CHAPTER

Bioanalytical method validation and bioanalysis in regulated settings

16

Krzysztof Selinger*, Eliza N. Fung†, Peter Bryan‡

** Forest Research Institute, Farmingdale, NY, USA, † Bioanalytical Sciences, Bristol-Myers Squibb, Princeton, NJ, USA, ‡ Mendham, NJ, USA*

CHAPTER OUTLINE

Specification of Drug Substances and Products. http://dx.doi.org/10.1016/B978-0-08-098350-9.00016-3

16.1 INTRODUCTION—BRIEF HISTORY OF HARMONIZATION EFFORTS

The subject of bioanalytical method validation has been extensively debated for several years and a number of papers describing various approaches to validation have been published.[1–20] The need to develop and accept a uniform approach to method validation has been generally recognized in the pharmaceutical industry and regulatory agencies, which resulted in several meetings and guidance documents such as:

- Food and Drug Administration (FDA) 2001,[1] referred to here as Method Validation Guidelines (MVG).
- European Medicine Agency (EMA) 2011,[2] referred to here as Guideline for Bioanalytical Method Validation (GBMV).

Starting in 1990, scientists representing the industry, academia, and regulatory bodies met regularly at conferences devoted to harmonization and refinement of bioanalytical methods and validations. These meetings were taking place initially in Crystal City, VA, and hence the reports from conferences have been called Crystal City I, II, III, etc.[3–5] Numerous workshops and other initiatives have been undertaken by various professional organizations, e.g. American Association of Pharmaceutical Scientists (AAPS) and Society of Quality Assurance (SQA), to work out problems even though so far no consensus has been reached or alternative solutions proposed.[6–13] Other professional associations, e.g. Société Française des Sciences et Techniques Pharmaceutiques (SFSTP), also contributed to the harmonization efforts.[14–16] Finally, numerous independent papers from academia and industry have been published.[17–20] The growing globalization of the economy and the pharmaceutical industry gave more impetus to the harmonization efforts. To this effect, a Global Bioanalytical Consortium (GBC) has been created,[21–25] which includes scientists from different regulatory regions. The goal of the GBC

is to merge existing bioanalytical guidances and industrial procedures into a unified document and procedures acceptable to the regulatory authorities in various countries. The resulting document would ensure that the bioanalytical work is done to the highest scientific standards, and that the review process would be uniform and objective in all countries involved.

Additionally, the Global Contract Research Organization Council for Bioanalysis has been formed[26] with broadly similar goals. Their efforts are fully supported by professional organizations such as AAPS, European Bioanalysis Forum, Canadian LC–MS Group/Calibration and Validation Group, and Applied Pharmaceutical Analysis—Boston Society for Advanced Therapeutics.

This chapter describes validation procedures which are based on the two major regulatory documents[1,2] augmented by various meeting reports, White Papers, and other independent publications[27–34]; these will be referred to as consensus papers. The following logic has been applied:

- If MVG and GBMV agree on an issue, this common position is presented in the chapter.
- If there is a difference between the MVG and GBMV, the stricter and more comprehensive position is presented here.
- If one of the documents remains silent on an issue, the position expressed in the other guidance is presented here.
- If both documents remain silent on issues, which in the meantime have been discussed and agreed on in professional forums, the point of view expressed in white papers and workshop reports is presented here.

The reader must be advised that the FDA has announced the publication of a new bioanalytical guidance. It has not been published as of this writing in the Spring of 2013.

The aim of this chapter is to provide in Sections 16.2–16.8 detailed instructions on method validation, which would facilitate acceptance by the regulatory authorities in most developed and developing countries. Method development and validation issues for which an agreement has not yet been reached, or have been under the radar of the bioanalytical community, are discussed in Section 16.7.

Most of the comments are related to chemical (chromatographic) methods; Ligand-Binding Assays (LBAs) are not addressed. The scope is limited to bioavailability, bioequivalence, and pharmacokinetic studies, including also some aspects of toxicology and metabolism.

The term "bioanalytical method validation" can be understood in a narrow or in a broad sense; this chapter will cover the broad sense, which includes:

- Proper validation exercise
- Application of the meetings validated method to routine drug analysis, which includes:
  a. Proper execution of a bioanalytical project
  b. Control of a method during its execution

16.2 GLP, GCP, AND REGULATED BIOANALYSIS

The MVG stipulates that bioanalytical work in support of pharmacology/toxicology and other preclinical studies for regulatory submissions must be done according to good laboratory practice (GLP),[35] while the support for human bioavailability, bioequivalence, pharmacokinetic, and drug

interaction studies shall be done according to good clinical practice (GCP).[36] However, the GCP document-21 CFR 320.29 is very short and demands only that a method used to support bioequivalence and bioavailability studies must be accurate, precise, and sensitive. There is another document from the EMA that describes in much greater detail the basic requirements of GCP[37] and requirements on the part of laboratories performing analysis in support of clinical trials. In practice both types of work are done according to the same set of rules with addition of some aspects of cGMP (current Good Manufacturing Practice) and standard industrial practices. For these reasons, the term "regulated bioanalysis" will be used here as it includes the hybrid of GLP, GCP, and cGMP.

16.2.1 Bioanalytical method validation according to the current regulatory and industrial standard

There are several kinds of validation:

- Original validation—the method is validated for the first time, immediately after method development
- Partial validation—when changes are introduced to a validated method, such as new personnel or equipment is used, or other minor changes in the method
- Crossvalidation—when a comparison is made between two laboratories or two procedures

These types of validation will be discussed in detail in Section 16.6. It is worthwhile to point out that the term "qualification" is frequently used. It can refer to a partial validation or a "fit-for-purpose" validation in which only selected parameters, such as accuracy and precision, are evaluated based on the intended use of the method. It has to be understood that the process of method development and validation is a continuum, and that method development and validation has a certain life cycle.

In the current highly cost-conscious pharmaceutical environment the balance of costs and benefits is a critical issue. New drug candidates, frequently from the same chemotype, are screened and tested and it would be wasteful to undertake a full validation for all of them. Hence, some initial pilot toxicology projects can be done using a method that has not been yet fully validated. Usually at the drug discovery stage doses administered to laboratory animals are rather high, enabling simple approaches, e.g. protein precipitation as sample pretreatment and a generic reversed-phase liquid chromatography (LC) method with liquid chromatography–tandem mass spectrometry (LC–MS/MS) detection. This is not to say that such a method is without scientific merit and does not absolve the analyst from understanding the chemistry of the compounds investigated. Testing at least some rudimentary level is needed, such as stability (is the compound stable for one day or one week that is needed to complete this pilot project?) or extraction efficiency (is recovery at least 20–30%?). At this stage a one-day testing procedure consisting of a single or duplicate calibration curve and a set of quality controls (QCs) is sufficient.

A project leaving the discovery stage and entering the development stage requires the regulated-bioanalysis treatment and a fully validated analytical method. Formal toxicology studies are pivotal, these include toxicological studies, chronic toxicology, and toxicokinetics. Although the ADME studies (absorption, distribution, metabolism, and elimination) of the parent drug do not formally require a validated method, such a method usually exists at this rather late stage of drug discovery. Many bioanalytical laboratories perform such studies as per regulated bioanalysis. All human studies submitted to regulatory authorities require a fully validated method.

Every laboratory needs to develop a validation protocol, which describes specific elements of validation in that particular laboratory and its acceptance criteria; this usually takes the form of a Standard Operation Procedure (SOP), a protocol, or a validation plan. The validation exercise has to be auditable and should be summarized in a validation report. A validation report may contain a detailed description of the analytical method, or the method can be described in a separate document (SOP, test method, or a method sheet).

16.3 METHOD DEVELOPMENT AND INITIAL VALIDATION OF THE CHEMICAL METHOD

Validation follows methods development, transfer, or modification, and is followed by a method application, i.e. a bioanalytical project. While validation logically follows the method development, the method application does not necessarily directly follow the validation. Continuity between these two elements is needed which means that if a method is not used on a regular basis it needs to be validated or qualified again before use; such a validation protocol may be abbreviated and limited to a single run. There is no set limit on how much time is allowed to elapse between testing occasions. The project manager of the study will need to use his or her judgment based on the circumstances. In the bioanalytical community, such a run is called a prestudy assay evaluation, a qualification, or even a validation run. Conversely, a method cannot be considered fully validated until it is applied to real clinical or animal samples with good reproducibility. To verify this attribute, an exercise called Incurred Sample Reassay or Reanalysis (ISR) is performed and will be discussed in Section 16.4.5.

A bioanalytical method cannot be developed and validated without considering the ultimate objective of a study. This means that during method development one has to keep in mind basic requirements and numerous details which may have a bearing on the project: range of concentrations needed and targeted lower limit of quantitation (LLOQ), matrix to be used, anticoagulants in the case of blood or plasma (ethylenediaminetetraacetic acid (EDTA)—at what concentration? which salts? heparin—lithium, sodium or potassium? any additives such as citrate, oxalate), volume of blood needed or available per assay (pediatric studies provide small sample volumes, usually <0.5 mL; the same limitation applies to small laboratory animals such as mice), stability, and safety considerations. The stability considerations are of utmost importance. For new chemical entities (NCEs), the stability data are frequently not available and neither is the method of analysis. One has to develop a skeleton of an assay and immediately employ it to evaluate stability in the sample matrix under different conditions (e.g. storage at room temperature, multiple freeze–thaw cycles, or at -20 °C). Once any instability issues are resolved, one has to ensure that appropriate sample collection and storage procedures are used in the clinic or animal rooms.

After a period of trials and tests, a bioanalytical method will be ready for validation. An analyst has to make sure that method development is complete and no additional changes will be introduced during validation; some examples of changes are provided in MVG.[1] Introduction of changes and modifications require starting the validation procedure from the beginning.

The essential parameters that need to be defined to ensure the acceptability of a bioanalytical method as per MVG are precision, accuracy, sensitivity, specificity, response function, recovery, reproducibility, and stability. The GBMV also demands evaluation of carryover, dilution, integrity, and matrix effects.

16.3.1 Precision and accuracy

The definitions of accuracy and precision are presented in Chapter 4. The goal of this chapter is to describe all necessary steps leading to precision and accuracy appropriate for trace analysis in biological matrices. The majority of analytical measurements are relative in nature, which means that a result is obtained by comparison of sample response with a standard response. Hence, the quest for accuracy begins with a reference standard.

The best reference standard for well-established and easily available drugs should be a compendial material (US pharmacopeia (USP), British Pharmacopoeia, European Pharmacopoeia), although USP does not provide a certificate of analysis (CoA). Standards from reputable commercial suppliers with CoAs are also acceptable. NCEs are available only from their originators, who should also provide a CoA. Nevertheless, it is the responsibility of the user to obtain the CoA. Ultimately, if no CoA is available, the user has to prepare a CoA by performing a number of tests confirming the identity and purity of the standard; these tests may be spectral and elemental analysis, high-performance liquid chromatography (HPLC) area summation, water and ash contents (see Chapter 2 for more details). One has to limit testing to a necessary minimum as many compounds are in short supply. It is worthwhile to consider using quantitative nuclear magnetic resonance spectroscopy as one-step analysis for purity. The use of drug formulations (tablets, injections) as a source of analytical standards is strongly discouraged. If a reference standard is difficult to obtain or purchase, and a drug formulation is the only source of the material, then such a secondary standard should be also characterized analytically by other techniques. Although a CoA is not necessary for the internal standards (ISs), it is obvious that the user must make sure that the potential IS is "fit for purpose". At the very least the analyst must ascertain that the IS does not contain any significant amount of the analyte of interest, or produce it by chemical reaction during the analytical procedure.

The next step is to ensure the correctness of calculations. Many substances exist in the form of salts and/or hydrates or solvates at different degrees of purity. Measurements in biological matrices should provide a result expressed in terms of a free base or acid. The calculations should be verified by a second analyst.

Weighing of the reference standard should be carried out on a properly maintained and currently calibrated analytical balance. For a typical weighing of 1–10 mg of material, sensitivity of at least 0.00001 g is needed. In order to avoid bias in the analytical technique, ideally two analysts should prepare a total of at least two weighings and two separate stock solutions. The master stock solutions should preferably be prepared in volumetric flasks of appropriate volume, or by accurately pipetting the required amount of solvent. After dilutions by the primary and secondary analysts, these stock solutions should be compared using an appropriate analytical technique; the technique does not have to be the same as in the final bioanalytical method. The acceptance criteria for stock/spiking solutions should be specified *a priori* in the laboratory's SOP; in general, stock solutions should be within 5% of one another.

The equation to calculate these differences is as follows:

$$(x1 - x2)100\%/\text{mean} \tag{16.1}$$

Should the differences be greater than the acceptable limit, records of preparation should be reviewed. New stock solutions might be prepared either by preparing a new stock solution or dilutions of existing stock solutions reprepared and compared. An outlier must be ultimately identified and eliminated.

Having made sure that the stock solutions accurately reflect the analyte concentration, one of the two stock solutions should be chosen to spike both the calibration standard and QCs. This approach has been endorsed by the Crystal City conference report,[5] although several laboratories in recent years received FDA 483 citations for following this advice.[34] While it is common in the industry to use separate stock solutions for the calibration standards and for the QCs, it is the opinion of the authors that, having verified the correctness of stock solutions prepared as described above, it is no longer necessary to use two sets of stock solutions because the use of separate stock solutions may introduce bias.

Spiking of calibration standards and QCs should involve a primary and secondary analyst. The primary analyst should prepare the calibration curve while the secondary analyst should prepare QC samples, or a combination thereof. Again, the secondary analyst is needed to make sure that a slightly different analytical (i.e. pipetting) technique would lead to the same results. Although the spiking into biological matrices appears trivial, it provides an ample opportunity for the introduction of errors. Here are some examples.

First, one may use either volumetric flasks (volumes restricted to certain values, sometimes difficult to mix efficiently) or deliver volumes using a pipette as biological matrices tend to be precious and should not be wasted; hence the smallest necessary volume should be used. Second, blood, plasma, and serum are viscous and relatively difficult to measure accurately. Third, blood, serum, plasma, and urine tend to foam while mixing; this makes the volume control difficult as the meniscus can be difficult to see. Fourth, stored plasma may contain precipitated proteins and fibrins, while stored urine may contain precipitated salts; both may block pipettes.

Pipettes (electronic or manual) used for dilutions should be maintained and calibrated according to the laboratory SOP. Usually delivering 98–102% of the nominal volume and with $<1\%$ coefficient of variation (CV) is acceptable for automated pipetting devices; calibration or performance verification should be performed every 3–6 months or more frequently when required (after maintenance or repair).

There are two schools of thought for the preparation of calibration standards. The first requires spiking of small volumes of standards on each analytical day using freshly prepared or diluted spiking solutions: such standards are not stored and are used on the day of preparation; this approach seems to be favored by the Crystal City I Conference Report I.[3] The second approach permits the preparation of standards in bulk, aliquoting them in separate tubes and storing under the same conditions as the QCs and study samples. In both situations, QCs are spiked in bulk, aliquoted, and stored with the study samples. The justification of the first case is that by always using new standards, the sample stability is monitored, and that calibration standards are distinct from the QCs. The justification for the second case is that this procedure is acceptable if stability data exist to support this approach. It is easier, more productive, and avoids an additional potential for bias to spike once only. Additionally, the difference between calibration standards and QCs is that a calibration curve is forced through the standards, but not the QCs. The authors of this chapter prefer the second approach.

The calibration standards should be prepared in the same matrix as the samples (whenever possible), and separate calibration curves should be established for each analyte. It is recommended to have six to eight nonzero standards for each calibration curve, with single or replicate samples, in each analytical batch. More standards are needed for complex, nonlinear calibration models. One calibration standard should be at the LLOQ; it is good practice to have the second standard at $2 \times$ LLOQ to define the lower end of the curve well, and all the other standards spread over the remaining range of

concentrations. It is also good practice to have two levels of standards between the concentration of the high QC sample and the upper limit of quantitation (ULOQ). It is recommended to have three standards per order of magnitude, i.e. a calibration curve over 1–200 ng/mL should have six standards, but a calibration curve over the range 0.1–100 ng/mL may require nine standards. Drug-free matrix (blank) and drug-free matrix with the IS added (standard zero) should be a part of every analytical batch. Calibration standards can be placed either all in the beginning of the run or dispersed throughout the batch. A run should start with system suitability samples. These are prepared by mixing the analytes of interest in the mobile phase or reconstitution solution at certain proportions representing either typical concentrations seen after the extraction, or the concentrations as observed in an LLOQ sample. The system suitability samples serve to verify the retention times, resolution (if critical), sensitivity, and stability of the system. The system suitability samples should meet predetermined acceptance criteria.

One of the issues frequently discussed is whether a calibration curve should be measured once or performed in duplicate in each analytical run. Again, a balance of costs and benefits is necessary. For robust assays with a stable response of the system, a single calibration curve may suffice. If a significant drift in the instrument's response is observed, the second calibration curve placed toward the end of the batch may be needed.

Quality controls should be prepared with at least three concentration levels and analyzed in duplicate with each analytical batch. One set of QCs should be close to the LLOQ (2–3 times higher than the LLOQ). The second set should be at approximately 40–60% of the ULOQ, and the third at 70–90% of the ULOQ. Quality controls should be spread evenly throughout the analytical batch. The selection of the calibration curve range and placement of the QCs must be judicious to represent the concentrations of samples encountered in the study. In other words, clinical or animal sample concentrations should not be clustered around the lower or upper quadrant of the calibration curve only, with one QC that falls within the concentration range of most study samples. Of course, this may not be known at the time of validation, but discovered only after first batches of a study have been analyzed. In this case, addition of the extra QCs of appropriate concentration are suggested.[2]

Assay accuracy is expressed as a percentage of the true value which is calculated according to the formula:

$$\text{Accuracy} = \text{observed value} \times 100\%/\text{true(nominal)value} \tag{16.2}$$

The term "recovery" is sometimes used to describe accuracy; this usage should be discouraged. The true value is assumed to be the nominal value at which the sample has been spiked; the accuracy can be expressed as a percentage of nominal or percent deviation from nominal (i.e. bias). The practice of using the observed values instead of the nominal ones is no longer considered acceptable.[1]

The use of samples from dosed subjects to assess accuracy has been suggested.[3] The rationale is that despite our best efforts, it is virtually impossible to mimic a clinical sample by the simple addition of a standard solution to appropriate medium, as a clinical sample may contain drug metabolites, concomitant medication and its metabolites, as well as endogenous substances, the level of which may be impacted upon by the drug administration. Hence, a pool of clinical samples could be used as an extra QC sample. Such a solution is generally impractical for the reason of availability, yet may be very useful in some situations. For example, drug conjugates (glucuronide and sulfate) are notoriously difficult to obtain. A solution to this problem is to hydrolyze chemically or enzymatically these conjugates and measure the concentration of the free drug. However, a hydrolysis control is needed to make sure that such a process remains reproducible; a pooled subject sample could play this role.

Another example is the possible conversion of unstable metabolites such as N-oxides back to the parent drug during the analytical process.

Precision is a measure of repeatability of a method and can be expressed by the relative standard deviation (RSD); this value is commonly known as the coefficient of variation (CV):

$$\mathrm{RSD} = \mathrm{SD} * 100\%/\mathrm{mean} \tag{16.3}$$

Here the standard deviation (SD) is calculated as per Eqn (2.2).

In today's bioanalytical assays, a CV of $\leq$5% characterizes a very precise method. A CV of 5–10% is probably the most common and represents an industrial norm in terms of precision, while a CV around 15% suggests either a method of extreme difficulty and unusually low LLOQ or some analytical problems; a CV of 20% may be acceptable only around the LLOQ.

The ultimate goal of any method is to assay samples; QCs are the best approximation of clinical samples. Hence, accuracy and precision of a method should be estimated using the percent nominal and CVs calculated for the QCs, and not the back calculated (interpolated) value of calibration standards, although these values should also be reported. A calibration curve is forced through the calibration points, and accuracy and precision based on standards always look somewhat better than those based on QCs. Back calculated standard concentrations are a useful and necessary tool in the evaluation and adherence of the system to the selected mathematical model.

Within-run precision and accuracy are evaluated during the validation by assaying a minimum of five replicate samples independent of standards at concentrations representative for the assay; a separate set of QCs could be used for that purpose.

16.3.2 Limits of quantitation

The LLOQ is frequently confused with the limit of detection (LOD). Both are a measure of the sensitivity. The LLOQ is the lowest concentration of the standard in the calibration curve and is higher than the LOD. The LOD has no practical use in regulated bioanalysis. Various ways of calculating the LLOQ are presented in Chapter 2. The consensus documents recommend a very pragmatic approach to the determination of LLOQ. It is the concentration which provides a CV $\leq$20%, and accuracy between 80% and 120%. The way to establish this experimentally is to prepare at least five samples independent of the standards at the concentration of the projected LLOQ, another set of five at concentration 2 $\times$ LLOQ, one more at 4 $\times$ LLOQ, and so on. These samples should be analyzed with a calibration curve. The concentration, which fits into specification, should be considered the LLOQ, and the lowest calibration standard should be set at this value. The conference also endorsed other approaches to LLOQ, and alternative models of LLOQ are presented in references.[38]

The MVG and GBMV specify that the minimal signal-to-noise ratio should be 5:1, which is a pragmatic recognition of the fact that it is nearly impossible to obtain acceptable precision and accuracy if the signal-to-noise ratio is <5:1. The practical way of calculating the signal-to-noise ratio is presented in Fig. 2.22 of the first version of this book. Many modern analytical computer programs are available to calculate this parameter. Most of the scientists working in the area of biological trace analysis are under constant pressure to improve sensitivity and lower the LLOQ; the question "Can you get lower than this?" is proverbial. As a practical rule of thumb, most bioanalytical assays only need to have sufficient sensitivity to quantify the main analyte of interest at concentrations estimated after five biological half-lives to adequately characterize a compound's pharmacokinetic profile. A chemical or

instrumental breakthrough answers today's questions, and more insight into the nature of things invites more questions.

It should be noted that there is not only an LLOQ but also an ULOQ, which is the highest concentration of the standard in the calibration curve. There are several reasons for the existence of the ULOQ: above a certain concentration a calibration curve may no longer be described by a chosen mathematical model (this usually means plateauing), large chromatographic peaks may be truncated if a detector is saturated, chromatographic peaks can be deformed by overloading of the system, or a method simply has not been validated above a certain concentration. How to handle results that are above the ULOQ (above the Upper Limit of Quantitation) is explained in Section 16.4.3.

16.3.3 Specificity

There are two components of specificity. First, a bioanalyst must prove that other components within the matrix do not generate (or contribute significantly to) the measured signal; second, that the signal (chromatographic peak) is indeed generated by the analyte of interest. In chromatographic methods with detectors other than MS, an analyte is identified solely on the basis of its retention time, wavelength in ultraviolet (UV) detectors, combination of excitation and emission wavelengths in fluorescence detectors or applied potentials in electrochemical detectors. These techniques are inherently nonspecific and can only limit the number of compounds which may be seen otherwise in the same time window. Only LC–MS/MS and to some lesser degree also LC–MS provide virtual certainty that the signal observed was generated by the analyte of interest. Extra caution should be exercised when developing methods for compounds with glucuronide and N-oxide metabolites. These metabolites are often difficult to obtain during early phase development and are prone to in-source fragmentation or may not be stable in the biological matrix. Chromatographic separation of these metabolites is required to ensure assay specificity. A good example of a case where chromatographic separation of a potential interfering compound is for analytes containing a glutarimide ring (e.g. thalidomide). The glutarimide ring is prone to hydrolysis and this degradation product often coelutes with the analyte and can be isobaric with the analyte. This problem can easily be solved by the addition of acid to the plasma samples, by only using fresh solvent for the preparation of stock solutions and the addition of dilute acid to the diluted working solutions.

Another type of specificity problem can be caused by the choice of a stable isotope labeled-IS (SIL-IS) for LC–MS or LC–MS/MS assays. Deuterated ISs can be prone to proton exchange with the solvent and need to have a sufficient number of substitutions to ensure there will be no interference in the analyte channel. Typically, a mass increase of four is sufficient to alleviate interference between the analyte and the IS for molecules containing only C, N, and O. The actual interference in selected reaction monitoring (SRM) can be calculated accurately based on chemical structures of the precursor and product ions, labeling positions, and concentration ratio.[39]

Six samples of the drug-free matrix obtained from six individuals should be used to prove lack of significant interference with the intended analyte. These blanks must be obtained from the relevant population, in the simplest case being split equally between the genders. It is becoming a common practice to evaluate hemolyzed and lipidemic plasma either as a part of or in addition to the six lots. Matrices from special populations, such as pediatric, may be evaluated at the discretion of the bioanalyst. The same biological matrix should be used for validation as that in the clinical/animal samples. In the case of blood, plasma, serum, or urine from humans or large animals, the matrix

availability does not present a problem. Some matrices, for example control cerebral–spinal fluid (CSF), bone marrow, sputum, bile, or samples from small animals may not be available in sufficient volumes or not at all; a surrogate matrix can be used instead for calibration standards and QC preparation.

A couple of issues require elaboration. There is no such thing as "no peak". If one amplifies electronically the baseline in the area of interest then oscillations of the baseline and minor spikes will become visible and in most cases there is something that could be integrated. The issue is how significant the contribution of the interference is allowed to be. The absolute minimum is that the interference should be $<20\%$ of the peak corresponding to the LLOQ. This requirement has to be specified in the appropriate SOP.

On the one hand, some projects involve dosing healthy volunteers whose diet is controlled and who provide samples that are relatively free of interference. On the other hand, in phases II and III of drug development or in oncology programs in which a drug is administered to patients who routinely take concomitant medications and whose general health condition may be poor, concomitant medications can be present in the samples for a variety of reasons and interference check is required. It is also prudent and in fact necessary to test whether or not the common over-the-counter (OTC) medications are present in the samples, even for normal healthy volunteers. A quick glance at the OTC shelves in a local pharmacy would indicate that these could be:

- Common pain killers (aspirin, ibuprofen, acetaminophen)
- Antiacids (cimetidine, ranitidine, famotidine, omeprazole, etc.)
- Antihistamines (loratadine, diphenhydramine, brompheniramine, chlorpheniramine, cetirizine)
- Components of cold medicines (pseudoephedrine, dextromethorphan)
- Caffeine

The situation may be much more complicated as one has to consider not only the parent drugs but also their metabolites, which can be numerous and difficult to obtain.

In the present era, when bioanalysis of small molecules means almost exclusively analysis by LC–MS/MS, the danger of interfering peaks originating from OTC medication is minimal. However, there is a danger of pharmacokinetic interactions between the tested drug and an OTC medication, cimetidine, for example.

In the simplest case, a bioanalytical laboratory could prepare a test mixture for evaluating the interference from OTC medication by mixing the easily obtainable reference standards and at least some of the metabolites at typical concentrations, such as C_{max}.

In order to prove that the substances being quantified are the analytes of interest in assays other than LC–MS/MS, one has to inject a reference solution containing only the intended analytes (drug, and/or metabolite, IS) in pure solutions at the beginning and the end of the run, and compare the obtained retention times with the biological samples. The best solvent is the mobile phase. A solvent in LC stronger than the mobile phase (e.g. methanolic solutions injected at volumes >10–20 μL into a typical reverse phase system) will produce a distorted peak with a shorter retention time. In gas chromatography, injection of simple reference solutions may not provide peaks at all if there are active sites in the system which adsorb analytes. The solution to this is to mix a blank extract with a reference solution or to include a "carrier" substance in large excess in the reference solutions, which would not interfere with the assay, but saturate the active sites. If the sample preparation involves a back extraction from diethyl ether, ethyl acetate, or other relatively water-soluble organic liquids, the

solvent in the reference solution should also be saturated with this reagent; otherwise the retention times will differ.

One must realize that today's supersensitive tandem mass spectrometers and even more sensitive detectors of tomorrow will make obtaining of the so-called "clean blanks" more and more difficult as it is very easy to introduce contamination to the samples. For laboratories that support development of a drug for a number of years, it is safe to assume that most of the laboratory benches, laboratory glassware, and LC–MS/MS systems come into contact with the analyte, and perhaps at high concentrations. Of course, the assays that are most prone to contamination are the ones with the LLOQ in the picograms per milliliter range. If there are persistent interfering peaks in the blanks or at lower concentrations, one would be well advised to follow precautions that are typical for the activities in a clean room. The bioanalyst should consider the following practical pieces of advice:

- Devote a separate room to this assay only, if possible
- If not possible, devote a corner of the laboratory to the assay and limit traffic there
- Start every day with a fresh laboratory coat or a disposable laboratory coat
- Clean the bench top with methanol every day or cover it with a paper mat
- Wipe the pipettes with methanol swabs frequently
- Change gloves and wash hands frequently
- Use disposable containers
- If glass volumetric flasks are needed, wash them personally with acid, water, and methanol
- If evaporation is involved in the processing, clean the jets of the evaporator with methanol
- Have devoted office supplies (pens, scissors) to the assay
- Never open sample tubes with stock solutions near the working space; if there is temperature difference a spray can be formed

16.3.4 Recovery

Recovery as defined here is the extraction efficiency. Only very uncomplicated samples in terms of concentration and matrix can be injected directly or after dilution onto an HPLC column; these could be CSF, urine, and saliva. Most biological samples have to be prepared in some way before entering a chromatographic column, on or off-line system. There is no formal requirement regarding how high recovery should be. A bioanalyst always tries to develop a method with recovery as close to 100% as possible. However, a recovery of $\leq$50% is also acceptable if it provides precise and accurate results and it is the best that can be achieved under the circumstances. Low recoveries are frequently associated with poor reproducibility, and are red flags for an analyst or a reviewer to watch for unexpected problems or outliers. A bioanalyst is often forced to work at the sensitivity limit of the system, and can hardly afford poor extraction efficiency, which in turn will decrease the sensitivity of the assay.

The absolute recovery can be calculated by comparing the peak areas (or peak area ratio of analyte/IS) from the extracted QCs (or calibration standards can be used) at three concentration levels in triplicate with those of unextracted solutions. The unextracted solutions should have the same concentration as those of the extracted QCs (or calibration standards). The unextracted solutions can be prepared by diluting neat solutions to concentrations representing 100% recovery with the extracted blank matrix, incorporating all the material losses due to the volume transfer. During the recovery study, all the volume transfers should be done quantitatively unless an IS is used. Another way of establishing recovery is to divide the slope of the extracted calibration curve by the slope of

unextracted standard curve representing 100% recovery; this approach applies only to rectilinear calibration curves.

16.3.5 Response function

The theoretical background of the establishment of the appropriate response function between the measured signal and the analyte concentration is described in Section 2.2.2 of the earlier version of this book. In general, the simplest response function should be selected, the fit should be statistically tested, and an appropriate algorithm or graph presented. What it means in practical terms is that during validation and/or study, the response function selected should remain the same, and not be changed from one batch to another.

Another issue related to the response function is the question of whether to use peak height or peak area. From a theoretical standpoint, only the peak area is proportional to the mass of the analyte, and the peak height is related to the mass only at the height of a triangle, which approximates an ideal peak. In today's highly computerized environment, measuring a peak area is very simple, and it should be used unless there are special circumstances, such as the presence of many overlapping peaks or severe tailing, where minor imperfections of the integration may cause significant errors. The alternative use of peak height is worth investigating at the prevalidation stage in order to find an optimal response function and range of calibration curve. The ultimate decision on using peak area or peak height, choice of regression, weighing factors, and linearity assessment belongs to the analyst and should be carefully documented.

16.3.5.1 Acceptance criteria

The validation is accepted or rejected as a whole with expectations that the overall precision and accuracy will be $\leq$15% at concentrations above the LLOQ, and $\leq$20% at the LLOQ.

16.3.6 Reproducibility

The methods used in regulated bioanalysis must be reproducible and repeatable. Briefly, repeatability is precision achieved in the same laboratory by the same operator using the same equipment, while reproducibility is precision in different laboratories by different operators.[40] Validation should be performed using a similar number of samples per batch, as in the study (so-called "run-size evaluation"). In the past, this number used to be considered as being close to 100. Today a discrete analytical run may consist of several 96-well plates as long as each plate contains its own set of six QCs. With typical analysis times of 2–4 min per sample, and some time devoted to the injection process which is highly dependent on the autosampler, one can expect the analysis of 96 samples to take 4–6 h. The rationale for run-size evaluation is to make sure that appropriate precision and accuracy are obtained by an operator (human or robot) challenged with a large number of samples, as well as to see if the system (chromatographic, robotic) performs correctly over the period of time needed to complete an analytical run.

There are no acceptance criteria for individual runs in the validation process. The validation exercise should be limited to a certain number of runs. It is the opinion of the authors that if three acceptable runs are needed to complete validation, no more than five attempts should be allowed. Should the fifth attempt fail to provide three acceptable runs, the method should be sent back for

further development, poorly defined parameters identified and optimized, and the validation should start from the beginning. Otherwise, a method would be a kind of game of chance rather than a science-driven process. Regulatory authorities consider this approach as "testing into compliance".

The bioanalyst always needs to maintain a balance of costs and benefits depending on the physicochemical properties of the analyte, concentrations required, and time considerations. The simplest solutions are quick, but do not necessarily provide the most robust methods. The order of extraction techniques from biological matrices according to increasing difficulty and time consumption may be as follows: direct injection, protein precipitation, single liquid–liquid extraction (LLE), simple solid phase extraction (SPE), LLE followed by back extraction, but the order in terms of chromatographic system robustness will be reversed.

Not much is said in MVG and GBMV about the ruggedness or robustness of a method, which is an important parameter and is discussed in Section 16.7.11.

16.3.7 Stability

The most common reasons for instability of drugs in biological matrices are chemical, enzymatic, and photochemical processes.[41] The chemical processes include hydrolysis of esters (diltiazam, aspirin), opening of the lactam ring in β-lactam antibiotics, opening of the lactone ring in campthotecin analogs, oxidation of phenols and naphthols, oxidation, dimerization, and side reactions of captopril, etc. The enzymatic processes include hydrolysis of esters such as procaine, esmolol, irinotecan, and remifentanil by esterases. The light sensitivity affects drugs such as nitrofurantoin, clomiphene, retinoids, and fluoroquinolones.

Stability has to be tested in the matrix of the study, under conditions encountered during the execution of a bioanalytical study. This includes the matrix in the presence of analytes of interest and/or their major metabolites which could potentially convert back to the parent drug, and includes bench top (processing) stability, freeze–thaw stability, on-instrument/autosampler stability, processed samples stability, and long-term storage stability. If a metabolite is included in the assay, then it should be included in the stability evaluation. It is also necessary to establish the stability of the stock and working solutions of the analytes and their ISs under the storage conditions.

Benchtop stability tests simulate situations during the sample collection and analytical processing, where samples typically remain at room temperature for a total of 3–6 h. At the end of that period, stability samples should be analyzed against freshly prepared calibration standards. In practice, one set of QCs is stored at room temperature (or ice bath, if needed) for 3–6 h, another set to 24 h, and then they are extracted along with freshly prepared calibration standards. Samples from HIV-positive patients are deactivated by heating at 56 °C for 3–5 h; hence stability at 56 °C should be also included in the validation exercise, if applicable.

The freeze–thaw stability test mimics the situation where samples undergo multiple freezing and thawing cycles either during sample collection, processing, or repeat analysis. The consensus is to subject the samples to at least three freeze–thaw cycles. These cycles should be at least 12 h apart with one cycle frozen for 24 h, if they are to simulate real-life situations. After the third thawing, the samples should be analyzed against freshly prepared calibration standards. The number of cycles should be adequate to cover the actual situation encountered in the study. The freeze–thaw test should be also performed with a dilution QCs (see section 16.3.9) in addition to the conventional low and high concentration QC samples to ensure that analyte precipitation at very high concentrations is not encountered.

The long-term storage stability test should be performed over a period of time that equals or exceeds the time between the date of sample collection and the date of analysis. Further evaluations may be made at later times; for example, 6 and 12 months of storage at the same storage temperature and in the same containers (geometry, caps) as the study samples. The stored stability samples are compared to freshly prepared calibration standards. What constitutes "fresh" is a matter of heated debate right now.[34] In an ideal situation, new calibration standards and new QCs would be prepared over 1–2 h, extracted, and injected to verify the correctness of the preparation (3–6 h), and then the new set of calibration standards and QCs would be extracted with the stored stability QC samples. From the author's most recent experience, it seems that the FDA insists on the definition of a "fresh standard" as that prepared and used on the same day. Bioanalysts would be well advised to complete the stability evaluation on the same day to avoid potential citation and Form 483, even though the logistics and practical aspects of the process may be challenging. On the other hand, even 24 h may be too long for unstable analytes, and other approaches are needed. In order to speed up the process, the bioanalyst can use the existing stock solutions, if their stability is known, or eliminate the verification of the new calibration standards and QCs.

The MVG[1] advises to use two concentration levels—high and low—in triplicate to evaluate stability, and the regular acceptance criteria of 15% applies.

Even if instability of a drug in the sample matrix is observed, the bioanalyst can take appropriate means to ensure sample integrity. For example, an antioxidant such as ascorbic acid or bisulfite can be added to avoid oxidation, or the pH can be lowered by the addition of citric acid to avoid hydrolysis of esters or lactam ring opening. Esterases can be inactivated by the addition of esterase inhibitors such as fluorides, physostigmine, or dichlorvos.[42] Addition of a derivatizing reagent can yield a stable entity. Lowering of temperatures is a good way to slow down degradation. Samples immediately after collection can be either flash frozen or kept in icy water and a refrigerated centrifuge can be used to harvest plasma. During sample processing, samples may be kept at 0–4 °C and processed quickly. Refrigerated autosamplers are readily available to ensure stability of extracted samples. If samples cannot be stored for any period of time, then samples may need to be analyzed immediately at the clinical site.[43] If the freeze–thaw stability is the problem, the samples can be divided into a number of aliquots at the clinical site, and reassays must be done using only separate aliquots.

An unusual case of instability is the situation in which a metabolite or degradant can convert back to the parent drug by undergoing a chemical reaction during the analytical process.[44] In this particular case, the instability may be detected as an increase in the parent drug concentration with repeat analysis.

16.3.8 Carryover

Carryover should be minimized and evaluated by injecting blank samples after high concentration calibration standards. Carryover should not be >20% of the LLOQ and 5% of the IS. If carryover is significant, then the analytical batch shall be organized in such a way that the carryover would be minimal and not significantly bias the results.[5] It should be stressed that carryover is highly compound and system specific. Some analytes tend to adhere to metal or polymer elements of the system and they may be difficult to eliminate. Sometimes an autosampler of a different design may provide carryover-free injections. Unfortunately, the carryover also depends on the maintenance condition of an autosampler and its history of use. Routine maintenance of the autosampler and components on the flow path such as replacement of worn-out components (e.g. injector syringe), polyether ether ketone

tubing, old columns (or guard columns if they are used) will help to keep carryover in check. It should be stressed that carryover is different from contamination. Carryover occurs after injection of a highly concentrated sample, while contamination is usually random in nature, not necessarily following a highly concentrated sample. Bioanalysts should carefully evaluate the situation and tackle the problem accordingly.

16.3.9 Dilution integrity

The process of diluting samples should not impact precision or accuracy. During the validation, the analyst should prepare an "Above the Upper Limit of Quantitation" Quality Control (sometimes called dilution QC) sample, dilute it to the expected concentration (e.g. by a factor of 10) and extract the sample five times. The acceptance criteria should be 15% both for accuracy and precision. Dilution should be done using the same matrix as the sample, although not necessarily from the same individual. If the dilution factor changes or study samples are at concentrations above that of the dilution QC, a new dilution experiment is needed.

16.3.10 Matrix effect

The matrix effect is a phenomenon where the signal of an MS/MS detector is different in the presence of coeluting components of the sample as compared to the neat sample. The signal can be increased or decreased and it is caused by changes in the efficiency of ionization and droplet formation in the MS source.[45] Coeluting interferences known to cause matrix effects include salts, phospholipids, additives (e.g. esterase inhibitors), metabolites, or other endogenous components. Phospholipids are a class of lipids that are commonly found in the blood and plasma. Two common structural classes of phospholipids are glycerophospholipids and sphingomyelins. Extensive work has been reported in the literature on the matrix effect caused by phospholipids.[46–48] Phospholipids can be monitored easily by using these transitions: positive precursor ion scan of m/z 184, positive neutral loss scan of 141 Da, and negative precursor ion scan of m/z 153. In general, protein precipitation is not effective in eliminating phospholipids since they tend to remain in the organic supernatant. A number of methodologies have been devised to eliminate phospholipids, such as the use of mixed-mode SPE, LLE, and the addition of selective trivalent cations.[49]

While it is desirable to eliminate the phospholipids from the extracted samples, it is more practical to avoid coeluting phospholipids and analytes of interest. This can be easily accomplished by performing a MS scan of the aforementioned transitions to locate the retention times of the phospholipids and adjust the chromatography accordingly (e.g. change the pH of the mobile phase, use a column with a different chemistry, change the gradient). It is also worthwhile to incorporate an organic wash (e.g. 100% Mobile Phase B for 1–2 min) after the elution of the analytes to wash out the phospholipids after each injection. When it becomes impossible to separate them, a more selective extraction method will be needed.

To investigate the matrix effect quantitatively, a bioanalyst should select six individual matrices (no pooling), one lipidemic lot and one hemolyzed lot, extract them as blank samples, and then spike them with the analytes(s) at concentrations of 3 × LLOQ and ULOQ in replicate; the IS should also be included.[2] The matrix factor (MF) will be calculated for each matrix by dividing the peak area in the presence of matrix by the peak area in the absence of matrix.[50] While there is no acceptance criterion

for the MF, the analyst should strive to have a consistent MF across all six lots of plasma for small molecules in plasma samples. The MF should be also normalized for the IS by dividing the MF of the analyte by the MF of the IS; the CV of this ratio should be $<15\%$ as per GBMV. If the method fails to meet these, more method development work will be needed.

The matrix effect should be studied in greater detail if the formulation administered contains polyethylene glycol, polysorbate, or surfactants. The effect should also be studied in matrices from special populations such as renally/hepatically impaired subjects. It is also recommended by EMA to assess matrix effects from excipients if the drug is administered by the intravenous route.

The matrix effect on the analyte can be compensated with an SIL-IS,[46,47] or minimized by a number of methods, such as a change in chromatography or employing a thorough sample extraction procedure (e.g. SPE and LLE).

In order to avoid revalidation or unexpected events during the sample analysis phase, it is the opinion of the authors that matrix effects should be evaluated carefully during method development and appropriate methodology adopted to eliminate or minimize their impact on the method. During method development, an infusion experiment[51] can be carried out to identify the regions that have severe matrix effects. The results can guide the development of appropriate chromatography. If it is impossible to separate the coeluting interference from the analytes of interest, a more thorough extraction method will be needed.

16.3.11 Template for a method validation

Table 16.1 shows an example of a method validation template that in the opinion of the authors is up to the current regulatory expectations and industrial standards.

16.3.12 Validation report

The validation report should contain at least the following elements:

- Summary of validation performance
- Operational description of the method with literature references, if any
- Description of reference standards (batch, CoA, storage conditions, expiry dates)
- Preparation of standards and QCs (dates, matrix, anticoagulant, storage temperature)
- Acceptance criteria
- Calibration range and response function
- Table of all analytical runs with dates and outcomes (pass–fail)
- Table of calibration standards derived from accepted runs with back calculated concentrations, accuracy, and precision
- Table of QC results derived from accepted runs with accuracy and precision (both within and between run); QCs outside the acceptance criteria must be clearly indicated
- Stability data on stock solutions, working solutions, QCs
- Data indicating selectivity, appropriate LLOQ, carryover, matrix effect, dilution integrity
- Explanation of unexpected results with a description of the action taken
- Deviation from the method or applicable SOPs
- Typical chromatograms and mass spectra
- Results of Incurred Samples Reassay data, if available at that stage

Table 16.1 Method Validation Template

Validation Day	Experiment	Purpose	N
1	System suitability	Verify retention times and sensitivity of the LC–MS/MS system	5
	Blank (pooled matrix)	Quality of processing, carryover	4
	Zero sample	Impact of int. standard	1 or 2
	Calibration standards, at least six levels	Agreement with the calibration model	6 × 1 or 6 × 2
	LLOQ	Precision and accuracy at LLOQ	6
	QC.1, QC.2, and QC.3	Precision and accuracy over the calibration range	3 × 6
	Individual blanks	Specificity	6 × 1
	Lipidemic blank (if plasma or blood)	Specificity in lipidemic matrix	1
	Hemolyzed plasma (if plasma)	Specificity in hemolyzed plasma	1
	QC.1, QC.2, and QC.3, extracted for recovery	Extraction recovery	3 × 3
	QC.1, QC.2, and QC.3, unextracted, for recovery	Extraction recovery	3 × 3
	Above the ULOQ–QC	Integrity of dilution	5
	QC.2	Autosampler stability	3
2	System suitability	Verify retention times and sensitivity of the LC–MS/MS system	5
	Blank (pooled matrix)	Quality of processing, carryover	4
	Zero sample	Impact of internal standard	1 or 2
	Calibration standards, at least six levels	Agreement with the calibration model	6 × 1 or 6 × 2
	LLOQ	Precision and accuracy at the LLOQ	6
	QC.1, QC.2, and QC.3	Precision and accuracy over the calibration range	3 × 6
	QC.1 and QC.3, in the presence of extracted blank	Matrix effect in six individual matrices	2 × 6 × 3
	QC.1 and QC.3, neat	Matrix effect	2 × 3
	QC.2	Autosampler stability	3
3	System suitability	Verify retention times and sensitivity of the LC–MS/MS system	5
	Blank (pooled matrix)	Quality of processing, carryover	4
	Zero sample	Impact of internal standard	1 or 2
	Calibration standards, at least six levels	Agreement with the calibration model	6 × 1 or 6 × 2
	LLOQ	Precision and accuracy at the LLOQ	6
	QC.1, QC.2, and QC.3	Precision and accuracy over the calibration range	3 × 6

Table 16.1 Method Validation Template *(continued)*

Validation Day	Experiment	Purpose	N
	QC. 1 and QC.3	Processing stability at a selected temperature	2 × 3 × 3
	QC.1 and QC.3	Freeze–thaw stability	2 × 3
	QC.1 and QC.3	Interference experiments	2 × 3
	QC.2	Autosampler stability	3
	QC.1 and QC.3	Interference by comedication	2 × 3

N = *number of samples.*

16.4 APPLICATION OF A VALIDATED METHOD TO ROUTINE DRUG ANALYSIS

16.4.1 Organization of the analytical batch

The size of an analytical batch is limited by practical concern such as instrument capacity and stability of the drug or system. In most studies, all clinical samples from the same subject should be analyzed, if possible, in the same run to avoid between-run variability, which tends to be greater than the within-run variability. This is of particular importance in bioequivalence studies.

An analytical batch should be started by injection of a system suitability/reference solution followed by crucial samples (LLOQ, blank, ULOQ) so as to provide an early indication of whether the run is under control, and still allow the analyst an opportunity to take corrective action, if necessary. The corrective actions can include a change of the in-line filter, guard column, analytical column, lamp, or cleaning the source of the MS detector.

As for the placement of calibration standards in the batch, there is no set rule. Some analysts prefer to disperse them evenly throughout the batch, while others would start a batch with the calibration curve. QCs should be prepared at a minimum of three concentration levels and analyzed in duplicate with each analytical batch. One set of QCs should be close to the LLOQ and two to three times higher than the LLOQ. The second set should be at approximately 40–60% of the ULOQ, and the third at 70–90% of the ULOQ. QCs should be spread evenly throughout the analytical batch. There shall be at least six QCs in a batch or 5% of all samples in the batch, whichever is greater. Additional sets of QCs at different concentrations can be prepared to mirror the concentrations of study samples observed in the study.

16.4.2 Acceptance criteria

The commonly used run acceptance criteria are as follows. For the calibration curve, at least six calibration standards constituting at least 75% of the total number of standards must be within ±15% of the nominal concentration; in the case of the LLOQ the difference can be ±20%. This means that if eight calibration standards are extracted, at least six (75%) must be used to set up the calibration curve. If nine standards are extracted, at least seven (78%) must be acceptable for the calibration curve to be valid.

QC samples are the ultimate tool to accept or reject a batch of samples. The rule "4–6–15" is generally accepted, which means that six QCs at three concentration levels in duplicate must be extracted with a batch of study samples (<100), four (4) out of these six (6) must be within ±15% from the nominal, and each QC level must be represented in these acceptable QCs.

Sometimes additional acceptance criteria are included. These, for example, can be:

- Required coefficient of determination (r^2) of the calibration curve of at least 0.99
- Lack of interferences in drug-free samples
- Consistency of the absolute peak area or height of an IS
- Bracketing of samples: if some of the QCs fail, the study samples between those failing shall be rejected and reassayed
- Special QCs, such as hydrolysis QCs, if the assay involves, for example, an enzymatic reaction liberating a drug from its glucuronide or sulfate metabolite
- Use of a dilution QC

16.4.3 Dilutions

There are three reasons for sample dilution in bioanalyses: (a) to bring samples with a concentration above the ULOQ within the calibration range, (b) for parallelism in live blood analysis (LBAs) (not covered in this chapter), and (c) insufficient volume of samples. In all cases, the dilution should be done using the same matrix as the study samples. One has to be careful in performing dilutions of samples with concentrations close to the LLOQ, as the diluted samples may be classified as below the quantitation limit (BQL). For example, if the LLOQ is 1 ng/mL, the dilution factor is 2, and the back calculated concentration (no dilution factor included) is 0.77 ng/mL, the reported concentration should be BQL, and not 1.54 ng/mL.

16.4.4 Reassays

Every company or laboratory should develop and describe in an SOP its policy on repeat analyses. This policy has to be made available before starting a study. First, it must decide who is making decisions with regard to reassays. As for the cases involving some kind of analytical or technical difficulties, the decision should be left in the hands of the bioanalytical personnel. These instances can be:

- Poor chromatography: which may include interfering peaks making the integration impossible; no peaks at all, a chromatographic pattern very different from the expected one, collapse of the stationary phase, etc.
- Lost sample (LS): which may include dropped samples at any step of processing, leaking pipette tips, leaking screw caps, etc.
- Bad processing (BP): which may include any kind of human or robotic error—not adding a reagent, forgetting to add an IS, adding two portions of IS, etc.

Any of the above-mentioned errors should leave an audit trail in the form of a deficient chromatogram, note to the file listing the LSs or describing errors in the processing, or computer printouts in the case of robotic systems. No numerical results are associated with such attempts and they should be repeated as a single sample.

Occasionally, clinical samples exhibit concentrations above the validated range (AQL). Such samples should be diluted with the same matrix and repeated as a single sample.

On the other hand, study samples sometimes provide results which formally and chromatographically look correct, yet defy logic or seriously contradict previous results. The SOP must specify who identifies these potential reassays—the pharmacokineticist or bioanalyst and on what grounds. The goal of a bioanalyst or a pharmacokineticist is neither to squeeze study results into a preconceived model nor to smooth out pharmacokinetic profiles. At the same time, it is a scientific duty to challenge suspected results. One may suspect a pharmacokinetic outlier, if a predose sample from naive subjects contains a measurable drug concentration, if the pharmacokinetic profile exhibits a split or double maximum contrary to known pharmacokinetics, or if concentrations are very different (500–1000%) than expected. Such samples, which could be called "suspected outliers", provide numerical values, and repeats should be done in duplicate. The institution should also develop a comprehensive decision tree dictating a verdict in every foreseeable case to eliminate arbitrary decisions. A very good decision tree has been developed by Lang and Bolton.[52,53] Briefly, a 15% agreement between data is considered a confirmation if the repeats are done in duplicate, or 30% if only one repeat was possible. If results are too far apart, no result is reported.

It should also be noted that infrequent and random outliers do not influence the outcome of a study, if an appropriate number of subjects is selected to ensure appropriate statistical power. Pharmacokinetic reassays are discouraged in bioequivalence studies.

16.4.5 Incurred sample reassay or reanalysis

The issue of bioanalytical method reproducibility has been on the agenda of regulatory authorities since the 1990s, when Health Canada requested reanalysis for bioequivalence and bioavailability studies. This was mentioned also in the Crystal City I conference report.[3] However, in 2003, Health Canada removed this requirement. In the meantime, the FDA gathered evidence and observations based on the analysis of pharmacokinetic repeats and repeats in multianalyte assays. It was quite obvious that there were examples of bioanalytical assays that were not reproducible when applied to individual clinical or animal samples despite being formally acceptable on the basis of calibration standards and Quality Controls. After meetings in 2007 and 2008, the incurred sample reassay or reanalysis (ISR) has been widely adopted by the industry.[13] The goal of the ISR is to provide evidence of the method reproducibility, and detect either poor methods or poor execution of a good method. One has also to keep in mind that a scientist's understanding of the molecule's behavior and its interactions increase as the drug candidate progresses through the drug development process, and crucial information or understanding may not be available at the time of a first study.

Essentially, selected samples from a given study are reanalyzed as soon as possible after their initial analysis so as to detect and correct problems immediately. The ISR runs must be separate from the regular production runs in such a sense that there should not be mixed production-ISR runs. If the sample was initially assayed in dilution, the ISR of this sample should be done with the same dilution ratio. In general, 10% of samples (or a minimum of 20) should be reanalyzed for studies with <1000 samples. For larger studies, other rules can apply aimed at reducing the number of samples selected for ISR. While there is no official guidance on selecting particular samples for a given study, samples should be selected from different subjects (animals), time points, and dose groups. No pooled samples should be used unless the matrix volume is very low.

White papers and EMA[2,13,33] provide enough advice to make correct decisions. The ISR should be performed when a method is applied for the first in human (FIH) studies, in a new target population

(disease state, renal and hepatic impairment), all bioequivalence studies, upon major changes to the method, after a method transfer to a new laboratory, or in any study where scientific rationale suggests reassaying. As for drug–drug interaction studies, the opinions are divided, perhaps proving that the coadministered drug and/or its metabolites do not interfere with the analyte of interest is sufficient. In animal studies, the ISR should be performed once per species and matrix. Each laboratory must have an SOP to provide general rules on selecting samples, e.g. samples at C_{max} from different subjects, across different doses, number of samples, and selecting studies for ISR.

The results from initial analysis are compared with the second analysis and calculated according to the equation:

$$\%\text{Diff (variability)} = \frac{\text{Reanalysis concentration} - \text{Original concentration}}{0.5 \times (\text{Reanalysis concentration} + \text{Original concentration})} \times 100\% \quad (16.4)$$

The difference between the two analyses should be <20% for chemical assays (e.g. LC–MS/MS-based assays), and 30% for LBAs for two-thirds of the total number of samples reanalyzed. The ISR results should be included in the bioanalytical study report, and/or in the validation report, if these results are available by the time the report is finalized.

In the event that the results do not meet the acceptance criteria, an unexpected event investigation should be initiated. Based on the findings, resolution plans will be devised and executed. The impact on the quality of the bioanalytical data generated will be assessed based on the findings from the investigation. An investigation report should be compiled to summarize the investigation. Even in successful ISR experiments the bioanalyst is well advised to pay close attention to results outside the acceptance criteria, in particular if these are concentrated in discrete runs.

16.5 BIOANALYTICAL REPORT

A study should end with a report describing the procedure, its performance, and study results, where applicable. The data should be presented as a narrative and in tabulated form, and include:

- Operational description of the assay procedure.
- Information on reference standards (batch, CoA, storage conditions) and ISs.
- Preparation of standards and QCs (dates, matrix, anticoagulant, storage temperature).
- Acceptance criteria.
- Sample tracking—dates of receipt, conditions, storage location, and temperature.
- Table of all analytical runs with dates and outcomes (pass–fail).
- Equations used for back calculating of results.
- Table of accepted runs with the calibration curve parameters and correlation coefficients.
  Parameters should contain a sufficient number of digits to back calculate concentrations accurately. Slope of the calibration curve may change from day to day, yet it remains a valuable diagnostic tool. A consistent value of the slope suggests a solid assay. Dramatic changes may suggest modification to the method, errors, or maintenance done on a detector. Values of the intercept consistently above zero may suggest an interfering peak hidden underneath the peak of interest.
- A table of calibration standards derived from accepted runs with back calculated concentrations, accuracy, and precision.

The table should be complete, with no empty spaces. If a standard has been rejected, its value should be provided anyway in brackets or with an asterisk, and an explanation should be provided. A sample lost or disqualified for whatever reasons should be flagged as such. Interday precision and accuracy should be calculated providing the mean, SD, CV, % of nominal and number of observations. An adequate number of significant digits should be provided, so a reviewer can verify calculations and arrive at the same results.

- A table of QC results derived from accepted runs with accuracy and precision: QCs outside the acceptance criteria must be clearly indicated.
  All the rules specified above apply to this table also. Additionally, all evaluable QC values have to be reported and included in the statistics, whether or not these QCs meet the acceptance criteria. The bioanalyst may provide two sets of statistics: one using all the data, and the other excluding the results outside the acceptance criteria. Precision and accuracy calculated on QC data represent the precision and accuracy of the method.
- Explanation of unexpected results with a description of the action taken.
- A list of requested PK reassays, including the reason, and original and reassay results.
- Reasons for missing samples.
- Deviation from the method or applicable SOPs.
- Documentation for reintegrated data, including the initial and repeat integration results, reason for reintegration, the requestor of reintegration, and the manager authorizing the reintegration.
- Chromatograms from 20% of subjects in the pivotal bioequivalence studies; 5% in other studies.
- ISR results from the study, if applicable.

In addition, the report should contain the list of abbreviations and codes used, reference list, copies of the references, and copies of relevant SOPs.

16.6 VALIDATION, PARTIAL VALIDATION, AND CROSSVALIDATION

There are several kinds of validation:

- Original validation (before study, immediately after method development); described in Section 16.3
- Within-study validation; described in Section 16.4
- Partial validation performed to an already validated method if changes are made to the method
  Some of the changes may include:
  a. Introducing new analyst(s) to the method
  b. Change of platform (e.g. change in the LC system, mass spectrometer) within the same laboratory
  c. Change in anticoagulant for blood or plasma samples
  d. Change of species within the matrix
  e. Change of volume taken for extraction
  f. Modification to the validated range of concentrations
  g. Modification to extraction procedure, etc.
  A validated method may be altered intentionally or inadvertently. In any case, the change should be described in a note to file, and its potential impact evaluated. Intentional modifications should

be authorized by an analytical director and rationale provided in writing before its implementation, for example, in a form of a method validation amendment.
It is recommended that the analyzing laboratory should have an appropriate SOP to determine when a partial validation, full validation, or revalidation is warranted.

- Crossvalidation: should be performed when two or more methods are used to produce data within the same study or across many studies, or the same method is used to support a study at two different laboratories. Crossvalidation should be performed before committing study samples to analysis under these circumstances. The crossvalidation can be done using spiked QCs or authentic study samples. In the former, the acceptance criteria of ±15% should be used. If study samples are used, the acceptance criteria typical for an ISR study are recommended, i.e. ±20% for at least two-thirds of the samples.

16.7 BIOANALYTICAL METHOD VALIDATION—OTHER PARAMETERS AND ISSUES

Sections 16.1–16.6 presented the current state of the art in the area of method validation and execution of a bioanalytical study. In this section the authors evaluate critically some aspects of validation where either there is no consensus, or an alternative solution may be needed, or issues that at this time have not been adequately addressed by the bioanalytical community.

16.7.1 Chromatography

It may appear strange that chromatography has to be brought up as an important but almost forgotten aspect of bioanalytical method validation. The tandem mass spectrometric detector is such a powerful tool and is used so commonly that the proper chromatography for many can be an afterthought. In many cases the tandem mass spectrometric detector can even correct deficiencies of poorly developed chromatographic methods. Very frequently, a generic gradient system on any C8 or C18 column is applied without much consideration of the analytes to obtain as short a run time as possible. There are several important reasons why chromatography should not be too simplistic:

- Peaks of interest should be separated from the area where most endogenous compounds in the matrix elute. These endogenous materials can be observed as a dip(s) in the Total Ion Current. In these areas of chromatograms there can be a huge competition for ionization, potential drop in the sensitivity, and significant matrix effects. One also cannot forget about late-eluting peaks, which may show up even a long time after the original injection and interfere with subsequent sample injections.[46]
- Conversion of unstable metabolites such as N-oxides or glucuronides, which typically elute before the parent drug, may convert back to the drug at a high temperature in the MS ion source. If there is no chromatographic separation between the parent and metabolites, this will result in artificially elevated concentrations of the parent drug.

It is opinion of the authors that a conscientious bioanalyst should find the right balance between good chromatography and the run time.

16.7.2 Alternative acceptance criteria

The fixed range, commonly used as run acceptance criteria "4–6–15", which means that for a run to be accepted it has to have an acceptable calibration curve and four out of six QCs within 15% of the nominal value with all QC levels being represented, is unfortunately arbitrary, ad hoc,[54,55] and unscientific. The only criterion is accuracy, and an assumption is made that precision will be satisfactory. However, it describes quite well what the industry and regulatory agencies are willing to accept in terms of a balance between the quality and efficiency. The expectation that the overall precision and accuracy obtained in a study based on these acceptance criteria will be always $\leq$15% is arguably overoptimistic and unfounded.

Hartmann et al.[56] calculated that in order to obtain mean values within the limits of $\pm$15% and with a probability of 95%, the bias and RSD (%CV) should be 8% with $n = 5$. The fixed range approach is totally pragmatic, not based on statistical principles, and confuses precision and accuracy. Acceptance criteria should be scientifically valid, able to detect errors and false alarms, easy to use, and provide immediate answers. The MVG recognized that a confidence interval (CI) approach is an alternative for acceptance criteria.

The analyst makes measurements which are related to the concentration, not the actual concentration itself. These measurements provide only a certain probability that the true concentration will be within a certain range.[57] If analytical errors are random they follow the normal Gaussian distribution. Hence, 68% of the results fall within one SD of the mean, 95% within 1.96 (popular 2) of the mean, and 99.7% within 3.09 (popular 3) of the mean.

The 99% CI is equal to

$$99\%\text{CI} = \text{mean} \pm 2.58s \tag{16.5}$$

where s is the SD (see also Chapter 2). Gross errors (e.g. bad chromatography or sample processing) should be eliminated from calculations. The acceptance criterion is simple, all QCs must fall within the CI.

The CI-based acceptance criteria are easy to use and provide an immediate answer, although they do not address accuracy. They can be even more liberal than the fixed range "4–6–15" rule. At low concentrations the RSD of many bioanalytical methods is on the order of 10–15%, and hence the acceptance criteria will be $\pm$25.8–38.7%.

Another concept of run acceptance criteria enjoys at least moderate support of the bioanalytical community. It is called "total error acceptance criteria" and requires the summation of the absolute values of the bias and CV and to apply the rule "4–6–20" or "4–6–25". In simplified form it is used in clinical analysis and has been incorporated into some Laboratory Information Management Systems (LIMS).

Much more sophisticated statistical acceptance criteria for a method validation based on total error was proposed by Hoffman and Kringle.[55] This approach proposes a formal statistical framework for evaluation of a bioanalytical method. The three reports[14–16] prepared by a commission of the SFSTP describing validation procedures and acceptance criteria are also based on the total error concept and rigorous statistics.

16.7.3 Regression selection

The MVG stipulates in at least two places that the relationship between the response and the concentration be established using statistical tests for goodness of fit. Though the bioanalytical

community generally treats the text of the MVG with great respect, this requirement in most of the cases is merrily ignored. Most of the laboratories apply without much consideration their common regression and the weighting scheme which in most cases is $1/x^2$ and to a much lesser extent, $1/y^2$. What matters in practice is the quality of the inverse prediction more than the goodness of fit.[16] There are a number of publications describing rational and statistically valid procedures for selecting the proper regression.[58–60]

The MVG leans heavily toward the use of the simplest response function, i.e. linear calibration curve. From the beginning, there was a lot of confusion with regard to linearity of the calibration curve. The bioanalytical community assumed that linearity means a rectilinear curve expressed by the equation

$$Y = ax + b \tag{16.6}$$

where a is the slope, and b is the intercept of the calibration line.

In statistical sense, a function is "linear" if it is a linear combination of its parameters. The quadratic function is also linear in its parameters although its graphical plot is not a straight line.[16]

This emphasis on rectilinearity may cause problems as well. A subjective judgment as to whether or not a set of points represents a linear model may be at variance with statistical tests, and this mistake can be in either direction.[61] A linear calibration curve may be forced on data that are slightly, but clearly nonlinear.

There could be several causes for the curving of calibration lines in chromatographic assays; receptor-binding assays are inherently nonlinear. Certain kinds of detectors provide nonlinear responses, such as the electron capture detector in gas chromatography, or in fact any detector if the range of calibration curve is excessive and covers concentrations of several orders of magnitude. To show the detector linearity one needs to inject increasing amounts of the unextracted analyte solutions and record responses. The analytical process may be also responsible for non-linearity; for example, variable extraction recovery or adsorption. To detect and document non-linearity one may use a number of techniques[62–64]:

- Visual assessment—subjective and requires an expertise in analytical methodology.
- Conventional analysis stemming from least squares regression—several approaches can be used such as components of variance, lack-of-fit testing, quadratic regression.
- Analysis of consecutive differences—simulates the visual assessment of linearity.
- Comparison of observed values against expected results (residuals, see 2.2.2.1).

A very simple test for linearity based on the residuals (% deviation from nominal) is called the "sign test".[61] The signs of residuals should be distributed at random between plus and minus, if no systematic error is involved. If a sequence of signs looks more like −++++-, a curvature of the regression line and a lack of linear fit could be suspected.

Particularly useful as a diagnostic tool is the analysis of consecutive differences, also called "deltas", or rather a variation of it. Peak height (area) ratio or absolute peak height (area) divided by the nominal concentration gives a value which is called a "response factor" or "unit ratio", and is readily available in some LIMS. This value represents the slope of the calibration curve at this point, and should be constant and equal to the overall slope of the rectilinear calibration curve. If a decreasing/increasing trend in the value is visible, the response function cannot be linear. Additionally, if response factors are constant over the whole calibration curve with the exception of the lowest standards, an interference hidden underneath the peak of the analyte should be suspected.

The coefficient of determination (r^2) is used frequently as a measure of the goodness of fit. However, it is a rather poor predictor of the fit. Even poor calibration curves may have coefficients of determination quite high and >0.99; this value is frequently used as an acceptance criterion or at least an indicator for the goodness of fit. Also, as discussed earlier in Chapter 2, one should not rely too heavily on the r^2 value as a measure of linearity because this parameter includes curvature as well as random errors. There is some practical use for the coefficient of determination in one sense that although a high correlation coefficient does not ensure a good calibration curve, but a low one, say <0.99, indicates that the calibration curve is biased with serious errors and probably is unacceptable.

To conclude, an automatic application of linear regression may be as wrong as the use of a complicated model to a simple chromatographic assay.

16.7.4 Blood samples

For convenience, plasma, rather than blood, is the most common matrix in bioanalysis. However, several types of experiments in blood are necessary at or before the stage of validation in order to provide meaningful data.[65]

In most cases, the blood/plasma ratio of analyte concentrations is approximately 1 ± 0.5. It is well known however, that several drugs bind preferentially to red-blood-cell membranes, hemoglobin, or carbonic anhydrase. In these cases the blood-to-plasma ratio is much >1. These, for example, are (blood/plasma ratios indicated in parentheses): pimobendan (3.2–4.5), cyclosporin A (2.0–4.6), tacrolimus (22.6–55.5), methazolamide (241), acetazolamide (2.9), chlorthalidone (30.7–32), chloroquine (3.5), rapamycin (14.3), and ribavirin (~100). The blood/plasma ratio must be known to ensure that plasma is indeed the most appropriate medium for bioanalysis and pharmacokinetic evaluation. A potential consequence of using plasma as a matrix for an analyte with a high blood to plasma ratio is that a small amount of hemolysis can cause an artificially high concentration of the analyte in the sample.

Another issue to consider is the thermodynamics and kinetics of red-blood-cell partitioning. There are several drugs (e.g. cyclosporin A, amitriptyline, and nortriptyline) that partition differentially at different temperatures. The erythrocyte/plasma ratio for amitriptyline at 2–10 °C was 0.3, and 0.5 at 40 °C. For nortriptyline, the ratio was approximately 0.85 at lower temperatures and 1.25 at 40 °C.[66] If the drug partitioning shows such a dependency on temperature, a bioanalyst has to devise a proper plasma-harvesting procedure for the clinical sites. The most common procedures of obtaining the blood, cooling in an ice bath, and centrifuging at 4 °C may not be appropriate, and perhaps centrifuging at 37 °C would provide better results. What is more, applying variable procedures with regards to the centrifuging temperature may introduce immediately a sizable variation of the drug concentration in plasma.

16.7.5 Drug stability in blood

Many drugs are actively metabolized in the blood cells of humans, e.g. chlorpromazine, captopril, haloperidol, heroin, isoproterenol, ribavirin, testosterone, and many more. It is important to devise a plasma-collecting procedure so that the integrity of samples will be maintained and the drug concentration in the plasma will be the same at the time of phlebotomy as at the end of the process when plasma samples are placed in a freezer. The bioanalyst must establish temperature conditions and time

limits for blood processing and plasma harvesting. A frequently encountered problem is that a bioanalytical method for plasma exists at this stage, but there are no plans for studies in blood and hence no need to validate the method in whole blood. It is recommended either to apply the validated plasma method to the plasma obtained from whole blood in a controlled stability experiment (so-called "whole-blood stability" evaluation), or to qualify the plasma method for use in whole blood in such an experiment.[34] Methods that involve liquid extraction in most cases will work equally well for plasma and blood. Methods that employ protein precipitation may require more robust precipitation and increasing the ratio of the precipitating agent (methanol, acetonitrile) to blood to about 10. No concentration data are needed, because the peak area ratio of the drug to the IS plotted against the time or temperature should reveal instability.

16.7.6 The "other" matrix effect

In today's bioanalysis dominated by the LC–MS/MS technique, the term "matrix effect" means almost exclusively the effect that endogenous extracted components of the matrix have on the ionization in an MS detector that results in the decreasing or increasing of the instrument's response. However, there are "other types" of matrix effects. Here are some of these as reported in the literature or experienced by the authors:

- Variable protein binding in FVB/N strain of mouse plasma

An LLE method has been developed to quantify reserpine in mouse plasma.[67] The method performed well for the standards and QCs prepared in control plasma, but for the study samples in the FVB/N mouse plasma, the IS was not detected in 30% of samples, and was decreased by a factor of 5–10 in additional 20% of samples. The phenomenon was attributed to specific protein binding in FVB/N plasma, which was eliminated by the addition of sodium EDTA.

- pH of samples

A liquid extraction was performed on plasma samples without buffering them at the physiological pH of plasma (7.4), which was an appropriate pH for that extraction. However, stored plasma releases carbon dioxide, which changes pH. Freeze thawed plasma may reach a pH of 8.5. A methanolic supernatant evaporated and redissolved can have a pH of 9.5.[68] These pH increases can have an adverse effect on the extraction and stability of analytes. Blood also may be affected by the same process but to a lesser extent.

- Ionic strength in the ion-exchange process

Urine samples were injected directly into a column switching system containing an ion-exchange column.[69] Some samples provided suspicious results. It was discovered that these were very concentrated urine samples with much higher levels of salts. The volume of injection was reduced by a factor of 10 and the matrix effect disappeared.

- Protein content

Recovery during protein precipitation is frequently incomplete; solubility of the drug and its protein binding plays a role. Total concentration of protein in human serum varies between 58

and 77 mg/mL depending on age and gender,[70] less in undernourished and sick individuals. In one experiment, recovery of triamterene from serum was measured at 50 ng/mL during precipitation with 10% perchloric acid. The serum was diluted with 0.9% sodium chloride solution in the following ratios (v/v): undiluted, 2:1, 1:1, and 1:2. The recoveries were 64%, 75%, 80%, and 88%. Obviously, the potential for significant errors exists due to differing protein content. An appropriate IS corrected the recovery problems, but the errors might have gone undetected without it.

- Presence of Protease Inhibitor Cocktail (PIC)

On occasion it was observed that in an LC–MS/MS-based assay, the absolute peak areas of the ISs in subject samples were different from those observed in standards and QCs. A hypothesis that the variability of adding a PIC to the samples was responsible for this effect was tested. The targeted concentration of PIC in the plasma samples was 1%. Control samples were spiked into plasma that contained PIC added at 0%, 1%, 2%, 4%, and 6%. The samples were extracted according to the method in triplicate. It appeared that the absolute peak area of the drug was inversely related to the concentration of the PIC in the plasma; the loss was at worst 24%. However, the IS compensated for this effect and the observed changes in the peak area ratio were approximately 5%; well within the precision and the accuracy of the assay.

16.7.7 Hemolyzed plasma

The blood from clinical or toxicology studies on occasions is sometimes partially hemolyzed due to drug action (e.g. ribavirin), disease state, addition of additive used to stabilize the drug, or technical errors while obtaining the blood. Hemoglobin and bilirubin are released from the red blood cells causing the plasma to appear pinkish to deep red. The degree of hemolysis is measured by the concentration of hemoglobin in the plasma. Plasma with concentration $<$30–50 mg/dL of hemoglobin is considered not hemolyzed, while plasma with hemoglobin $>$300 mg/dL is considered badly (grossly) hemolyzed. The degree of hemolysis can be easily estimated using the Becton Dickinson scale as presented by Hughes.[71] According to this paper, the impact of hemolysis can be considered a special case of matrix effect, and is caused by either the presence of additional interfering peaks or serious suppression of the MS/MS signal due to the presence of hemoglobin, bilirubin, and other endogenous components of erythrocytes. Better extraction procedures (LLE, or SPE instead of protein precipitation), more selective chromatography, replacing of an analog IS with an SIL-IS, or dilution of hemolyzed plasma eliminated these effects. The authors propose doing validation experiments using simulated hemolyzed plasma that is produced by adding 2% of totally hemolyzed plasma to regular plasma and which corresponds to approximately 550 mg/dL hemoglobin. If the back calculated concentrations of QCs prepared in hemolyzed plasma are outside of the regular $\pm$15% limit such samples cannot be analyzed, and the experiment should be repeated using 1% of blood in plasma.

In the experience of the authors, one of their assays could not provide satisfactory results in plasma containing approximately 1060 mg/dL of hemoglobin. In fact, the precipitation of proteins was not complete with the supernatant being visibly pinkish. The chromatographic column could not handle the protein load and failed in the middle of the run. The experiment was performed successfully at 530 mg/dL of hemoglobin.

16.7.8 Lipidemic plasma

Lipidemic plasma should be considered as part of method validation for several reasons:

- Specificity: Brazilian Agência Nacional de Vigilância Sanitária document demands to run at least one lot of lipidemic plasma and one lot of hemolyzed plasma as a part of the specificity experiment.[72]
- Postprandial plasmas have elevated levels of lipids and proteins; it is very easy to see as the plasma can be very cloudy. The extra lipids can be extracted during the extraction process, and eventually injected into the LC–MS/MS system and cause potential signal suppression, increase in the column backpressure, or even column overload and collapse.
- The extra lipids present in the plasma can be considered as one more reagent or organic solvent in the system. Hence, it may change extraction recovery during the LLE or SPE. An SIL-IS will compensate for it, but an analog IS may not.

16.7.9 Internal standard consistency

The consistency of an IS peak area or height is another parameter of increasing interest in the bioanalytical community and regulatory agencies.[34] The fact is that in today's mostly automated methods there is no good reason for the great variability of the IS, though the response from the tandem mass spectrometer has certain inherent variability and is nowhere close to the stability of simpler UV or fluorescence detectors. It is assumed that the variability is caused by the unidentified matrix effects, and that such a method is perhaps underdeveloped and not sufficiently robust. Two approaches have emerged so far on how to deal with a highly variable IS response:

- Set up a fixed range of 50–150% around the mean of all samples (standards, QC's, zeroes, study samples), and flag and possibly repeat samples outside the range.
- Trend the absolute peak area/height in the known samples (standards, zeroes, QCs) and investigate unknown samples in which the ISs are outside these expected limits.

It must be stressed that the nature of an ideal IS, a C^{13} stable isotope label, is to compensate for all variables in the analytical process. The highly variable IS cannot be an automatic ground for sample or batch rejection, but can trigger an investigation whether or not the IS indeed compensates for these variables. One may use a standard addition method to prove the point.

The most commonly used SIL-ISs are the deuterated analytes, which contain three to six deuterium atoms. The deuterated analytes are a little more hydrophilic as compared to the analyte, their retention times are normally 1–2 s shorter than those of the analytes, and in rare cases such an IS may not behave in the same way as the analyte.[73]

16.7.10 Tubes and containers

In the opinion of the authors the bioanalytical community devotes too little attention to the tubes and containers. Even though proteins and lipids can help to form an emulsion which could aid the solubility, or homogeneity of poorly soluble analytes, or cover the active sites that may bind to a drug, they still do not ensure the avoidance of adsorptive losses.[67] It is even worse in the case of matrices that contain very little or no proteins or lipids, such as urine, CSF, bronchoalveolar lavage, tears, and so on.

In the laboratory of one of the authors, the aqueous solution of an analyte well known for adhering to various surfaces was placed in 11 different types of glass and polymer tubes, at the volume of 1 mL each, at a concentration 100 ng/mL. The analyte was known not to be light sensitive. The solutions were left in the tubes for 1 h and vortexed from time to time. After 1 h, aliquots were injected into a simple HPLC/UV system to ensure the best reproducibility of the response.

The differences were quite drastic. Selecting a wrong type of tube to collect the CSF samples decreased the apparent concentration by as much as 50%. The bioanalyst should propose the proper type of tubes to collect the samples and the proper procedures. In this example, addition of 10% isopropanol to CSF was needed to avoid adsorption to the tubes and maintain the integrity of samples.

16.7.11 Robustness testing

According to the ICH (Q2(R1)), the robustness/ruggedness of an analytical procedure is a measure of its capacity to remain unaffected by small, but deliberate variations in method parameters and provides an indication of its reliability during normal usage. The factors influencing the assay can be quantitative (pH, concentration, temperature, time) or qualitative (batch of HPLC columns).[74] It is impractical to investigate all parameters of a method. The bioanalyst who is the originator of a method certainly knows the crucial factors of the method. At least one or two of them should be investigated. The introduced changes should mimic potential errors in day-to-day laboratory operations. These, for example could be:

- pH of mobile phase

  Prepare mobile phase which deliberately is 0.2 pH unit off the target on the positive or negative side.

- Percentage of organic solvent in the mobile phase

  Prepare mobile phase which deliberately is slightly off in terms of organic content, such as 62% of methanol instead of 65%.

- Composition of extracting solvent

  If for example the extracting solvent is a mixture of hexane/isopropanol at a ratio of 90:10, try extracting with a mixture at a ratio of 95:5.

- If a method calls for the completion of extraction within certain time limits for the reason of stability, set aside a set of samples and complete it within time limits that are 50% longer.

Vander Heydan et al.[74] provided comprehensive and elaborate guidance on robustness testing, including the appropriate statistics. As the bioanalytical community, generally speaking, tends to opt for simple solutions, the pragmatic acceptance criteria of $\pm 15\%$ used for the batch acceptance and stability testing can be applied here. The robustness testing does not need to be extensive or costly. In many cases, the existing extracts can be reused and reinjected under different conditions of the assay.

16.7.12 Bioanalysis in tissues and homogenates

While plasma, serum, and urine are the most common matrices in bioanalysis, it is not uncommon to analyze tissue samples such as liver, brain, heart, and kidney. In general, the workflow

"grind-extract-measure" applies to tissue sample analysis. Representative tissue sample is excised from the organ, homogenized, followed by extracting the homogenates and analysis by LC–MS/MS.[75]

It is important to obtain representative tissue samples. Unlike plasma or serum samples, drug concentrations can vary between different parts of an organ, e.g. a drug-coated stent that releases the drug slowly will have a concentration gradient in the surrounding tissue, with higher concentration closer to the stent.[76,77] Therefore, a thorough understanding of which part of the organ provides meaningful drug measurement is needed. For some cases, whole organ (e.g. liver) is required while in other cases, a slice of the whole organ (e.g. brain) is sufficient.

Generally, a buffer such as phosphate buffer or phosphate buffered saline (PBS) is used to homogenize the tissue samples. In some cases, water or plasma or even whole blood is used instead. It is recommended that the bioanalysts carefully evaluate the recovery of the analytes in different reagents. While water and PBS are relatively easy to handle, they may not be able to provide good recovery over the range of the curve due to nonspecific binding.

Another key factor is carryover. An adequate washing procedure should be put in place to avoid carryover of drugs from previous samples during homogenization. The homogenate can be extracted directly with the designated extraction method, or it can be diluted with plasma (e.g. 1:10 or an even higher dilution factor) and then extracted accordingly with the plasma method. The extracted samples will then be analyzed by LC–MS/MS or other techniques.

In general, the matrix effect is more severe with tissue homogenates due to the presence of a large number of endogenous components. It is recommended that the bioanalysts carefully evaluate the matrix effect and adopt an appropriate extraction procedure to ensure ruggedness of the assay. The use of a SIL-IS is highly recommended if available.

In general, there are no "official" acceptance criteria for tissue sample bioanalysis. Zhang et al.[76] reported a validated method on zotarolimus in stented swine arteries. Nonetheless, a "fit-for-purpose" qualification is usually adopted. It can include an accuracy and precision run, recovery and matrix effect, and stability evaluation. It is not necessary to use the same criteria as plasma methods such as $\pm 15\%$ from nominal concentration for stability evaluation. The bioanalysts should carefully consider the performance of the assay before deciding on acceptance criteria.

16.8 EMERGING TECHNOLOGIES IN BIOANALYSIS

16.8.1 Dried blood spot

Dried blood spot (DBS) is a microsampling technique that was proposed in 1963 by Guthrie[78] to collect finger-pricked capillary blood from neonates for detecting genetic metabolic disorders. Over the last several years the technique has been applied to drug development to analyze human and animal samples. There are several practical advantages of this approach:

- Reduction in sample volume; very important in pediatric studies and small animals
- Reduction in the number of animals for toxicokinetic studies; no more satellite groups needed
- Reduction in costs of collecting, storing, and shipping
- Better sample stability for certain chemotypes

DBS method development and validation are based on the same principles as regular chemical or chromatographic methods with several technique-specific modifications.[79–81] Before a bioanalyst embarks on employing such a strategy there must be a clear understanding that at this time there has not been any drug submission accepted by regulatory agencies based solely on DBS. At the time of this writing, the FDA considers DBS methods as supporting evidence, and the early adopters must provide the conventional plasma data along with the bridging studies supporting the use of DBS methods.[82]

The special requirements in DBS method development and validation are:

- Selection of the paper for blood collection, both treated with stabilizing agents and untreated; variability of the cards
- Assay robustness related to pipetting variance (10–30 μL)
- Whole-blood stability
- On card stability
- Dilution technique—smaller punch or dilution with control matrix extract
- Application of IS—no IS on the card, IS over DBS, IS under DBS
- Intraspot and interspot homogeneities
- Carryover from punch and from mat
- Hematocrit effect—normal hematocrit is 45%, but the range is 25–75%.

16.8.2 Liquid chromatography–high-resolution mass spectrometry

Recently, there have been numerous discussions at scientific conferences (American Society for Mass Spectrometry, American Association of Pharmaceutical Scientists) on the use of high-resolution mass spectrometers (e.g. time-of-flight (TOF) or orbitrap-based spectrometers) to simultaneously perform qualitative and quantitative analyses of the same sample. This represents a significant paradigm shift. Triple quadrupole-based mass spectrometers with their superior sensitivity are the workhorses for quantitative analysis. High-resolution platforms (e.g. TOF or orbitrap-based mass spectrometers), with their superior mass resolution have been used extensively in qualitative analysis. With the newer generation of high-resolution mass spectrometers that combine high resolution and good sensitivity, and competitive pricing, a single platform can be used to perform both qualitative and quantitative (so-called "Qual/Quan") analysis.[83]

In liquid chromatography–high-resolution mass spectrometry (LC–HRMS), total ion chromatograms (TICs) are acquired over a predefined m/z range (e.g. 100–2000 m/z) with a preset mass resolution (e.g. 20,000) on the mass spectrometer. Extracted ion chromatograms (EICs) are generated post data acquisition from the TICs with the exact masses of the target analytes and a predefined mass extraction window (MEW). Quantitative information is then obtained from the EICs, similar to that of the triple quadrupole-based method. Unlike triple quadrupole-based methods, in which the mass spectrometers are typically set at unit resolution, with full width at half maximum of 0.7 Da for data acquisition, different mass resolutions are typically available on full-scan mass spectrometers, depending on the type of mass spectrometers used. Higher mass resolution in general provides better selectivity, especially in a complex sample matrix.

There are a couple of advantages to using HRMS. (1) When using HRMS for method development, there is no need to determine the most favorable product ions since data from high-resolution full-scan acquisition should provide sufficient selectivity. This should expedite the method development

process, especially when a large number of compounds are monitored, as in the case of discovery settings. (2) In triple quadrupole-based methods, data from selected SRM transitions are acquired. In HRMS-based methods, data from a wide mass range is collected during data acquisition, the data can be "mined" post-acquisition for different analytes of interest such as "unknown" metabolites, phospholipids, etc.

The key parameters for HRMS-based methods are the mass resolution setting (R) of the mass spectrometer during data acquisition and the MEW used to extract ion chromatograms during data processing. The interplay between these parameters has been discussed in the literature.[84] Additional work needs to be done to provide further understanding of these parameters and clearly understand their impact on the quality of the data generated. Bioanalysts should carefully investigate these parameters during method development.

Thus far, HRMS has been successfully applied to determine pharmaceutical compounds, pesticides, veterinary drugs, and peptides in both discovery and development settings.[85,86] Based on currently published data, HRMS can provide sufficient sensitivity, selectivity, and ruggedness for routine bioanalysis. At the time of writing this chapter, regulatory agencies have not provided any formal guidance on the use of HRMS in regulated bioanalysis. Another area that needs to be addressed is the amount of data generated from full-scan data acquisition. Thousands of samples are analyzed during the course of development of a drug candidate and the amount of data accumulated at the end will require careful consideration of storage space and retrieval of data for review. Finally, in-depth discussions with regulatory agencies will be needed to gain perspectives and feedback on this new technology platform, in particular with regard to post-acquisition data mining.

16.8.3 Bioanalysis of therapeutic proteins by LC–MS/MS

Another emerging area in the bioanalysis field is quantitative determination of therapeutic proteins (e.g. monoclonal antibodies, domain antibodies) by LC–MS/MS.[87–92] Unlike traditional small molecules with molecular weights $<$1000 Da, therapeutic proteins in general have molecular weights $>$10,000 Da. This poses different challenges to bioanalysts attempting to use LC–MS/MS for analysis:

- Therapeutic proteins in general have similar physiochemical properties as other endogenous proteins; therefore, traditional sample cleanup techniques for small molecules such as protein precipitation are not suitable to use.
- With their high molecular weights, it is in general not feasible to monitor the intact molecule in their $[M + H]^{+}$ or $[M + 2H]^{2+}$ charge states since it is likely outside of the mass range of the mass spectrometer, especially if triple-quadrupole mass spectrometers are used. On the other hand, proteins in general have multiple charge states (e.g. $[M + H]^{8+}$, $[M + 2H]^{9+}$, etc.), and some of these charge states will fall into the mass range, although the sensitivity may suffer as a result.
- With their relatively large size, traditional narrow-bore HPLC or UHPLC columns used for small molecule analyses may not be ideal candidates for separation of proteins.

The most frequently used strategy for quantitative analysis of proteins is to digest the proteins enzymatically or chemically to form smaller peptides, preferably in the mass range of 1000–3000 Da. These small peptides are used as surrogates for the proteins (and in general referred to as "surrogate peptides") and can be extracted by LLE, SPE, or protein precipitation and analyzed by

LC–MS/MS. Stable isotopically labeled ISs can now be custom synthesized for use in the assay at a reasonable cost.

In terms of sample cleanup from serum samples, there are four major types:

- Immunocapture: an antibody specific to the protein of interest is used to capture the protein, while other proteins in the serum samples are washed out. The protein of interest is then eluted out for enzymatic digestion.
- Differential precipitation by organic solvent: this method explores the different solubility of pegylated proteins and nonpegylated proteins in organic solvents. For example, Wu et al.[91] reported that pegylated proteins are soluble in 0.1% formic acid in 2-propanol while other endogenous proteins are not. The serum samples are thus treated with 0.1% formic acid in 2-propanol to remove endogenous proteins.
- Precipitating out with all other proteins with an organic solvent: in this method, all proteins including the protein of interest and other endogenous proteins are precipitated out with an organic solvent e.g. methanol. The precipitated proteins are then resuspended in a digestion buffer for enzymatic digestion. It is to be noted that this method does not result in clean samples.
- SPE (both on-line and off-line): this is more applicable to peptides and small proteins.

Regarding digestion by endoproteases, a number of endoproteases (e.g. trypsin, chymotrypsin, GluC, AspN, LysC, protease K, and pepsin) have been reported. The most common one thus far is trypsin. It specifically hydrolyzes peptide bonds at the carboxyl side of lysine and arginine residues. Other enzymes hydrolyze peptide bonds at other specific amino acids. The resulting surrogate peptides can then be analyzed by LC–MS/MS. It is the authors' experience that the best surrogate peptides are between 10 and 30 amino acids in length for good retention on reversed-phase LC columns and reasonable sensitivity.

The samples can be further extracted post-digestion by SPE or 2D-HPLC, or simply injected directly to the LC–MS/MS system for analysis. It is recommended that the bioanalysts evaluate these possibilities carefully during method development.

Thus far, a number of LC–MS/MS-based methods for therapeutic protein analysis have been published. It is gaining traction in discovery settings because it mitigates the needs of precious reagents used in LBAs, and thus expedites assay development. Further work needs to be done to make this technology applicable to routine use, especially in the development arena. To name a few, improvement in sample preparation techniques, LC separation, MS sensitivity are some of the areas to be focused on.

Another key area that needs to be addressed is to establish the link between data generated by LBA and LC–MS/MS, which are two fundamentally different but complementary techniques. The data generated by each technique represent unique properties of the protein. For LBA, it relies on the binding of the capturing reagent. For LC–MS/MS, it relies on the generation of a surrogate peptide that is representing the protein of interest. How the two sets of data relate to each other is highly linked to the protein of interest and the capture reagents used in LBA and specific region of the surrogate peptide. A thorough understanding of the link between the two sets of data is needed.

At the time of this writing, regulatory agencies have not provided any formal guidance on the use of LC–MS/MS-based data for filing. However, it is certainly an area that has tremendous growth potential.

16.9 CONCLUSIONS

Bioanalytical method validation and regulated bioanalysis are an integral part of a drug development program. They have evolved over the years in terms of technological platforms and regulations. Different technological platforms have been used to analyze chemical-based drug candidates, from LC–UV and LC–Fluorescence to LC–MS to LC–MS/MS. Guidance documents from regulatory bodies across the globe are revised to reflect the current technologies but are not yet fully harmonized. At the onset of method development, bioanalysts should carefully evaluate the physiochemical properties of the analyte of interest, its metabolites, assay requirements such as LLOQ, matrix, against the currently available technological platforms. Different parameters that can affect the assay performance should be carefully evaluated. Once a desired method is developed, the bioanalysts can then proceed with method validation and bioanalysis in accordance with the different regulatory guidelines and laboratory-specific SOPs. In addition, methods can be amended when new information is available, for example, discovery of a new metabolite that requires monitoring, or as more data are produced, e.g. clinical pharmacokinetic data from an FIH study may drive a lower LLOQ, or other unforeseeable issues. It has to be understood that the process of method development and validation is a continuum. There is a life cycle to a bioanalytical assay as drug development progresses, and it should be science driven.

Acknowledgments

Krzysztof Selinger would like to thank Dr Daksha Desai-Krieger for friendly encouragement and helpful discussions. The authors would like to thank Dr Anne-Francoise Aubry for reviewing the manuscript and providing valuable feedback.

References

1. Food and Drug Administration. *Guidance for Industry: Bioanalytical Method Validation;* US Department of Health and Human Services, FDA Center for Drug Evaluation and Research: Rockville MD, 2001.
2. European Medicines Agency, Guideline on Bioanalytical Method Validation, London, UK, Committee for Medicinal Products for Human Use, 2011.
3. Shah, V. P.; Midha, K. K.; Dighe, S.; McGilveray, I. J.; Skelly, J. P.; Yacobi, A.; Layloff, T.; Viswanathan, C. T.; Cook, C. E.; McDowell, R. D., et al. *Pharm. Res.* **1992,** *9,* 588–592.
4. Shah, V. P.; Midha, K. K.; Findlay, J. W. A.; Hulse, J. D.; McGilveray, I. J.; McKay, G.; Miller, K. J.; Patnaik, R. N.; Powell, M. L., et al. *Pharm. Res.* **2000,** *17,* 1551–1557.
5. Viswanathan, C. T.; Bansal, S.; Booth, B.; DeStefano, A. J.; Rose, M. J.; Sailstad, J.; Shah, V. P.; Skelly, J. P.; Swann, P. G.; Weiner, R. *Pharm. Res.* **2007,** *24,* 1962–1973.
6. Miller, K. J.; Bowsher, R. R.; Celniker, A.; Gibbons, J.; Gupta, S.; Lee, J. W.; Swanson, S. J.; Smith, W. C.; Weiner, R. S. *Pharm. Res.* **2001,** *18,* 1373–1383.
7. DeSilva, B.; Smith, W.; Weiner, R.; Kelley, M.; Smolec, J.; Lee, B.; Khan, M.; Tacey, R.; Hill, H.; Celniker, A. *Pharm. Res.* **2003,** *20,* 1885–1900.
8. Smolec, J.; DeSilva, B.; Smith, W.; Weiner, R.; Kelly, M.; Lee, B.; Khan, M.; Tacey, R.; Hill, H.; Celniker, A. *Pharm. Res.* **2005,** *22,* 1425–1431.
9. Abbott, R. W. *Bioanalysis* **2010,** *2,* 703–708.

10. Abbott, R. W.; Gordon, B.; van Amsterdam, P.; Lausecker, B.; Brudny-Kloeppel, M.; Smeraglia, J.; Romero, F.; Globig, S.; Globig, M.; Knutsson, M., et al. *Bioanalysis* **2011,** *3,* 833–838.
11. Abbott, R. W.; Brudny-Kloeppel, M. *Bioanalysis* **2009,** *1,* 273–276.
12. Tudan, C. Highlights of 4th Regulated Bioanalysis Workshop, SQA Technical Document 2010-2, Charlottesville, VA; Society of Quality Assurance, 2010.
13. Fast, D. M.; Kelley, M.; Viswanathan, C. T.; O'Shaughnessy, J.; King, S. P.; Chaudhary, A.; Weiner, R.; DeStefano, A.; Tang, D. *The AAPS J.* **2009,** *11,* 238–241.
14. Hubert, Ph.; Nguyen-Huu, J.-J.; Boulanger, B.; Chapuzet, E.; Chiap, P.; Cohen, N.; Compagnon, P. A.; Dewé, W.; Feinberg, M.; Lallier, M., et al. *J. Pharm. Biomed. Anal.* **2004,** *36,* 579–586.
15. Hubert, Ph.; Nguyen-Huu, J. J.; Boulanger, B.; Chapuzet, E.; Chiap, P.; Cohen, N.; Compagnon, P. A.; Dewé, W.; Feinberg, M.; Lallier, M., et al. *J. Pharm. Biomed. Anal.* **2007,** *45,* 70–81.
16. Hubert, Ph.; Nguyen-Huu, J.-J.; Boulanger, B.; Chapuzet, E.; Chiap, P.; Cohen, N.; Compagnon, P. A.; Dewé, W.; Feinberg, M.; Lallie, M., et al. *J. Pharm. Biomed. Anal.* **2007,** *45,* 82–96.
17. Braggio, S.; Barnaby, R. J.; Grossi, P.; Cugola, M., et al. *J. Pharm. Biomed. Anal.* **1996,** *14,* 375–388.
18. Dadgar, D.; Burnett, P. E.; Choc, M. G.; Gallicano, K.; Hooper, J. W. *J. Pharm. Biomed. Anal.* **1995,** *13,* 89–97.
19. Wieling, J.; Hendriks, G.; Tamminga, W. J.; Hempenius, J.; Mensink, C. K.; Oosterhuis, B.; Jonkman, J. H. *J. Chromatogr. A* **1996,** *12,* 381–394.
20. James, C. A.; Breda, M.; Frigerio, E. *J. Pharm. Biomed. Anal.* **2004,** *35,* 887–893.
21. Timmerman, P.; Lowes, S.; Fast, D. M., et al. *Bioanalysis* **2010,** *2,* 683.
22. Bansal, S. K.; Arnold, M.; Garofolo, F. *Bioanalysis* **2010,** *2,* 685–687.
23. Lausecker, B.; van Amsterdam, P.; Brudny-Kloeppel, M.; Luedtke, S.; Timmerman, P. *Bioanalysis* **2009,** *1,* 873–875.
24. van Amsterdam, P.; Lausecker, B.; Luedtke, S.; Timmerman, P.; Brudny-Kloeppel, M. *Bioanalysis* **2010,** *2,* 689–691.
25. van Amsterdam, P.; Arnold, M.; Bansal, S.; Fast, D.; Garofolo, F.; Lowes, S.; Timmerman, P.; Woolf, E. *Bioanalysis* **2010,** *2,* 1801–1803.
26. Premkumar, N.; Lowes, S.; Jersey, J.; Garofolo, F.; Dumont, I.; Masse, R.; Stamatiou, B.; Caturla, M. C.; Steffen, R.; Malone, M., et al. *Bioanalysis* **2010,** *2,* 1797–1800.
27. Savoie, N.; Booth, B. P.; Bradley, T.; Garofolo, F.; Hughes, N. C.; Hussain, S.; King, S. P.; Lindsay, M.; Lowes, S.; Ormsby, E., et al. *Bioanalysis* **2009,** *1,* 19–30.
28. Savoie, N.; Garofolo, F.; van Amsterdam, P.; Booth, B. P.; Fast, D. M.; Lindsay, M.; Lowes, S.; Masse, R.; Mawer, L.; Ormsby, E., et al. *Bioanalysis* **2010,** *2,* 53–68.
29. Savoie, N.; Garofolo, F.; van Amsterdam, P.; Bansal, S.; Beaver, C.; Bedford, P.; Booth, B. P.; Evans, C.; Jemal, M.; Lefebvre, M., et al. *Bioanalysis* **2010,** *2,* 1945–1960.
30. Garofolo, F.; Rocci, M. L., Jr.; Dumont, I.; Martinez, S.; Lowes, S.; Woolf, E.; van Amsterdam, P.; Bansal, S.; Gomes Barra, A.; Bauer, R., et al. *Bioanalysis* **2011,** *3,* 2081–2096.
31. Timmerman, P.; Anders Kall, M.; Gordon, B.; Laakso, S.; Freisleben, A.; Hucker, R. *Bioanalysis* **2010,** *2,* 1185–1194.
32. Freisleben, A.; Brudny-Klöppel, M.; Mulder, H.; de Vries, R.; de Zwart, M.; Timmerman, P. *Bioanalysis* **2011,** *3,* 1333–1336.
33. Timmerman, P.; Luedtke, S.; van Amsterdam, P.; Brudny-Kloeppel, M.; Lausecker, B.; Fischmann, S.; Globig, S.; Sennbro, C.; Jansat, J. M.; Mulder, H., et al. *Bioanalysis* **2009,** *1,* 1049–1056.
34. Lowes, S.; Jersey, J.; Shoup, R.; Garofolo, F.; Savoie, N.; Mortz, E.; Needham, S.; Caturla, M. C.; Steffen, R.; Sheldon, C., et al. *Bioanalysis* **2011,** *3,* 1323–1332.
35. Good Laboratory Practice for Nonclinical Laboratory Studies, Code of Federal Regulations, Title 21, Chapter I, Subchapter A, Part 58.

36. Analytical Methods for an in vivo Bioavailability or Bioequivalence Study, Code of Federal Regulations, Title 21, Volume 5, Chapter I, Subchapter D, Section 320.29 (a).
37. EMA/INS/GCP/532137/2010, Reflection Paper on Guidance for Laboratories that Perform the Analysis or Evaluation of Clinical Trial Samples, London, UK, February 2012.
38. a. Kaiser, H. *Anal. Chem.* **1970,** *42,* 24A–38A. b. Kaiser, H. *Anal. Chem.* **1970,** *42,* 26A–58A.
39. Gu, H.; Wang, J.; Aubry, A. F.; Jiang, H.; Zeng, J.; Easter, J.; Wang, J. S.; Dockens, R.; Bifano, M.; Burrell, R., et al. *Anal. Chem.* **2012,** *84,* 4844–4850.
40. International Organization for Standardization, in Accuracy (Trueness and Precision) of Measurement Methods and Results, ISO 5725-1 and 5725-3, 1994.
41. Heizman, P.; Zinapold, K.; Geshke, R. *Methodol. Surv. Biochem. Anal.* **1994,** *23,* 351–357.
42. Fung, E. N.; Zheng, N.; Arnold, M. E.; Zeng, J. *Bioanalysis* **2010,** *4,* 733–743.
43. Scott, D. O.; Bindra, D. S.; Stella, V. J. *Pharm. Res.* **1993,** *10,* 1451–1457.
44. Jersey, J. A.; Duyan, S. A.; Davis, I. M. *Pharm. Res.* **1994,** *11,* S-58.
45. Dams, R.; Huestis, M. A.; Lambert, W. E.; Murphy, C. M. *J. Am. Soc. Mass Spectrom.* **2003,** 1290–1294.
46. Xia, Y.; Jemal, M. *Rapid Commun. Mass Spectrom.* **2009,** *23,* 2125–2138.
47. Liang, Z. *Bioanalysis* **2012,** *4,* 1227–1234.
48. Silvester, S.; Smith, L. *Bioanalysis* **2012,** *4,* 879–895.
49. Wu, S. T.; Schoener, D.; Jemal, M. *Rapid Commun. Mass Spectrom.* **2008,** *22,* 2873–2881.
50. Matuszewski, B. K. *J. Chromatogr. B* **2006,** *830,* 293–300.
51. De Nardi, C.; Bonelli, F. *Rapid Commun. Mass Spectrom.* **2006,** *20,* 2709–2916.
52. Lang, J. R.; Bolton, S. *J. Pharm. Biomed. Anal.* **1991,** *9,* 357–361.
53. Lang, J. R.; Bolton, S. *J. Pharm. Biomed. Anal.* **1991,** *9,* 435–442.
54. Kringle, R. *Pharm. Res.* **1994,** *11,* 556–560.
55. Hoffman, D.; Kringle, R. *Pharm. Res.* **2007,** *24,* 1157–1164.
56. Hartmann, C.; Massart, D. L.; McDowall, R. D. *J. Pharm. Biomed. Anal.* **1994,** *12,* 1337–1343.
57. Thompson, M.; Howarth, R. J. *Analyst* **1980,** *105,* 1188–1195.
58. Kimanani, E. K. *J. Pharm. Biomed. Anal.* **1998,** *16,* 1117–1124.
59. Kimanani, E. K.; Lavigne, J. *J. Pharm. Biomed. Anal.* **1998,** *16,* 1107–1115.
60. Singtoroj, T.; Tarning, J.; Annerberg, A.; Ashton, M.; Bergqvist, Y.; White, N. J.; Lindegardh, N.; Day, N. P. J. *J. Pharm. Biomed. Anal.* **2005,** *11,* 219–227.
61. Thompson, M. *Analyst* **1982,** *107,* 1169–1180.
62. Tholen, D. W. *Arch. Pathol. Lab. Med.* **1992,** *116,* 746–756.
63. Karnes, H. T.; March, C. *J. Pharm. Biomed. Anal.* **1991,** *9,* 911–918.
64. Krouwer, J. S.; Schlain, B. *Clin. Chem.* **1993,** *39,* 1689–1693.
65. Hinderling, P. H. *Pharmacol. Rev.* **1997,** *49,* 279–295.
66. Fisar, Z.; Fuksová, K.; Sikora, J.; Kalisová, L.; Velenovská, M.; Novotna, M. *Neuro. Endocrinol. Lett.* **2006,** *27,* 307–313.
67. Ke, J.; Yancey, M.; Zhang, S.; Lowes, S.; Henion, J. *J. Chromatogr. B Biomed. Sci. Appl.* **2000,** *9,* 369–380.
68. Fura, A.; Harper, T. W.; Zhang, H.; Fung, L.; Shyu, W. C. *J. Pharm. Biomed. Anal.* **2003,** *14,* 513–522.
69. Morris, D. M.; Selinger, K. *J. Pharm. Biomed. Anal.* **1994,** *12,* 255–264.
70. Lentner, C., Ed. Ciba-Geigy Ltd: Basel, Switzerland, 1984.
71. Hughes, N. C.; Bajaj, N.; Fan, J.; Wong, E. Y. *Bioanalysis* **2009,** *1,* 1057–1066.
72. *Guide for Validation of Analytical and Bioanalytical Methods*. Resolution-RE n. 899; ANVISA, May 29, 2003.
73. Wang, S.; Cyronak, M.; Yang, E. *J. Pharm. Biomed. Anal.* **2007,** *17,* 701–707.
74. Vander Heydena, Y.; Nijhuisb, A.; Smeyers-Verbekea, J.; Vandeginsteb, B.G.M.; Massart, D.L. Guidance for Robustness/Ruggedness Tests in Method Validation, http://www.vub.ac.be/fabi/tutorial/guideline.pdf.

75. Smith, K. M.; Yan, X. *Bioanalysis* **2012,** *4,* 741–749.
76. Zhang, J.; Reimer, M. T.; Ji, Q. C.; Chang, M. S.; El-Shourbagy, T. A.; Burke, S.; Schwartz, L. *Anal. Bioanal. Chem.* **2007,** *387,* 2745–2756.
77. Ji, Q. C.; Zhang, J.; Rodila, R.; Watson, P.; El-Shourbagy, T. *Rapid Commun. Mass Spectrom.* **2004,** *18,* 2293–2298.
78. Guthrue, R.; Suzi, A. *Pediatrics* **1963,** *23,* 338–343.
79. Evans, C. Current Technology and Use of Dried Blood Spots, 12th Annual Land O'Lakes Bioanalytical Conference, Merrimac, WI, July 2011.
80. Needham, S. Method Development and Validation for Dried Blood Spots, 12th Annual Land O'Lakes Bioanalytical Conference, Merrimac, WI, July 2011.
81. Brewer, E. Special Analytical Challenges and Solutions for Implementation, 12th Annual Land O'Lakes Bioanalytical Conference, Merrimac, WI, July 2011.
82. Viswanathan, C. Regulatory Perspective on Dried Blood Spots, 12th Annual Land O'Lakes Bioanalytical Conference, Merrimac, WI, July 2011.
83. Ramanathan, I. R.; Jemal, M.; Ramagiri, S.; Xia, Y. Q.; Humpreys, W. G.; Olah, T.; Korfmacher, W. A. *J. Mass Spectrom.* **2011,** *46,* 595–601.
84. Xia, Y. Q.; Lau, J.; Olah, T.; Jemal, M. *Rapid Commun. Mass Spectrom.* **2011,** *15,* 2863–2878.
85. Fung, E. N.; Xia, Y. Q.; Aubry, A. F.; Zeng, J.; Olah, T.; Jemal, M. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2011,** *1* (879), 2919–2927.
86. Kaufmann, J. A.; Butcher, P.; Maden, K.; Walker, S.; Widmer, M. *Anal. Chim. Acta* **2011,** *700* (1–2), 86–94.
87. Berna, K. M. J.; Zhen, Y.; Watson, D. E.; Hale, J. E.; Ackermann, B. L. *Anal. Chem.* **2007,** *1* (79), 4199–4205.
88. Lu, Q.; Zheng, X.; McIntosh, T.; Davis, H.; Nemeth, J. F.; Pendley, C.; Wu, S. L.; Hancock, W. S. *Anal. Chem.* **2009,** *1* (81), 8715–8723.
89. Li, H.; Ortiz, R.; Tran, L.; Hall, M.; Spahr, C.; Walker, K.; Laudemann, J.; Miller, S.; Salimi-Moosavi, H.; Lee, J. W. *Anal. Chem.* **2012,** *7* (84), 1267–1273.
90. Xu, Y.; Mehl, J. T.; Bakhtiar, R.; Woolf, E. J. *Anal. Chem.* **2010,** *15* (82), 6877–6886.
91. Wu, OS. T.; Ouyang, Z.; Olah, T.; Jemal, M. *Rapid Commun. Mass Spectrom.* **2011,** *30* (25), 281–290.
92. Rauh, P. M. *J. Chromatogr. B* **2012,** *883–884,* 59–67.