;1(3):537–48.
10. Sennbro CJ, Knutsson M, van Amsterdam P, Timmerman P. Anticoagulant counter ion impact on bioanalytical LC-MS/MS assays: results from discussions and experiments within the European Bioanalysis Forum. Bioanalysis. 2011;3(21):2393–99.
11. Sennbro CJ, Knutsson M, Timmerman P, van Amsterdam P. Anticoagulant counter ion impact on bioanalytical LC-MS/MS assay performance: additional validation required? Bioanalysis. 2011;3(21):2389–91.
12. World Health Organization. Handbook: good laboratory practice (GLP): quality practices for regulated non-clinical research and development. 2nd ed. Geneva: WHO Special Programme for Research and Training in Tropical Diseases; 2009. 978-92-4-154755-0.