EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

21 July 2011
EMEA/CHMP/EWP/192217/2009
Committee for Medicinal Products for Human Use (CHMP)

Guideline on bioanalytical method validation

Draft agreed by the Efficacy Working Party	September 2009
Adoption by CHMP for release for consultation	19 November 2009
End of consultation (deadline for comments)	31 May 2010
Agreed by Pharmacokinetics Working Party (PKWP)	June 2011
Adoption by CHMP	21 July 2011
Date for coming into effect	1 February 2012

Comments should be provided using this template. The completed comments form should be sent to PKWPsecretariat@ema.europa.eu.

Keywords	*CHMP, EMEA, Guideline, validation, bioanalytical method, analyses*

7 Westferry Circus ● Canary Wharf ● London E14 4HB ● United Kingdom
Telephone +44 (0)20 7418 8400 **Facsimile** +44 (0)20 7418 8613
E-mail info@ema.europa.eu **Website** www.ema.europa.eu

Guideline on bioanalytical method validation

Table of contents

Executive summary

This guideline defines key elements necessary for the validation of bioanalytical methods. The guideline focuses on the validation of the bioanalytical methods generating quantitative concentration data used for pharmacokinetic and toxicokinetic parameter determinations. Guidance and criteria are given on the application of these validated methods in the routine analysis of study samples from animal and human studies.

1. Introduction (background)

Measurement of drug concentrations in biological matrices (such as serum, plasma, blood, urine, and saliva) is an important aspect of medicinal product development. Such data may be required to support applications for new actives substances and generics as well as variations to authorised drug products. The results of animal toxicokinetic studies and of clinical trials, including bioequivalence studies are used to make critical decisions supporting the safety and efficacy of a medicinal drug substance or product. It is therefore paramount that the applied bioanalytical methods used are well characterised, fully validated and documented to a satisfactory standard in order to yield reliable results.

Acceptance criteria wider than those defined in this guideline may be used in special situations. This should be prospectively defined based on the intended use of the method.

2. Scope

This guideline provides recommendations for the validation of bioanalytical methods applied to measure drug concentrations in biological matrices obtained in animal toxicokinetic studies and all phases of clinical trials. As ligand binding assays differ substantially from chromatographic analytical methods, separate validation recommendations for ligand binding assays are provided.

In addition, specific aspects for the analysis of study samples will be addressed.

Furthermore, this guideline will describe when partial validation or cross validation should be carried out in addition to the full validation of an analytical method.

Methods used for determining quantitative concentrations of biomarkers used in assessing pharmacodynamic endpoints are out of the scope of this guideline.

3. Legal basis

This guideline has to be read in conjunction with the introduction and general principles (4) and Part I and II of the Annex I to Directive 2001/83 as amended. It applies to Marketing Authorisation Applications for human medicinal products submitted in accordance with the Directive 2001/83/EC as amended, and Regulation (EC) No. 726/2004, in which the analysis of drug concentrations in a biological matrix is part of the application.

The validation of bioanalytical methods and the analysis of study samples for clinical trials in humans should be performed following the principles of Good Clinical Practice (GCP). Further guidance that will help clinical laboratories develop and maintain quality systems which will comply with relevant European Union Directives, national regulations and associated guidance documents can be found in the “Reflection Paper for Laboratories That Perform The Analysis Or Evaluation Of Clinical Trial Samples.” (EMA/INS/GCP/532137/2010).

Non-clinical (pharmaco-toxicological) studies submitted in a marketing authorisation application shall be carried out in conformity with the provisions related to Good Laboratory Practice, Directive 2004/10/EC on the harmonisation of laws, regulations and administrative provisions relating to the application of the principles of good laboratory practice and the verification of their applications for tests on chemical substances and Directive 2004/9/EC on the inspection and verification of good laboratory practice (GLP). Normally, the validation of bioanalytical methods used in non-clinical pharmacotoxicological studies that are carried out in conformity with the provisions related to Good Laboratory Practice should be performed following the Principles of Good Laboratory Practice. Aspects of method validation not performed according to GLP should be clearly identified and their potential impact on the validation status of the method indicated. Methods used in pre-clinical studies not required to be performed to GLP should be fit for purpose but not necessarily developed in a GLP facility.

4. Method validation

4.1. Full validation of an analytical method

A full method validation should be performed for any analytical method whether new or based upon literature.

The main objective of method validation is to demonstrate the reliability of a particular method for the determination of an analyte concentration in a specific biological matrix, such as blood, serum, plasma, urine, or saliva. Moreover, if an anticoagulant is used, validation should be performed using the same anticoagulant as for the study samples. Generally a full validation should be performed for each species and matrix concerned.

In some cases, it may be problematic for validation purposes to obtain an identical matrix compared to the matrix of the study samples. A suitable alternative matrix may be used, e.g. synthetically prepared cerebrospinal fluid, if justified.

The main characteristics of a bioanalytical method that are essential to ensure the acceptability of the performance and the reliability of analytical results are: selectivity, lower limit of quantification, the response function and calibration range (calibration curve performance), accuracy, precision, matrix effects, stability of the analyte(s) in the biological matrix and stability of the analyte(s) and of the internal standard in the stock and working solutions and in extracts under the entire period of storage and processing conditions.

Usually one analyte or drug has to be determined, but on occasions it may be appropriate to measure more than one analyte. This may involve two different drugs, but can also involve a parent drug with its metabolites, or the enantiomers or isomers of a drug. In these cases the principles of validation and analysis apply to all analytes of interest.

Reference standards

During method validation and analysis of study samples, a blank biological matrix will be spiked with the analyte(s) of interest using solutions of reference standard(s) to prepare calibration standards, quality control samples and stability samples. In addition, suitable internal standard(s) (IS) can be added during sample processing in chromatographic methods.

It is important that the quality of the reference standard and IS is ensured, as the quality (purity) may affect the outcome of the analysis, and therefore the outcome of the study data. Therefore the reference standards used during the validation and study sample analysis should be obtained from an authentic and traceable source. Suitable reference standards, include certified standards such as

compendial standards (EPCRS, USP, WHO), commercially available standards, or sufficiently characterised standards prepared in-house or by an external non-commercial organisation. A certificate of analysis is required to ensure purity and provide information on storage conditions, expiration date and batch number of the reference standard.

The use of certified standards is not needed for IS, as long as the suitability for use is demonstrated, e.g. lack of analytical interference is shown for the substance itself or any impurities thereof. A certificate of analysis is not required.

When mass-spectrometry (MS) detection is used in the bioanalytical method, a stable isotope-labelled IS is recommended to be used whenever possible. However, it is essential that the labelled standard is of the highest isotope purity and that no isotope exchange reaction occurs. The presence of any unlabelled analyte should be checked and if relative amounts of unlabelled analyte are detected the potential influence has to be evaluated during method validation.

4.1.1. Selectivity

The analytical method should be able to differentiate the analyte(s) of interest and IS from endogenous components in the matrix or other components in the sample. Selectivity should be proved using at least 6 individual sources of the appropriate blank matrix, which are individually analysed and evaluated for interference. Use of fewer sources is acceptable in case of rare matrices. Normally, absence of interfering components is accepted where the response is less than 20% of the lower limit of quantification for the analyte and 5% for the internal standard.

It may also be necessary to investigate the extent of any interference caused by metabolites of the drug(s), interference from degradation products formed during sample preparation, and interference from possible co-administered medications. Co-medications normally used in the subject population studied which may potentially interfere should be taken into account at the stage of method validation, or on a study specific and compound specific base.

The possibility of back-conversion of a metabolite into parent analyte during the successive steps of the analysis (including extraction procedures or in the MS source) should also be evaluated, when relevant (i.e. potentially unstable metabolites e.g. acidic metabolites to ester, unstable N-oxides or glucuronide metabolites, lactone-ring structures). The extent of back-conversion should be established and the impact on the study results discussed. It is acknowledged that this evaluation will not be possible early during drug development of a new chemical entity when the metabolism is not yet evaluated. However, it is expected that this issue is taken into account and a partial validation is performed if relevant as further knowledge regarding metabolism of the active substance is gained during drug development.

It is recognized that in some cases it is very difficult to obtain the metabolites of interest. Alternatively, back-conversion of a metabolite can be checked by applying incurred sample reanalysis. However, in this case potential back conversion during sample processing cannot be ruled out.

4.1.2. Carry-over

Carry-over should be addressed and minimised during method development. During validation carry-over should be assessed by injecting blank samples after a high concentration sample or calibration standard at the upper limit of quantification. Carry over in the blank sample following the high concentration standard should not be greater than 20% of the lower limit of quantification (LLOQ; see below) and 5% for the internal standard. If it appears that carry-over is unavoidable, study samples should not be randomised. Specific measures should be considered, tested during the validation and

applied during the analysis of the study samples, so that it does not affect accuracy and precision. This could include the injection of blank samples after samples with an expected high concentration, before the analysis of the next study sample.

4.1.3. Lower limit of quantification

The lower limit of quantification (LLOQ) is the lowest concentration of analyte in a sample which can be quantified reliably, with an acceptable accuracy and precision. The LLOQ is considered being the lowest calibration standard (see Accuracy and Precision). In addition, the analyte signal of the LLOQ sample should be at least 5 times the signal of a blank sample. The LLOQ should be adapted to expected concentrations and to the aim of the study. As an example, for bioequivalence studies the LLOQ should be not higher than 5% of the Cmax, while such a low LLOQ may be not necessary for exploratory pharmacokinetic studies.

4.1.4. Calibration curve

The response of the instrument with regard to the concentration of analyte should be known, and should be evaluated over a specified concentration range. The calibration standards should be prepared in the same matrix as the matrix of the intended study samples by spiking the blank matrix with known concentrations of the analyte. There should be one calibration curve for each analyte studied in the method validation and for each analytical run.

Ideally, before carrying out the validation of the analytical method it should be known what concentration range is expected. This range should be covered by the calibration curve range, defined by the LLOQ being the lowest calibration standard and the upper limit of quantification (ULOQ), being the highest calibration standard. The range should be established to allow adequate description of the pharmacokinetics of the analyte of interest.

A minimum of six calibration concentration levels should be used, in addition to the blank sample (processed matrix sample without analyte and without IS) and a zero sample (processed matrix with IS). Each calibration standard can be analysed in replicate.

A relationship which can simply and adequately describe the response of the instrument with regard to the concentration of analyte should be applied. The blank and zero samples should not be taken into consideration to calculate the calibration curve parameters.

The calibration curve parameters should be reported (slope and intercept in case of linear fit). In addition, the back calculated concentrations of the calibration standards should be presented together with the calculated mean accuracy values (see definition of Accuracy below). All the available (or acceptable) curves obtained during validation, with a minimum of 3 should be reported.

The back calculated concentrations of the calibration standards should be within ±15% of the nominal value, except for the LLOQ for which it should be within ±20%. At least 75% of the calibration standards, with a minimum of six calibration standard levels, must fulfil this criterion. In case replicates are used, the criteria (within ±15% or ±20% for LLOQ) should also be fulfilled for at least 50% of the calibration standards tested per concentration level. In case a calibration standard does not comply with these criteria, this calibration standard sample should be rejected, and the calibration curve without this calibration standard should be re-evaluated, including regression analysis. In case all replicates of the LLOQ or the ULOQ calibration standard are rejected then the batch should be rejected from the validation, the possible source of the failure be determined and the method revised (if necessary). If the next validation batch also fails, then the method should be revised before restarting validation.

Although the calibration curve should preferably be prepared using freshly spiked samples, it is allowed to use previously prepared and stored calibration samples, if supported by appropriate stability data.

4.1.5. Accuracy

The accuracy of an analytical method describes the closeness of the determined value obtained by the method to the nominal concentration of the analyte (expressed in percentage). Accuracy should be assessed on samples spiked with known amounts of the analyte, the quality control samples (QC samples). The QC samples should be spiked independently from the calibration standards, using separately prepared stock solutions, unless the nominal concentration(s) of the stock solutions have been established.

The QC samples are analysed against the calibration curve, and the obtained concentrations are compared with the nominal value. The accuracy should be reported as percent of the nominal value. Accuracy should be evaluated for the values of the QC samples obtained within a single run (the within run accuracy) and in different runs (the between-run accuracy).

To enable evaluation of any trends over time within one run, it is recommended to demonstrate accuracy and precision of QC samples over at least one of the runs in a size equivalent to a prospective analytical run of study samples.

Within-run accuracy

Within-run accuracy should be determined by analysing in a single run a minimum of 5 samples per level at a minimum of 4 concentration levels which are covering the calibration curve range: the LLOQ, within three times the LLOQ (low QC), around 50% of the calibration curve range (medium QC), and at least at 75% of the upper calibration curve range (high QC). The mean concentration should be within 15% of the nominal values for the QC samples, except for the LLOQ which should be within 20% of the nominal value.

Between –run accuracy

For the validation of the between-run accuracy, LLOQ, low, medium and high QC samples from at least three runs analysed on at least two different days should be evaluated. The mean concentration should be within 15% of the nominal values for the QC samples, except for the LLOQ which should be within 20% of the nominal value.

Reported method validation data and the determination of accuracy and precision should include all results obtained except those cases where errors are obvious and documented.

4.1.6. Precision

The precision of the analytical method describes the closeness of repeated individual measures of analyte. Precision is expressed as the coefficient of variation (CV). Precision should be demonstrated for the LLOQ, low, medium and high QC samples, within a single run and between different runs, i.e. using the same runs and data as for the demonstration of accuracy.

Within-run precision

For the validation of the within-run precision, there should be a minimum of five samples per concentration level at LLOQ, low, medium and high QC samples in a single run. The within-run CV value should not exceed 15% for the QC samples, except for the LLOQ which should not exceed 20%.

Between –run precision

For the validation of the between-run precision, LLOQ, low, medium and high QC samples from at least three runs analysed on at least two different days should be evaluated. The between-run CV value should not exceed 15% for the QC samples, except for the LLOQ which should not exceed 20%.

4.1.7. Dilution integrity

Dilution of samples should not affect the accuracy and precision. If applicable, dilution integrity should be demonstrated by spiking the matrix with an analyte concentration above the ULOQ and diluting this sample with blank matrix (at least five determinations per dilution factor). Accuracy and precision should be within the set criteria, i.e. within ±15%. Dilution integrity should cover the dilution applied to the study samples.

Evaluation of dilution integrity may be covered by partial validation. Use of another matrix may be acceptable, as long as it has been demonstrated that this does not affect precision and accuracy.

4.1.8. Matrix effect

Matrix effects should be investigated when using mass spectrometric methods, using at least 6 lots of blank matrix from individual donors. Pooled matrix should not be used.

For each analyte and the IS, the matrix factor (MF) should be calculated for each lot of matrix, by calculating the ratio of the peak area in the presence of matrix (measured by analysing blank matrix spiked after extraction with analyte), to the peak area in absence of matrix (pure solution of the analyte). The IS normalised MF should also be calculated by dividing the MF of the analyte by the MF of the IS. The CV of the IS-normalised MF calculated from the 6 lots of matrix should not be greater than 15 %. This determination should be done at a low and at a high level of concentration (maximum of 3 times the LLOQ and close to the ULOQ).

If this approach cannot be used, for instance in the case of on-line sample preparation, the variability of the response from lot to lot should be assessed by analysing at least 6 lots of matrix, spiked at a low and at a high level of concentration (maximum of 3 times the LLOQ and close to the ULOQ). The validation report should include the peak areas of the analyte and of the IS and the calculated concentration for each individual sample. The overall CV calculated for the concentration should not be greater than 15 %.

If the matrix is difficult to obtain, less than 6 different lots of matrix may be used, but this should be justified. However, matrix effects should still be investigated.

If a formulation for injection to be administered to the subjects or animals contains excipients known to be responsible for matrix effects, for instance polyethylene glycol or polysorbate, matrix effects should be studied with matrix containing these excipients, in addition to blank matrix. The matrix used for this evaluation should be obtained from subjects or animals administered the excipient, unless it has been demonstrated that the excipient is not metabolised or transformed *in-vivo*. The effect of the excipients can be studied by the determination of the MF or by a dilution study of a study sample with a high concentration with blank matrix not containing the excipient.

In addition to the normal matrix it is recommended to investigate matrix effects on other samples e.g. haemolysed and hyperlipidaemic plasma samples. If samples from special populations (such as renally or hepatically impaired populations) are to be analysed it is also recommended to study matrix effects using matrix from such populations.

4.1.9. Stability

Evaluation of stability should be carried out to ensure that every step taken during sample preparation and sample analysis, as well as the storage conditions used do not affect the concentration of the analyte.

Stability should be ensured for every step in the analytical method, meaning that the conditions applied to the stability tests, such as sample matrix, anticoagulant, container materials, storage and analytical conditions should be similar to those used for the actual study samples. Reference to data published in the literature is not considered sufficient.

Stability of the analyte in the studied matrix is evaluated using low and high QC samples (blank matrix spiked with analyte at a concentration of a maximum of 3 times the LLOQ and close to the ULOQ) which are analysed immediately after preparation and after the applied storage conditions that are to be evaluated. The QC samples are analysed against a calibration curve, obtained from freshly spiked calibration standards, and the obtained concentrations are compared to the nominal concentrations. The mean concentration at each level should be within ±15% of the nominal concentration.

Stability of the stock and working solutions should be tested with an appropriate dilution, taking into consideration the linearity and measuring range of the detector.

Stability studies should investigate the different storage conditions over time periods that equal or exceed those applied to the actual study samples.

The following stability tests should be evaluated:

- stability of the stock solution and working solutions of the analyte and internal standard,
- freeze and thaw stability of the analyte in the matrix from freezer storage conditions to room temperature or sample processing temperature,
- short term stability of the analyte in matrix at room temperature or sample processing temperature,
- long term stability of the analyte in matrix stored in the freezer,

In addition the following tests should be performed if applicable:

- stability of the processed sample at room temperature or under the storage conditions to be used during the study (dry extract or in the injection phase),
- on-instrument/ autosampler stability of the processed sample at injector or autosampler temperature.

Regarding the freeze and thaw stability: The QC samples are stored and frozen in the freezer at the intended temperature and thereafter thawed at room or processing temperature. After complete thawing, samples are refrozen again applying the same conditions. At each cycle, samples should be frozen for at least 12 hours before they are thawed. The number of cycles in the freeze-thaw stability should equal or exceed that of the freeze/thaw cycles of study samples.

Regarding long term stability of the analyte in matrix stored in the freezer: The QC samples should be stored in the freezer under the same storage conditions and at least for the same duration as the study samples. For small molecules it is considered acceptable to apply a bracketing approach, i.e. in case stability has been proved for instance at -70°C and -20°C, it is not necessary to investigate the stability at temperatures in between. For large molecules (such as peptides and proteins) stability should be studied at each temperature at which study samples will be stored. Study samples may be used in addition to QC samples, but the exclusive use of study samples is not considered sufficient as

the nominal concentrations of those samples is not known. The results of the evaluation of long term stability should be available before the study report is issued.

Regarding the stability of stock and working solutions: It is not needed to study the stability at each concentration level of working solutions and a bracketing approach can be used. It is not needed to study the stability of stable-isotope labelled internal standards if it is demonstrated that no isotope exchange reactions occur under the same conditions as the stability of the analyte was demonstrated.

In case of a multi-analyte study and specific for bioequivalence studies, attention should be paid to stability of the analytes in the matrix containing all the analytes.

Sufficient attention should be paid to the stability of the analyte in the sampled matrix directly after blood sampling of subjects and further preparation before storage, to ensure that the obtained concentrations by the analytical method reflect the concentrations of the analyte in the subject at the moment of sampling. A demonstration of this stability may be needed on a case-by-case basis, depending on the structure of the analyte.

4.2. Partial validation

In situations where minor changes are made to an analytical method that has already been validated, a full validation may not be necessary, depending on the nature of the applied changes. Changes for which a partial validation may be needed include transfer of the bioanalytical method to another laboratory, change in equipment, calibration concentration range, limited sample volume, another matrix or species, change in anticoagulant, sample processing procedure, storage conditions etc. All modifications should be reported and the scope of revalidation or partial validation justified.

Partial validation can range from as little as the determination of the within-run precision and accuracy, to an almost full validation.

4.3. Cross validation

Where data are obtained from different methods within and across studies or when data are obtained within a study from different laboratories, applying the same method, comparison of those data is needed and a cross validation of the applied analytical methods should be carried out. Differences in sample preparation or the use of another analytical method may result in different outcomes between the study sites. Cross validation should be performed in advance of study samples being analysed if possible. For the cross validation, the same set of QC samples or study samples should be analysed by both analytical methods. For QC samples, the obtained mean accuracy by the different methods should be within 15% and may be wider, if justified. For study samples, the difference between the two values obtained should be within 20% of the mean for at least 67% of the repeats. The outcome of the cross validation is critical in determining whether the obtained data are reliable and whether they can be compared and used.

5. Analysis of study samples

After full validation of the analytical method, analysis of study or subject samples can be carried out. Before start of the analysis of the study samples the performance of the bioanalytical method should have been verified.

The study samples, QC samples and calibration standards should be processed in accordance with the validated analytical method to ensure the acceptability of the analytical run.

5.1. Analytical run

An analytical run consists of the blank sample (processed matrix sample without analyte and without IS) and a zero sample (processed matrix with IS), calibration standards at a minimum of 6 concentration levels, at least 3 levels of QC samples (low, medium and high) in duplicate (or at least 5 % of the number of study samples, whichever is higher), and study samples to be analysed. As indicated before the calibration standards and QC samples should have been spiked independently using separately prepared stock solutions, unless the nominal concentration(s) of the stock solutions have been established. All samples (calibration standards, QC samples, and study samples) should be processed and extracted as one single batch of samples in the order in which they intend to be submitted or analysed. A single batch is comprised of samples which are handled at the same time, i.e. subsequently processed without interruption in time and by the same analyst with the same reagents under homogeneous conditions. Analysing samples, which were prepared separately as several batches, in a single analytical run should be avoided. If such an approach cannot be avoided, for instance due to bench-top stability limitations, each batch of samples should include low, medium and high QC samples. Acceptance criteria should be pre-established in a Standard Operating Procedure (SOP) or in the study plan and should be defined for the whole analytical run and the separate batches in the run.

For bioequivalence studies it is advised to analyse all samples of one subject together in one analytical run to reduce the variability in outcome. The QC samples should be divided over the run in such a way that the accuracy and precision of the whole run is ensured.

5.2. Acceptance criteria of an analytical run

Criteria for acceptance or rejection of an analytical run should be defined in the protocol, in the study plan or in a SOP. In case a whole run consist of more batches, acceptance criteria should be applied to the whole run and to the individual batches. The run can be acceptable, although a batch might have to be rejected, as criteria were not met.

The following acceptance criteria should apply:

Accuracy:

The back calculated concentrations of the calibration standards should be within ±15% of the nominal value, except for the LLOQ for which it should be within ±20%. At least 75% of the calibration standards, with a minimum of six, must fulfil this criterion. If one of the calibration standards does not meet these criteria, this calibration standard should be rejected and the calibration curve without this calibration standard should be re-evaluated, and regression analysis performed.

If the rejected calibration standard is the LLOQ, the LLOQ for this analytical run is the next lowest acceptable calibration standard of the calibration curve. If the highest calibration standard is rejected, the ULOQ for this analytical run is the next acceptable lower calibration standard of the calibration curve. The revised calibration range must cover all QC samples (low, medium and high).

The accuracy values of the QC samples should be within ±15% of the nominal values. At least 67% of the QC samples and at least 50% at each concentration level should comply with this criterion. In case these criteria are not fulfilled the analytical run should be rejected, and the study samples re-extracted and analysed.

In the case of the simultaneous determination of several analytes, there should be one calibration curve for each analyte studied. If an analytical run is acceptable for one analyte but has to be rejected for another analyte, the data for the accepted analyte can be used, but the samples should be re-extracted and analysed for determination of the rejected analyte.

If replicate calibration standards are used and only one of the LLOQ or ULOQ standards fails, the calibration range is unchanged.

The overall (mean) accuracy and precision of the QC samples of all accepted runs should be calculated at each concentration level and reported in the analytical report. In case the overall mean accuracy and precision exceeds 15%, this should lead to additional investigations justifying this deviation. In the case of bioequivalence trials it may result in the rejection of the data.

5.3. Calibration range

If a narrow range of analyte concentrations of the study samples is known or anticipated before the start of study sample analysis, it is recommended to either narrow the calibration curve range, adapt the concentrations of the QC samples, or add new QC samples at different concentration levels as appropriate, to adequately reflect the concentrations of the study samples.

If a narrow range of analysis values is unanticipated, but observed after the start of sample analysis, it is recommended that the analysis is stopped and either the standard calibration range narrowed, existing QC concentrations revised, or QC samples at additional concentrations are added to the original curve before continuing with study sample analysis. It is not necessary to reanalyse samples analysed before optimising the standard curve range or QC concentrations.

The same applies if it appears that a large number of the analyte concentrations of the study samples appear to be above the ULOQ. The calibration curve range should be extended, if possible, and QC samples added or their concentrations modified.

At least 2 QC sample levels should fall within the range of concentrations measured in study samples. If the calibration curve range is changed, the bioanalytical method should be revalidated (partial validation) to verify the response function and to ensure accuracy and precision.

5.4. Reanalysis of study samples

Possible reasons for reanalysis of study samples and criteria to select the value to be reported should be predefined in the protocol, study plan or SOP, before the actual start of the analysis of the samples. The number of samples (and percentage of total number of samples) that have been reanalysed should be discussed in the study report.

The following are examples of reasons for study sample reanalysis:

- rejection of an analytical run because the run did not fulfil the acceptance criteria with regard to accuracy of the calibration standards and/or the QC samples,
- internal standard response significantly different from the response for the calibration standard and QC samples, if such criteria have been pre-defined in a SOP,
- improper sample injection or malfunction of equipment,
- the obtained concentration is above the ULOQ or below the run's LLOQ, in runs where the lowest standard sample has been rejected from a calibration curve, resulting in a higher LLOQ compared with other runs,
- identification of quantifiable analyte levels in pre-dose samples or placebo sample,
- poor chromatography.

For bioequivalence studies, normally reanalysis of study samples because of a pharmacokinetic reason is not acceptable, as this may affect and bias the outcome of such a study. In this case, reanalysis

might be considered as part of laboratory investigations, to identify possible reasons for results considered as abnormal and to prevent the recurrence of similar problems in the future.

In case of reanalysis because of positive pre-dose samples or because of a pharmacokinetic reason, the reanalysed samples should be identified and the initial value, the reason for reanalysis, the values obtained in the reanalyses, the finally accepted value and a justification for the acceptance should be provided.

Re-injection of samples can be made in case of instrument failure if reinjection reproducibility and on-injector stability have been demonstrated during validation. Re-injection of a full analytical run or of individual calibration standard samples or QC samples, simply because the calibration or QCs failed, without any identified analytical cause, is not acceptable.

The safety of trial subjects should take precedence over any other aspect of the trial. Consequently, there may be other circumstances when it is necessary to re-extract and/or re-analyse specific study samples, for example where an unexpected or anomalous result is identified that may impact on patient safety.

5.5. Integration

Chromatogram integration and re-integration should be described in a SOP. Any deviation from this SOP should be discussed in the analytical report. Chromatogram integration parameters and in case of re-integration, initial and the final integration data should be documented at the laboratory and should be available upon request. For further guidance reference is made to the “Reflection Paper for Laboratories That Perform The Analysis Or Evaluation Of Clinical Trial Samples.” (EMA/INS/GCP/532137/2010).

6. Incurred samples reanalysis

The use of calibration standards and QC samples during validation may not mimic the actual study samples. Differences for instance in protein binding, back-conversion of known and unknown metabolites, sample inhomogeneity or concomitant medications, may affect the accuracy and precision of the analyte in such samples during processing and storage. It is therefore recommended to evaluate accuracy of incurred samples by reanalysis of study samples in separate runs at different days. The extent of testing depends on the analyte and the study samples, and should be based upon in-depth understanding of the analytical method and analyte. However, as a guide, 10% of the samples should be reanalysed in case the number of samples is less than 1000 samples and 5% of the number of samples exceeding 1000 samples. Furthermore, it is advised to obtain samples around Cmax and in the elimination phase.

The concentration obtained for the initial analysis and the concentration obtained by reanalysis should be within 20% of their mean for at least 67% of the repeats. Large differences between results may indicate analytical issues and should be investigated.

In case incurred sample analysis showed deviating results, this should be investigated, and adequate steps should be taken to minimize inaccuracy (and imprecision).

Incurred sample reanalysis should be done at least in the following situations:

- toxicokinetic studies once per species
- all pivotal bioequivalence trials
- first clinical trial in subjects

- first patient trial
- first trial in patients with impaired hepatic and/or renal function

For animal studies, the incurred sample reanalysis may be done only in early Phase studies, if these are representative for pivotal studies in terms of dose administered and concentrations obtained.

Samples should not be pooled, as pooling may limit anomalous findings.

7. Ligand binding assays

7.1 *Method validation*

Ligand-binding assays (LBA) or immunoassays are especially used for macromolecules. The validation principles and the considerations with regard to analysis of study samples, as indicated before should also be applied in general for ligand-binding assays. However ligand binding assays pose several challenges. Due to the inherent characteristics and complex structure of the macromolecules, the extraction process is problematic and as such these assays are often run without prior separation of the analyte of interest. In addition these assays do not directly measure the macromolecule itself but indirectly measure a binding reaction with reagents employed in the assay. For these reasons, several issues need special attention.

7.1.1. Full validation

7.1.1.1. Reference standards

Macromolecules are heterogeneous and their potency and immunoreactivity may vary. The reference material should be well characterised and documented (e.g. certificate of analysis and origin).The purest reference standard available at the time should be procured. It is strongly recommended that the batch of the reference standard used for the preparation of calibration standards and QC samples is the same as used for dosing in the non clinical and clinical studies. In case of change of batch, an analytical characterisation and bioanalytical evaluation should be carried out prior to use to ensure that the performance characteristics of the method are not altered.

7.1.1.2. Specificity

Specificity of the binding reagent(s) refer(s) to its (their) ability to bind solely to the analyte of interest. Specificity is related to the concept of cross-reactivity. Ideally the binding reagent should be specific such that no cross-reactivity occurs with structurally “related compounds” (e.g. endogenous compounds, isoforms, variants forms of the analyte, or physico-chemically similar compounds) or with anticipated concomitant medication. During method development and validation, frequently these “related molecules” are not available. Evaluation of specificity may be conducted after the original validation is completed as more data on the behaviour of the analyte become available. Specificity should be tested with QC samples by adding increasing concentrations of available “related molecules” or drugs expected to be concomitantly administered, into drug-naive sample matrix (matrix obtained from animals or subjects never exposed to the analyte) and measuring the accuracy of the macromolecule of interest at both LLOQ and ULOQ. The assay acceptance criteria of the QC samples should be within 25% of the nominal values.

7.1.1.3. Selectivity

Selectivity of a ligand-binding assay is the ability to measure the analyte of interest in the presence of unrelated compounds in the matrix. Generally there is no extraction due to the inherent characteristics of macromolecules. Then, unrelated compounds present in matrix e.g. degrading enzymes, heterophilic antibodies or rheumatoid factor, may interfere with the analyte of interest in the ligand binding assay. Selectivity is tested by spiking at least 10 sources of sample matrix at or near the LLOQ. These sources should include lipemic and haemolysed samples. It is also strongly recommended that sources from relevant disease population be included. Selectivity should be evaluated at the low end of an assay where problems occur in most cases. It may be prudent also to evaluate selectivity at higher analyte concentrations. In cases where interference is concentration dependent, it is essential to determine the minimum concentration where interference occurs. It may be necessary to adjust the lower level of quantification accordingly, before assay validation. The accuracy should be within 20% (25% at the LLOQ) of the nominal spiked concentration in at least 80% of the matrices evaluated.

7.1.1.4. Carry-over effect

If robotic liquid handling systems are used, potential for carry-over should be investigated by placing blank samples after samples with a high analyte concentration or calibration standard at the upper limit of quantification.

7.1.1.5. Matrix selection

The measurement of some macromolecules may not be possible in complex matrices without extraction due to high interferences with high levels of structurally related endogenous compounds. Although the use of extracted matrix (e.g. charcoal, immuno-affinity) or alternative matrix (e.g. protein buffers, dialysed serum) is not recommended, the use of such matrices may be necessary when there is no other strategy to quantify the analyte of interest. The calibration standard curve may be prepared in these surrogate matrices. QC samples should be prepared in the actual sample matrix and the accuracy should be calculated to demonstrate the absence of matrix effect.

7.1.1.6. Minimum required dilution

Because matrices may exhibit a high background signal, it may be necessary to determine the minimum required dilution. The minimum required dilution is the smallest dilution to which a sample must be diluted in buffer to optimize accuracy and precision in an assay run by reducing the signal to noise ratio. Spiked samples should be prepared in the same matrix as the study samples for determination of the minimum required dilution.

7.1.1.7. Calibration curve

The response function of the calibration curve is measured indirectly and is generally non linear and often sigmoidal. A minimum of 6 calibration standards should be run at least in duplicate. The calibration standards should be spaced approximately evenly on a logarithmic scale within the anticipated range. In addition to the calibration standards, anchor points outside the range of quantification can be used to facilitate the fitting of the curve. A minimum of 6 independent runs should be evaluated during the validation. The results must be reported in a table to establish the overall robustness of the regression model of the calibration curve. A calibration standard may be excluded from the curve due to a technical error with an assignable cause (e.g. pipetting error).

The target back-calculated concentrations of the calibration standards should be within 20% of the nominal value (25% at LLOQ and ULOQ) for at least 75% of calibration standards. The anchor calibrators do not require acceptance criteria since they are beyond the quantifiable range of the curve.

7.1.1.8. Precision and accuracy

For the estimation of precision and accuracy QC samples should not be freshly prepared, but should be frozen and treated the same way as for the analysis of study samples. At least 5 QC samples (anticipated LLOQ, less than 3 times the LLOQ, mid, high and anticipated ULOQ) should be used to assess accuracy, precision and the total error of the method. Validation should mimic the actual study samples analysis, i.e. in case a study sample is measured twice (i.e. using 2 wells) as recommended then during validation QCs should be analysed twice (i.e. using 2 wells per QC sample). Measurements should be made across at least 6 independent assay runs over several days. Regarding within-run and between-run accuracy, the mean concentration should be within 20% of the nominal value at each concentration level (25% at the LLOQ and ULOQ). The within-run and between-run precision should not exceed 20% (25 % at LLOQ and ULOQ). Furthermore the total error (i.e. sum of absolute value of the % relative error and % coefficient of variation) should not exceed 30% (40% at LLOQ and ULOQ).

7.1.1.9. Dilutional linearity

Because the narrow range of the calibration standard curve, it is necessary to demonstrate with QC samples that the analyte of interest, when present in concentrations exceeding the range of quantification (above ULOQ), can be accurately measured by the assay after dilution in blank matrix to bring the analyte concentrations into the validated range for analysis. An additional reason for conducting dilutional experiments is to detect a possible prozone or “hook effect” i.e. a signal suppression caused by high concentrations of analyte. The back-calculated concentration for each dilution should be within 20% of the nominal concentration after correction for dilution and the precision of the final concentrations across all the dilutions should not exceed 20%.

7.1.1.10. Parallelism

If study samples are available, parallelism between the calibration standard curve and serially diluted study samples should be assessed to detect possible matrix effect or differing affinities for metabolites. A high concentration study sample (preferably close to Cmax) should be diluted to at least three concentrations with blank matrix. The precision between samples in a dilution series should not exceed 30%. In case the sample does not dilute linearly (i.e. in a non parallel manner), a procedure for reporting a result should be defined a priori. If study samples are not available during the validation of the method, parallelism should be evaluated as soon as study samples become available.

7.1.1.11. Stability of the samples

Stability of the analyte is evaluated using samples of the low and high level QC samples as described before (section 4.1.9). As previously mentioned, the investigation of stability should cover short-term stability at room temperature or sample processing temperature and freeze-thaw stability. In addition, long-term freezer stability should be studied at each temperature at which study samples will be stored. A bracketing approach may be considered.

The mean concentration at each level should be within 20% of the nominal concentration.

7.1.1.12. Reagents

Critical reagents, including binding reagents (e.g. binding proteins, aptamers, antibodies or conjugated antibodies) and those containing enzymatic moieties have direct impact on the results of the assay and therefore their quality must be assured. Accordingly, when changing reagent batches during validation or sample analysis the analytical performance of the method must be verified to ensure that it is not altered compared with the original or previous batch.

Conditions guaranteeing the maintenance of the stability of both non critical reagents (e.g. buffers, diluents or acidification reagents) and more importantly of the critical reagents should be documented in order to ensure that the performance of the method is not affected over time.

7.1.1.13. Commercial kits

Commercial kits may have been developed for purposes other than to support pharmacokinetics. Therefore, commercial kits need to be revalidated to ensure that the LLOQ and the QC samples in the actual concentration range to be used for sample analysis perform accurately and precisely. The principles of validation listed above apply.

7.2. Partial validation and cross-validation

All the validation aspects reported in previous sections 4.2 and 4.3 are applicable to ligand binding assays.

7.3. Analysis of study samples

7.3.1. Analytical run

Most often microtiter plates are used for LBA. An analytical run may comprise several individual plates, but each plate should contain an individual set of calibration standards and QC samples to compensate for difference in plate performance. The sample capacity in some platforms may be limited. Then, it may be acceptable that a set of calibration standards be placed in the first and the last platform and QC samples on every single platform.

It is recommended to assay a study sample in replicate, i.e by using at least 2 wells instead of 1.

7.3.2. .Acceptance criteria for study sample analysis

The back calculated concentrations of the calibration standards should be within 20% of nominal value, except for LLOQ and the ULOQ for which it should be within 25%. At least 75 % of the calibration standards with a minimum of 6, must fulfil this criterion. This requirement does not apply to anchor calibrators.

Each plate should contain at least 3 levels of QC samples (low, medium and high) at least in duplicate. Also during within study validation, the QCs should mimic the analysis of the study sample with regard to the number of wells used per study sample. At least 67% QC samples and 50% at each concentration level should be within 20% of the nominal value. Exceptions to this criterion should be justified.

7.3.3. Incurred samples reanalysis

All the considerations regarding the incurred sample analysis reported in previous section 6 are applicable to ligand binding assays. The concentration obtained for the initial analysis and the concentration obtained by reanalysis should be within 30% of their mean for at least 67% of the repeats.

8. Reports

Information regarding conducted audits/inspection should be included in the report(s).

8.1. Validation report

Depending on the level of detail of the information provided in the validation report, reference to the SOPs for relevant analysis specific procedures may be sufficient. Otherwise these SOPs should be appended to the report.

All source data should be available in its original format and available on request.

Any deviation from the validation protocol should be recorded.

The validation report should include at least the following information:

- summary of the validation performances,
- details of the applied analytical method and where appropriate, the source of the analytical method (references from literature and/or modifications in the procedure),
- details of the assay procedure (analyte, IS, sample pre-treatment, extraction and analysis),
- reference standards (origin, batch number, certificate of analysis, stability and storage conditions),
- calibration standards and QC samples (matrix , anticoagulant if applicable, preparation, preparation dates, and storage conditions),
- run acceptance criteria,
- analysis:
    - table of all analytical runs with analysis dates, whether passed or failed and the reason for the failure
    - table of calibration results of all accepted analytical runs, including calibration range, response function, back-calculated concentrations, and accuracy,
    - table of QC results of all accepted analytical runs (within- and between-run precision and accuracy); values outside acceptance criteria should be clearly marked,
    - stability data of stock solution, working solution, QC, covering the applied storage conditions,
    - data on selectivity, LLOQ, carry-over, matrix effect if applicable, dilution integrity;
- unexpected results obtained during validation with full justification of the action taken,
- deviations from method and/or SOPs (description of deviations, impact on study, supportive data).

All measurements with the individual calculated concentrations have to be presented in the validation report.

8.2. Analytical report

The analytical report should include a reference to the validation report(s) applicable to the analysis of the study samples. Furthermore it should include a detailed description of the analysis of the study samples.

If the analytical report provides detailed information, a reference to the analysis specific SOPs in the analytical report is sufficient. Otherwise the SOPs should be appended to the analytical report.

All source data should be available in its original format and available on request.

Any deviation from the protocol, analytical procedure or SOPs should also be discussed in the analytical report.

The analytical report should include at least the following information:

- reference standards (origin, batch, certificate of analysis, stability, storage conditions)
- calibration standards and QC samples (storage conditions)
- run acceptance criteria (short description, reference to specific protocol or SOP)
- assay procedure (short description)
- sample tracking (dates of receipt and contents, sample conditions on receipt, storage location and conditions, if applicable)
- study sample analysis:

  - content of the analytical run,
    - table identifying all analytical runs and study samples, with run dates and results,
    - table of calibration results of all (passed) analytical runs,
    - table of QC results of all (passed) analytical runs; values outside acceptance criteria should be clearly marked;
- failed analytical runs (identity, assay date, reason for failure),
- deviations from method and/or SOPs (description of deviations, impact on study, supportive data),
- reassay, excluding reassay due to analytical reasons, such as failed run (table of sample identification, reason for re-assay, original and re-assay values).

The results of incurred sample reanalysis may be supplied either in the validation report, in the analytical report or in a stand alone report.

For bioequivalence studies, all chromatograms from the runs which include 20% of the subjects, including the corresponding QC samples and calibration standards should be appended to the analytical study report. For other studies representative chromatograms should be appended to the report. Additional chromatograms should be available on request.

Definitions

1. Accuracy

The accuracy of an analytical procedure expresses the closeness of the determined value to the value which is accepted either as a conventional true value or an accepted reference value. Accuracy is defined as (determined value/true value) x100%.

2. Analyte

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

3. Analytical run

A complete set of analytical and study samples with appropriate number of calibration standards and QC samples for their validation. Several runs may be completed in one day, or one run may take several days to complete.

4. Analytical Procedure

The analytical procedure refers to the way of performing the analysis. It should describe in detail the steps necessary to perform each analysis.

5. Anchor calibrators

Anchor calibrators are standards points outside of the range of quantification, used to assist in fitting the non linear regression of the standard curve in ligand-binding assays.

6. Calibration range

The range of an analytical procedure is the interval between the upper and lower concentration (amounts) of analyte in the sample (including these concentrations) for which it has been demonstrated that the analytical procedure meets the requirements for precision, accuracy and response function.

7. Calibration standard

A matrix to which a known amount of analyte has been added or spiked. Calibration standards are used to construct calibration curves.

8. Carry over

Carry-over is the appearance of an analyte signal in blank sample after the analysis of samples with a high analyte concentration.

9. Cross validation

Comparison of validation parameters of two bioanalytical methods.

10. Full validation

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

11. Incurred samples

Study samples from dosed subjects or animals.

12. Incurred sample reanalysis

The analysis of a portion of the incurred samples to determine whether the original analytical results are reproducible.

13. Internal standard

Test compound(s) (e.g. a structurally similar analogue, or stable isotope labelled compound) added to calibration standards, QC samples and study samples at a known and constant concentration to correct for experimental variability during sample preparation and analysis.

14. Lower limit of quantification (LLOQ)

The lower limit of quantification of an individual analytical procedure is the lowest amount of analyte in a sample which can be quantitatively determined with pre-defined precision and accuracy.

15. Matrix effect

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

16. Nominal concentration

Theoretical or expected concentration.

17. Partial validation

Series of analytical experiments where only relevant parts of the validation are repeated after modifications are made to the validated bioanalytical method.

18. Precision

The precision of an analytical procedure expresses the closeness of agreement (degree of scatter) between a series of measurements obtained under the prescribed conditions. Precision is defined as the ratio of standard deviation/mean (%).

19. Quality control (QC) sample

A spiked sample used to monitor the performance of a bioanalytical method and to assess the integrity and validity of the results of the unknown samples analysed in an individual batch.

20. Response function

Response function is a function which adequately describes relationship between instrument response (e.g. peak area or height ratio) and the concentration (amount) of analyte in the sample. Response function is defined within a given range.

21. Selectivity

Selectivity is the ability of the bioanalytical method to measure and differentiate the analyte(s) of interest and internal standard in the presence of components which may be expected to be present in the sample.

22. Specificity

Specificity is the ability to measure the analyte unequivocally in the presence of other compounds, either exogenous or endogenous, in the matrix.

23. Stability

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

24. Standard Operating Procedure

Document which describes the regularly recurring operations relevant to the quality of the investigation and enabling to carry out the operations correctly and always in the same manner..

25. Upper limit of quantification (ULOQ)

The upper limit of quantification of an individual analytical procedure is the highest amount of analyte in a sample which can be quantitatively determined with pre-defined precision and accuracy.