写

事　務　連　絡
平成 26 年 5 月 30 日

各都道府県薬務主管課　御中

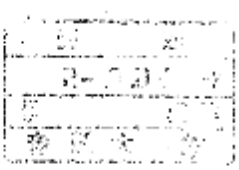
厚生労働省医薬食品局審査管理課

「医薬品開発における生体試料中薬物濃度分析法のバリデーションに関するガイドライン（リガンド結合法）」等の英文版の送付について

標記について、別添１及び２のとおり取りまとめましたので、貴管下関係業者に対して周知方お願いします。

別添１　Guideline on Bioanalytical Method（Ligand Binding Assay）Validation in Pharmaceutical Development

別添2　Questions and Answers (Q&A) for the Guideline on Bioanalytical Method (Ligand Binding Assay) Validation in Pharmaceutical Development



別添１

Guideline on Bioanalytical Method (Ligand Binding Assay) Validation in Pharmaceutical Development

Table of Contents

1. Introduction
2. Scope
3. Reference Standard
4. Analytical Method Validation
4.1. Full validation
4.1.1. Specificity
4.1.2. Selectivity
4.1.3. Calibration curve
4.1.4. Accuracy and precision
4.1.5. Dilutional linearity
4.1.6. Stability
4.2. Partial validation
4.3. Cross validation
5. Analysis of Study Samples
5.1. Calibration curve
5.2. QC samples
5.3. Incurred samples reanalysis (ISR)
6. Points to note
6.1. Calibration range
6.2. Reanalysis
6.3. Carry-over
6.4. Cross-talk
6.5. Critical reagents
6.6. Interfering substances
7. Documentation and Archives
List of relevant guidelines
Glossary

1. Introduction

In the development of medicinal products, bioanalytical methods are used in clinical and non-clinical pharmacokinetic studies (including toxicokinetic studies) to evaluate the efficacy and safety of drugs and their metabolites. Drug concentrations determined in biological samples are used for the assessment of characteristics such as *in vivo* pharmacokinetics (absorption, distribution, metabolism, and excretion), bioavailability, bioequivalence, and drug-drug interaction.

It is important that these bioanalytical methods are well characterized throughout the analytical procedures to establish their validity and reliability.

This guideline serves as a general recommendation for the validation of ligand binding assays (LBAs) as bioanalytical methods to ensure adequate reliability. It also provides a framework for analysis of study samples by using validated methods to generate study results supporting applications for drug marketing authorization.

Flexible adjustment and modification can be applied in the case of specialized analytical method or depending on the intended use of the result of analysis. Adjustments and modifications may include appropriate predefinition of acceptance criteria based on scientific rationale.

2. Scope

This guideline is applicable to the validation of LBAs as analytical methods for the measurement of drugs in biological samples obtained in toxicokinetic studies and clinical trials, as well as to the analysis of study samples using such methods. The information in this guideline generally applies to the quantification of peptides and proteins as well as low-molecular-weight drugs that are analyzed by LBAs. A typical example of an LBA is an immunological assay based on antigen-antibody reaction, such as enzyme immunoassay (EIA).

This guideline is not mandatory for analytical methods used in non-clinical studies that are beyond the scope of "Ministerial Ordinance Concerning the Standards for the Conduct of Non-clinical Studies on the Safety of Drugs (Ministry of Health and Welfare ordinance No. 21, dated March 26, 1997)," but could be used as a reference for conducting the required validation of such methods.

3. Reference Standard

A reference standard serves as the scale in quantifying an analyte, and is mainly used to prepare calibration standards and quality control (QC) samples, which are relevant blank matrix spiked with a known concentration of the analyte of interest. The quality of the reference standard is critical, as the quality affect measurement data. A certificate of analysis or an alternative statement that provides information on lot number, content (amount, purity, or potency) and storage conditions should accompany the standard. Also, the expiration date or its equivalent is preferably clarified. As for a reference standard, it is important that the material is procured from an authenticated source and is of well-controlled quality.

4. Analytical Method Validation

An analytical method validation should be performed at every relevant facility when establishing a bioanalytical method for quantification of a drug or its metabolite(s).

4.1. Full validation

A full validation should be performed when establishing a new bioanalytical method for quantification of an analyte/analytes. A full validation is also required when implementing an analytical method that is disclosed in literature or commercialized as a kit product.

The objective of a full validation is to demonstrate the assay performance of the method, e.g., specificity, selectivity, calibration curve, accuracy, precision, dilutional linearity, and stability. Generally, a full validation should be performed for each species or matrix (mainly plasma or serum) to be analyzed.

The matrix used in analytical validation should be as close as possible to the target study samples, including anticoagulants and additives. When an analytical method is to be established for a matrix of limited availability (rare matrix, e.g., tissue, cerebrospinal fluid, bile) and a sufficient amount of matrix cannot be obtained from sufficient number of sources (subjects or animals), a surrogate matrix may be used to prepare calibration standards and QC samples. However, the use of a surrogate matrix should be justified as much as possible in the course of establishing the analytical method.

In an LBA full validation, the minimum required dilution (MRD) should be defined *a priori* (i.e., in the course of method development) to dilute samples with buffer solution.

When using a plate-based LBA, analysis should generally be performed in at least 2 wells per sample; a sample concentration should then be determined either by calculating a mean of responses from the wells or by averaging the concentrations calculated from each response.

4.1.1. Specificity

Specificity is the ability of an analytical method to detect and differentiate the analyte from other substances, including its related substances (i.e., substances that are structurally similar to the analyte). For an LBA, it is important that the binding reagent specifically binds to the target analyte but does not cross-react with coexisting related substances. If presence of related substances is anticipated in biological samples of interest, the extent of the impact of such substances should be evaluated. Specificity may be evaluated in the course of method development. In some cases, an additional specificity testing may have to be conducted after a method validation is completed.

Specificity is evaluated using blank samples (matrix samples without analyte addition) and blank samples spiked with the related substance at concentration(s) anticipated in study samples; in addition, QC samples with the analyte concentrations near the lower limit of quantification (LLOQ) and near the upper limit of the quantification (ULOQ) of calibration curve should be evaluated after spiking with the related substance at anticipated concentration(s).

Assay results for the "near" blank sample and blank samples spiked with the related substance should be below the LLOQ; and accuracy in the measurements of the QC samples spiked with the related substance should demonstrate an accuracy of within ±20% of the theoretical concentration (or within ±25% of the theoretical concentration at the LLOQ and ULOQ).

4.1.2. Selectivity

Selectivity is the ability of an analytical method to detect and differentiate the analyte in the presence of other components in the samples.

Selectivity is evaluated using blank samples obtained from at least 10 individual sources and near-LLOQ QC samples (i.e., QC samples at or near the LLOQ) prepared using the individual blank samples. In the case of a matrix with limited availability, it may be acceptable to use matrix samples obtained from less than 10 sources.

Assay results for at least 80% of the blank samples should be below the LLOQ; at least 80% of the near-LLOQ QC samples should demonstrate an accuracy of within ±20% of the theoretical concentration (or within ±25% at the LLOQ).

4.1.3. Calibration curve

The calibration curve demonstrates the relationship between a theoretical analyte concentration and its resulting response variable.

A calibration curve should be prepared by using the same matrix as the intended study samples, whenever possible, by spiking the blank matrix with known concentrations of the analyte. A calibration curve should be generated with at least 6 concentration levels of calibration standards, including LLOQ and ULOQ samples, and a blank sample. Anchor point samples at concentrations below the LLOQ and above ULOQ of the calibration curve may also be used to improve curve fitting. A 4- or 5-parameter logistic model is generally used for the regression equation of a calibration curve. The validation report should include the regression equation and weighting conditions used.

The accuracy of back-calculated concentration of each calibration standard should be within ±25% deviation of the theoretical concentration at the LLOQ and ULOQ, and within ±20% deviation at all other levels. At least 75% of the calibration standards excluding anchor points, and a minimum of 6 levels of calibration standards, including the LLOQ and ULOQ, should meet the above criteria.

4.1.4. Accuracy and precision

Accuracy of an analytical method describes the degree of closeness between the analyte concentration determined by the method and its theoretical concentration. Precision of an analytical method describes variation between individual concentrations determined in repeated measurements.

Accuracy and precision are assessed by performing analysis with QC samples, i.e., samples spiked with known amounts of the analyte. In the validation, QC samples with a minimum of 5 different concentrations (LLOQ, low-, mid-, high-levels, and ULOQ) within the calibration range are prepared. The low-level should be within 3 times the LLOQ, the mid-level is near the midpoint on the calibration curve, and the high-level should be at least one-third of the ULOQ of the calibration curve. Accuracy and precision should be evaluated by repeating the analysis in at least 6 analytical runs.

The mean within-run and between-run accuracy at each concentration level should be within ±20% deviation of the theoretical concentration, except at the LLOQ and ULOQ, where it should be within ±25%. Within-run and between-run precision of concentrations determined at each level should not exceed 20%, except at the LLOQ and ULOQ, where it should not exceed 25%. Furthermore, a total error (sum of the absolute value of the relative error [i.e., accuracy minus 100%] and precision) at each level should not exceed 30%, except at the LLOQ and ULOQ, where it should not exceed 40%.

4.1.5. Dilutional linearity

Dilutional linearity is assessed to confirm the following: (i) the method can appropriately analyze samples at concentrations exceeding the ULOQ of a calibration curve without influence of a hook effect or prozone; (ii) measured concentrations are not affected by dilution within the calibration range. Dilutional linearity is evaluated by analyzing a QC sample exceeding the ULOQ of a calibration curve and its serial dilutions at multiple concentrations. The absence or presence of response reduction (hook effect or prozone) is checked in the analyzed samples, and if discovered, measures should be taken to eliminate response reduction in study sample analysis. Accuracy and precision in the measurements corrected for the dilution factor should be within ±20% deviation of the theoretical concentration and not more than 20%, respectively.

4.1.6. Stability

Analyte stability should be evaluated to ensure that the concentration is not affected through each step of the process from the sample collection to the analysis. The stability of the anlayte should be assessed under conditions that are as close as possible to the actual circumstances, e.g. sample storage and sample analysis. Careful consideration should be given to the solvent or matrix type, container materials, and storage conditions used in the stability-determination process.

Validation studies should determine analyte stability after freeze and thaw cycles, and after short-term (e.g., at room temperature, on ice, or under refrigeration) and long-term storage. All stability experiments should be performed on samples that have been stored for a time that is longer than the actual storage period.

Stability of the analyte in the stock and working solutions is evaluated using solutions at or near the highest and lowest concentration levels for the actual solution storage

situation. Stability of the analyte in the studied matrix is evaluated using low- and high-level QC samples. The QC samples should be prepared using a matrix that is as close as possible to the actual study samples, including anticoagulant and additives. Stability is evaluated by analysis of at least 3 replicates per QC concentration level before and after stability storage. The mean accuracy in the measurements at each level should be within ±20% deviation of the theoretical concentration, in principle. Other criteria could be used if they are deemed scientifically more appropriate for the evaluation of a specific analyte.

4.2. Partial validation

Partial validation may be performed when minor changes are made to an analytical method that has already been fully validated. The items in a partial validation are determined according to the extent and nature of the changes made to the method.

Typical bioanalytical method changes subjected to a partial validation are as follows: analytical method transfers between laboratories, changes in analytical instruments, changes of the critical reagent lot, changes in calibration range, changes in the MRD, changes in anticoagulant, changes in analytical conditions, changes in sample storage conditions, confirmation of impact by concomitant drugs, and use of rare matrices.

Acceptance criteria used in partial validation should, in principle, be the same as those employed in the full validation.

4.3. Cross validation

Cross validation is primarily conducted when data are generated in multiple laboratories within a study or when comparing analytical methods used in different studies, after a full or partial validation. The same set of QC samples spiked with the analyte or the same set of study samples is analyzed, and the mean accuracy at each concentration level of QC samples or the assay variability in the measurements of study samples is evaluated.

In the cross validation among 2 or more laboratories within a study, the mean accuracy of QC samples (low-, mid-, and high-levels) evaluated by at least 3 replicates at each level, should be within ±30% deviation of the theoretical concentration, in principle, considering intra- and inter-laboratory precision. When using a set of study samples, the assay variability should be within ±30% for at least two-thirds of the samples.

When conducting cross validation between different analytical methods based on different assay principles, both validation procedure and acceptance criteria (i.e., mean accuracy or assay variability) should be separately defined based on scientific judgment according to the type of the analytical methods.

5. Analysis of Study Samples

Study samples are biological specimens that are obtained from toxicokinetic studies and clinical trials for analysis. Analysis of study samples should be carried out using a fully validated analytical method. During analysis, study samples should be handled under conditions that are validated for adequate stability, and analyzed within a confirmed stability period, along with a blank sample, calibration standards at a minimum of 6 concentration levels, and QC samples at a minimum of 3 concentration levels. In a plate-based LBA, assay should generally be performed in at least 2 wells per sample prepared. A sample concentration should then be determined either by calculating a concentration from an average of each response or by averaging the concentrations calculated from each response.

Validity of the analytical method during study sample analysis should be evaluated in each analytical run by using the calibration curve and QC samples. In a plate-based assay, each plate represents a single analytical run. In studies that serve pharmacokinetic data as a primary endpoint, reproducibility of the analytical method should be confirmed for each representative study per matrix by performing incurred sample reanalysis (ISR: reanalysis of incurred samples in a separate analytical run on a different day to determine whether the original analytical results are reproducible).

5.1. Calibration curve

A calibration curve is used to determine the concentration of the analyte of interest in study samples. A calibration curve used in study sample analysis should be generated for each analytical run by using the validated analytical method. The same model as in the bioanalytical method validation should be used for the regression equation and weighting conditions of the calibration curve.

The accuracy of back-calculated concentrations of calibration standards at each level should be within ±25% deviation of the theoretical concentration at the LLOQ and ULOQ of the calibration curve, and ±20% deviation at all other levels. At least 75% of

the calibration standards excluding anchor points, with a minimum of 6 levels, should meet the above criteria.

If the calibration standard at the LLOQ or ULOQ does not meet the criteria in study sample analysis, the next lowest/highest-level calibration standard may be used as the LLOQ or ULOQ of the calibration curve. Even though narrowed, the modified calibration range should still cover at least 3 different QC sample levels (low-, mid-, and high-levels).

5.2. QC samples

QC samples are analyzed to assess the validity of the analytical method used for calibration curve and study sample analysis.

QC samples with a minimum of 3 different concentration levels (low-, mid-, and high-levels) within the calibration range are analyzed in each analytical run. Usually, the low-level is within 3 times the LLOQ, the mid-level is in the midrange of the calibration curve, and the high-level needs to be at least one-third of the ULOQ of the calibration curve. QC samples are processed in the same manner as with study samples. The analysis requires 2 QC samples at each QC level or at least 5% of the total number of study samples in the analytical run, whichever is the greater.

The accuracy of measurement of QC samples should be within ±20% deviation of the theoretical concentrations. At least two-thirds of the QC samples and at least 50% at each concentration level should meet the above criterion.

5.3. ISR (Incurred sample reanalysis)

In bioanalysis, it can happen that the results of analyses of study samples are not reproducible, even when the method validation is successfully conducted and the validity of at each analytical run is confirmed by calibration standards and QC samples. Such failures can be attributed to various factors, including inhomogeneity of study samples, contamination and other operational errors, and interference of biological components unique to the study samples or of unknown metabolites. ISR refers to reanalysis of incurred samples in separate analytical runs on different days to check whether the original analytical results are reproducible. Confirmation of the reproducibility by ISR improves the reliability of the analytical data. In addition, a failure to demonstrate the reproducibility of the original data in ISR can trigger a cause investigation and remedial

measures for the analytical method.

Usually, ISR is performed for representative studies selected for each matrix in studies that use pharmacokinetic data as the primary endpoint. For instance, ISR should be conducted in the following situations: non-clinical toxicokinetic studies for each different species, representative clinical pharmacokinetic studies in healthy volunteers and patients with renal/hepatic impairment, as well as bioequivalence studies. For non-clinical studies, ISR may be performed with samples obtained in a independent non-GLP study, if the study design is similar to the relevant toxicokinetic study in terms of sampling conditions.

ISR should be performed with samples from as many subjects or animals as possible, including those near the maximum blood concentration (C_{max}) and the elimination phase, within a time window that ensures the analyte stability. As a guide, approximately 10% of the samples should be reanalyzed in cases where the total number of study samples is less than 1000 and approximately 5% of the number of samples exceeding 1000.

The results of ISR are evaluated using assay variability. Assay variability can be calculated as the difference between the concentration obtained by ISR and that in the original analysis divided by their mean and multiplied by 100. The assay variability should be within +30% for at least two-thirds of the samples analyzed in ISR. In case the ISR data failed to meet the above criteria, root cause investigation should be conducted for the analytical method and necessary measures should be taken by considering the potential impact on study sample analysis.

It should be noted that ISR is performed to monitor assay variability. The original data should never be discarded or replaced with the reanalysis data even if the assay variability exceeds ±30% in a specific sample.

6. Points to note

6.1. Calibration range

In LBAs, calibration range is largely dependent on the characteristics of the binding reagent and it may be difficult to arbitrarily determine the range. In addition, because the calibration range of LBA is comparably narrow, dilutional linearity should be appropriately established to bring the concentrations of analyte in diluted study samples within the range of the calibration curve.

In case the calibration range is changed, partial validation should be performed.

However, it is not necessary to reanalyze the study samples that have been quantified prior to the change in the calibration range.

6.2. Reanalysis

Possible reasons and procedures for reanalysis, as well as criteria for handling of concentration data should be defined *a priori* in the protocol or standard operating procedure (SOP).

Examples of reasons for reanalysis are as follows: calibration curve or QC samples failed to meet the criteria for validity of the analytical run; the obtained concentration exceeded the ULOQ of the calibration curve or fell below the LLOQ due to excess dilution; the analyte of interest was detected in pre-dose or placebo samples; improper analytical operation or malfunction of analytical instrument; and causal investigation on abnormal values.

Reanalysis of study samples for pharmacokinetic reasons should be avoided, whenever possible. Particularly in bioequivalence studies, it is not acceptable to reanalyze study samples and modify the concentration data only because the initial data were pharmacokinetically questionable. However, reanalysis of specific study samples is acceptable when, for instance, the initial analysis yielded an unexpected or anomalous result that may affect the safety of subject in a clinical trial.

In any case, when reanalysis is performed, the analytical report should provide information of the reanalyzed samples, the reason for reanalysis, the data obtained in the initial analysis, if any, the data obtained in the reanalysis, and the final accepted values and the reason and method of selection.

6.3. Carry-over

Carry-over is an alteration of a measured concentration due to residual analyte in the analytical equipment.

Carry-over is not an issue for analyses performed in plates and tubes, while it should be taken into account in analyses that use a single flow cell, flow path, and/or autosampler.

If carry-over is inevitable, its impact needs to be examined, and appropriate measures should be taken to avoid any impact on the actual study sample analysis. Should there be any concern that carry-over may affect the quantification of analyte in study samples, it

should be evaluated during the actual study sample analysis to assess the impact on the concentration data.

6.4. Cross-talk

Cross-talk is an alteration of a measured concentration due to a leak of fluorescence or luminescence to adjacent wells in plate-based assay.

If cross-talk is inevitable, its impact needs to be examined, and appropriate measures should be taken to avoid any impact on the actual study sample analysis. Should there be any concern that cross-talk may affect the quantification of analyte in study samples, this should be evaluated during the actual study sample analysis to assess the impact on the concentration data.

6.5. Critical reagents

A critical reagent is the one that has a direct impact on the results of an LBA-based bioanalytical method and usually includes, but is not limited to, binding reagents (e.g., unlabeled or labeled antibodies).

A critical reagent should be selected by considering the specificity for the analyte and should be stored under conditions that ensure consistent quality. The quality of critical reagent should be appropriately maintained throughout the period of use in analytical method validation and study sample analysis. Partial validation is in principle required when the critical reagent lot is changed.

6.6. Interfering substances

Interfering substances are those that may affect the concentration data in study sample analysis and may include, but are not limited to, soluble ligands and anti-drug antibodies.

If interfering substances are potentially present in study samples, it is advisable to examine the impact of interfering substances on the concentration data.

7. Documentation and Archives

In order to ensure adequate reproducibility and reliability of bioanalysis, results obtained in analytical method validations and study sample analyses should be

documented in a validation report and a study sample analysis report as described below. The reports should be stored along with relevant records and raw data in an appropriate manner.

All relevant records and raw data should be kept, including those obtained in rejected analytical runs, specifically record of reference materials, blank matrices, and critical reagents (receipt/release, use, and storage), record of samples (receipt/release, preparation, and storage), record of analyses, record of instrument (calibration and settings), record of deviations, record of communications, and raw data such as analytical data.

Validation report

- Summary of the validation
- Information on the reference standards
- Information on the blank matrices
- Information on the critical reagents
- Analytical method (including description related to the MRD)
- Validated parameters and the acceptance criteria
- Validation results and discussion
- Rejected runs together with the reason for rejection
- Information on reanalysis
- Deviations from the protocol and/or SOP, along with the impact on study results
- Information on reference study, protocol, and literature

Study sample analysis report

- Summary of the study sample analysis
- Information on the reference standards
- Information on the blank matrices
- Information on receipt and storage of study samples
- Information on the critical reagents
- Analytical method (including description related to the MRD)
- Parameters, acceptance criteria, and results of the validity evaluation
- Results and discussion of study sample analysis

- Rejected runs together with the reason for rejection
- Information on reanalysis
- Deviations from the protocol and/or SOP, along with impact on study results
- Information on reference study, protocol, and literature

List of relevant guidelines

1) Regarding the "Guideline on Bioanalytical Method Validation in Pharmaceutical Development" PFSB/ELD Notification No. 0711-1 dated July 11, 2013.
2) Regarding the "Questions and Answers (Q&A) for the Guideline on Bioanalytical Method Validation in Pharmaceutical Development" Office Communication dated July 11, 2013.
3) Regarding "the Guidance on Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals (ICH M3(R2))" PFSB/ELD Notification No. 0219-4 dated February 19, 2010.
4) Regarding the "Note for Guidance on Toxicokinetics: The Assessment of Systemic Exposure in Toxicity Studies" PAB/ELD Notification No. 443 dated July 2, 1996.
5) Regarding the "Guideline on Nonclinical Pharmacokinetics" PNSB/ELD Notification No. 496 dated June 26, 1998.
6) Regarding "Partial Revision of the Guideline on Bioequivalence Studies for Generic Pharmaceuticals, etc." PFSB/ELD Notification No. 0229-10 dated February 29, 2012.
7) Regarding Revision of the "Questions and Answers (Q &A) Concerning the Guideline on Bioequivalence Studies for Generic Pharmaceuticals" Office Communication dated February 29, 2012.
8) "Note on Clinical Pharmacokinetic Studies of Pharmaceuticals" etc. PFSB/ELD Notification No. 796 dated June 1, 2001.
9) US FDA: Guidance for Industry, Bioanalytical Method Validation, U.S. Department of Health and Human Services, FDA, Center for Drug Evaluation and Research, Center for Veterinary Medicine (2001).
10) EMA: Guideline on bioanalytical method validation, EMEA/CHMP/EWP/192217/2009, Committee for Medicinal Products for Human Use (2011).

Glossary

Accuracy: The degree of closeness of a concentration determined by the method to the theoretical concentration. Accuracy is expressed as a percentage relative to the theoretical concentration.

Accuracy (%) = (Measured concentration/Theoretical concentration) × 100.

Analysis: A series of analytical procedures from sample dilution to measurement on an analytical instrument.

Analyte: A specific compound being analyzed. It can be a drug, biomolecule or its derivative, metabolite, and/or degradation product in a sample.

Analytical run: A set of samples comprising calibration standards, QC samples, and study samples. Usually, a set of samples is prepared without interruption in time by a single analyst with the same reagents under the same conditions, and subsequently analyzed in a single plate as a single analytical run.

Anchor point: Samples set at concentrations below the LLOQ or above the ULOQ of the calibration curve and analyzed concurrently with calibration standards to improve curve fitting.

Assay variability: The degree of difference between the duplicate concentrations determined for a single sample. The difference is expressed as a percentage relative to the mean of the two.

Assay variability (%) = [(Concentration in comparative analysis - Concentration in reference analysis) / Mean of the two] × 100.

Binding reagent: A reagent that directly binds to the analyte in LBA-based bioanalytical methods.

Blank sample: A matrix sample that is analyzed without analyte addition.

Calibration curve: Presentation of the relationship between concentration and response for an analyte. A calibration curve is generated with at least 6 concentration levels of calibration standards, including LLOQ and ULOQ samples, as well as a blank sample. Anchor points may be added outside the calibration range.

Calibration standard: A sample spiked with the analyte of interest to a known concentration, which is used to generate calibration curves. Calibration standards are used to generate calibration curves, from which the concentrations of the analyte in QC

samples and study samples are determined.

Carry-over: An alteration of the measured concentration due to a residual analyte in the analytical instrument used.

Critical reagent: A reagent that has a direct impact on the results of an LBA-based bioanalytical method, an example being binding reagents (e.g., unlabeled or labeled antibodies).

Cross validation: A validation conducted when data are generated in multiple laboratories within a study or when comparing analytical methods used in different studies. Cross validation is performed after a full or partial validation.

Cross-reactivity: Interaction of the binding reagent with components other than the analyte.

Cross-talk: An alteration of a measured concentration due to leak of fluorescence or luminescence to adjacent wells in plate-based assay.

Dilutional linearity: A parameter demonstrating that the method can appropriately analyze samples at a concentration exceeding the ULOQ of the calibration curve without influence of hook effect or prozone effect and that these measured concentrations are not affected by dilution within the calibration range.

Full validation: A validation that evaluates a full panel of performance parameters, i.e., specificity, selectivity, calibration curve, accuracy, precision, dilutional linearity, and stability. A full validation is usually performed when establishing a new bioanalytical method.

Hook effect: Suppression of response due to very high concentrations of a particular analyte. Precautions should be taken when the hook effect is expected, because samples at a concentration exceeding the ULOQ of calibration curve may falsely produce results below the upper limit of the calibration range. A hook effect may occur in LBAs that utilize a liquid-phase reaction step for incubating the binding reagents with the analyte.

Incurred sample: A study sample that is obtained from a subject or animal dosed with an active study drug.

Incurred sample reanalysis (ISR): Reanalysis of a portion of the incurred samples in separate analytical runs on different days to determine whether the original analytical results are reproducible.

Interfering substance: A substance that is present in the matrix that may affect interaction between the binding reagent and the analyte.

Ligand binding assay: A method to analyze an analyte of interest using reagents that specifically bind to the analyte. The antigen-antibody reaction is utilized in the majority of ligand binding assays. The analyte is detected using reagents labeled with e.g., an enzyme, radioisotope, fluorophore, or chromophore. Reactions are carried out in 96-well microtiter plates, test tubes, disks, etc.

Lower limit of quantification (LLOQ): The lowest concentration of analyte in a sample at which the analyte can be quantified with reliable accuracy and precision.

Matrix: Whole blood, plasma, serum, urine, or other biological fluid or tissue selected for analysis. A matrix not containing exogenous chemicals (except anticoagulant) and their metabolites is called a blank matrix.

Minimum required dilution (MRD): A dilution factor where biological samples are diluted with buffer solution for the analysis by LBAs. The MRD may not necessarily be the ultimate minimum dilution but should be identical for all samples including calibration standards and QC samples.

Partial validation: A validation performed when minor changes are made to an analytical method that has already been fully validated. A set of parameters to be evaluated in a partial validation should be determined according to the extent and nature of the changes made to the method. It can range from as little as accuracy and precision evaluation in a day to a nearly full validation.

Precision: The degree of closeness between individual concentrations determined in repeated measurements. Precision is expressed as the coefficient of variation (CV) or the relative standard deviation (RSD) in percentage.

$$\text{Precision (\%)} = (\text{Standard deviation}/\text{Mean}) \times 100.$$

Prozone: Suppression of response due to very high concentrations of a particular analyte. This is the same phenomenon as the hook effect.

Quality control (QC) sample: A sample spiked with the analyte of interest to a known concentration used to assess the reliability of an analytical method. In analytical runs, QC samples are analyzed to assess the validity of the analytical method used for calibration curve and study sample analysis.

Quantification range: The range of concentration of analyte in which the analyte can

be quantified with reliable accuracy and precision. The quantification range of a bioanalytical method is ensured by the range of calibration curve (calibration range) and the dilutional linearity.

Reanalysis: Repetition of a series of analytical procedures from the dilution step to final analysis on samples that have been analyzed once.

Reference standard: A compound used as the standard in quantifying an analyte mainly used to prepare calibration standards and QC samples.

Response variable: A response obtained from the detector of the analytical instrument. In LBAs, response is generally monitored by a spectroscopic technique, which converts the response into an electrical signal, such as absorbance or luminescent intensity.

Selectivity: The ability of an analytical method to detect and differentiate the analyte in the presence of other components in biological samples.

Specificity: The ability of an analytical method to detect and differentiate the analyte from its related substances. In LBAs, it is largely dependent on the properties of binding reagents.

Stability: The chemical or biological stability of an analyte in a given solvent or matrix under specific conditions over given time intervals. Analyte stability is evaluated at each step to ensure that the analyte concentration is not affected by the samples move through each step of the process from sample collection to final analysis.

Stock solution: A non-matrix solution of reference material at the highest concentration prepared in an appropriate solvent.

Study sample: A biological specimen that is obtained from a toxicokinetic study or clinical trial for bioanalysis.

Surrogate matrix: A matrix used as an alternative to a matrix of limited availability (e.g., tissue, cerebrospinal fluid, bile). Surrogate matrix may also be used as an alternative to a matrix that contains endogenous substances that have the same structure to the analyte.

Total error: Sum of the absolute value of the relative error (i.e., accuracy minus 100%) and precision.

Upper limit of quantification (ULOQ): The highest concentration of analyte in a sample at which the analyte can be quantified with reliable accuracy and precision.

Validation: Demonstration of adequate reproducibility and reliability of an analytical method through various evaluations.

Working solution: A non-matrix solution prepared by diluting the stock solution in an appropriate solvent. It is mainly added to matrix to prepare calibration standards and QC samples.

別添2

Questions and Answers (Q&A) for the Guideline on Bioanalytical Method (Ligand Binding Assay) Validation in Pharmaceutical Development

<<Scope>>

Q1. Is this guideline applicable to a drug having an amino acid sequence identical to an endogenous substance?

A1. Yes, this guideline is applicable. However, selection of a blank matrix requires special precautions, such as the use of a surrogate matrix or a matrix that has been depleted of the endogenous substance concerned (see Q&A No. 9 and No. 13). Same precautions should be taken for a drug which is indistinguishable by a ligand binding assay (LBA) from the endogenous substance.

<< Reference Standard>>

Q2. What procedures should be followed if the expiration date for the reference standard is yet to be established?

A2. If the expiration date of the reference standard cannot be established, quality of the reference standard should be ensured by other appropriate means, such as setting a re-test date *a priori*.

Q3. What procedures should be followed in renewing the reference standard lot?

A3. Confirm comparability of the current and new reference standard lots by referring to the relevant Certificate of Analysis (CoA), or any appropriate documentation. If no information regarding the lot comparability is available from the CoA or other relevant documentation, it should be confirmed by an LBA.

Q4. Does the reference standard lot have to be the same as the drug substance lot used for dosing in the non-clinical or clinical studies?

A4. Any lot may be used as the reference standard as long as it conforms to the same quality specifications based on information available from a CoA or other appropriate document. In an early stage of non-clinical studies where a quality specifications for a standard material are yet to be established, it is preferable that the reference standard lot is the same as the drug substance lot used for dosing in the non-clinical studies; if this is not the case, the lot comparability has to be confirmed by an LBA.

<<Full validation>>

Q5. What is the difference between the MRD and sample dilution stated in the dilutional linearity section?

A5. The MRD is a dilution factor where samples are diluted with buffer solution to reduce the matrix effect on the analysis: an identical MRD should be applied to all samples, including calibration standards and QC samples. On the other hand, sample dilution is a procedure where samples with high analyte concentrations are diluted with blank matrix or diluted blank matrix to bring the analyte concentration within the calibration range.

Q6. When a sample is analyzed in two or more wells and when there is a large variation in their responses or determined concentrations, what procedures should be followed?

A6. Specify handling procedures for data with a large variation *a priori* in the protocol or standard operating procedure (SOP). Such a variation will compromise the data's reliability.

<<Specificity>>

Q7. Is there any case where specificity evaluation is not necessary?

A7. Specificity is important information for analytical data interpretation. In an LBA, specificity is dependent on the reactivity of the binding reagent. Therefore, if the characteristics of the binding reagent are well known from its development phase, it may not be necessary to repeat the specificity test in validation.

<<Selectivity>>

Q8. Should the use of disease-derived, hemolyzed, or lipemic matrix be necessary?

A8. If such a factor is likely to affect the assay system, consider using a relevant matrix, although this is not mandatory.

<<Calibration curve>>

Q9. What should be used as a blank matrix in analysis of a drug having an amino acid sequence identical to an endogenous substance?

A9. If the presence of the endogenous substance is anticipated in the study sample matrix, a surrogate matrix or a matrix that has been depleted of the endogenous substance can be used. When such matrices are used, their validity should be shown.

<<Accuracy and precision>>

Q10. For QC samples, it is stated that "the mid-level is in the midrange on the calibration curve, and the high-level needs to be at least one-third of the upper limit of quantification (ULOQ) of the calibration curve." Please specify how to set these concentration levels.

A10. The mid-level, in the midrange on the calibration curve/calibration range, generally means a level near the geometric mean of the ULOQ and the lower limit of quantification (LLOQ); a level adjacent to the arithmetic mean may be used for a more balanced QC sample distribution. The high-level, at least one-third of the ULOQ of the calibration curve, is intended to equally distribute QC samples within the range of a calibration curve where the nominal analyte concentrations are plotted on a logarithmic scale. A level around 75% of the ULOQ may be used depending on the overall balance of the QC sample distribution.

Q11. How should the number of replicates be set in each analytical run when assessing the accuracy and precision?

A11. When calculating within-run accuracy and precision, at least triplicate analyses (n=3) per analytical run are required. When calculating between-run accuracy and precision, a single analysis per analytical run would be accepted. An alternative method, such as analysis of variance (ANOVA), is also accepted. When using ANOVA, at least duplicate analyses are set per analytical run to evaluate accuracy and precision.

Q12. Why is a total error required?

A12. The absolute value of the relative error (i.e., accuracy minus 100%) represents the systematic error and precision reflects the random error. Evaluation of the total error allows early elimination of an analytical method having a large variation that could compromise data reliability (DeSilva et al, Pharm. Res., 2003). Although the acceptance criteria in accuracy and precision for LBA are set wider than those for chromatography/small molecules, it is anticipated that, by evaluating the total error, a LBA-based analytical method whose accuracy and precision are both on the edge of acceptability can be eliminated; this should ensure the reliability of the analytical results obtained.

Q13. What are the points to consider in assessing accuracy for the analysis of a drug having an amino acid sequence identical to an endogenous substance?

A13. Use a surrogate matrix or a matrix that has been depleted of the endogenous substance. Alternatively, determine the concentration of the endogenous substance in the blank sample and calculate the accuracy using one of the following formulas.

$$\text{Accuracy (\%)} = \frac{\text{Measured concentration of drug in sample}}{\text{Conc. of endogenous substance} + \text{Conc. of reference standard}} \times 100$$

$$\text{Accuracy (\%)} = \frac{\text{Measured conc. of drug in sample} - \text{Conc. of endogenous substance}}{\text{Concentration of reference standard}} \times 100$$

<<Dilutional linearity>>

Q14. What is the difference between dilutional linearity and dilution integrity?

A14. Dilution integrity is tested to confirm that the dilution procedure has no impact on the measured concentration, while dilutional linearity is tested to confirm not only dilution integrity, but also the absence or presence of a hook effect or prozone.

<<Cross validation>>

Q15. What is the rationale for setting the acceptance criteria as "the mean accuracy at each level should be within ±30% deviation of the theoretical concentration, in principle"?

A15. The guideline requires the mean accuracy to be within ±20% deviation of the theoretical concentration for analytical method validation. The acceptance criterion was relaxed to 30% for cross validation to accommodate additional factors, i.e., intra- and inter-laboratory precision. When study samples within a single study are to be analyzed at multiple laboratories, measures should be taken to minimize inter-laboratory variation in study sample analyses, such as by defining criteria for handling of study samples and reference standards in the relevant protocol or SOP, separate from the analytical method validation.

<< Study sample analysis >>

Q16. Should a calibration curve analysis be required for each plate?

A16. In principle, each plate should contain a set of calibration standards. If, however, assay integrity is demonstrated during the validation process, it is also acceptable to apply a single calibration standard set to multiple plates as far as each plate contains a set of QC samples.

Q17. Is it not necessary to evaluate parallelism?

A17. Parallelism is defined as an established parallel relationship between a dose-response curve from a study sample dilution series and a curve from a calibration standard series, with no difference among back-calculated

concentrations for multiple dilutions of a study sample. As of the issuance of this guideline, domestic and international knowledge has neither accumulated nor discussion yet matured regarding cases in which parallelism was not established, causes for failing to establish parallelism, and the extent of impact the failure might have on pharmaceutical development. Therefore, evaluation of parallelism is not necessarily required for all analytical methods. However, if parallelism is an intrinsic issue for an LBA-based bioanalytical method and is likely to cause a problem based on the nature of the analyte or method or data accumulated in the course of pharmaceutical development, scientifically valid evaluation and assessment of the impact on measured concentrations should be considered to the extent possible.

<<QC samples in study sample analysis>>

Q18. What should be considered in the arrangement of calibration standards and QC samples within a plate?

A18. It may sometimes occur that a certain pattern is inevitably seen in the assay results due to sample arrangement within a plate (e.g., edge effect); in such cases measures should be taken to mitigate the impact on the analysis results. Measures include the arrangement of calibration standards, QC samples, and study samples on a plate and the number of replicates for each sample prepared.

<<ISR>>

Q19. How should I perform ISR in toxicokinetic studies?

A19. In a GLP toxicokinetic study, ISR should be performed once per matrix for each animal species. If an analytical method is modified or analysis is performed in a different laboratory, ISR should be performed again.

In addition, ISR can be performed during a bioanalytical method validation using study samples obtained from a non-GLP study such as a dose-finding study performed before a GLP toxicokinetic study. In this case, the study design, including dose and regimen, should be comparable to that of the GLP study.

Q20. How should I perform ISR in clinical trials?

A20. ISR should be performed in representative clinical trials whose pharmacokinetic data as a primary endpoint. To evaluate the validity of an analytical method in an early stage, ISR should be performed as early as possible in the process of pharmaceutical development.

In a clinical trial with a different population of subjects with altered matrix composition, ISR should be performed again. In a bioequivalence study which serves pharmacokinetic data as the primary endpoint, ISR should be performed in the study.

Q21. If study samples obtained from clinical trials are already available at the time of analytical method validation, can I use the samples for ISR?

A21. If you have already obtained study samples from a clinical trial at the time of analytical method validation, you can use the samples for ISR. For example, a metabolite is added to the analyte(s), or reanalysis is performed with an improved analytical method after a failure to meet ISR acceptance criteria. However, an informed consent must be obtained from each subject who provides the study samples. The procedures of ISR and related items should be predefined.

Q22. If overall results meet the ISR acceptance criteria, but the assay variability of a specific sample exceeds the threshold of ±30%, is it required to reanalyze the samples to correct first value?

A22. ISR is intended to confirm the validity of an analytical method using study samples. Therefore, reanalysis of individual study samples is not required to correct the first value, even though the assay variability exceeds the threshold of ±30% when overall result meets the ISR acceptance criteria.

Q23. Where in a report is appropriate to provide ISR results?

A23. When the ISR is performed in the study sample analysis, ISR results should be reported in a study sample analytical report to prove the validity of an analytical method. When the ISR is performed in the analytical method validation, ISR results should be reported in a validation report.

Q24. In non-clinical studies, it is often the case that the sample volume is not sufficient for ISR evaluation. What procedures should be followed in that case?

A24. The study plan should always assume ISR, even in non-clinical studies. Even if the volume becomes insufficient in some samples due to reanalysis or for other reasons, ISR can still be performed, for example, by using samples from other time points. ISR may also be performed by utilizing samples from a preliminary study in which samples were collected under comparable conditions. In any case, even in non-clinical studies, every practical effort should be made to evaluate the reproducibility of the measured concentrations by ISR.

<< Critical reagents >>

Q25. Should an expiration date be established for critical reagents?

A25. Expiration date is not necessarily required for a critical reagent, as long as the quality is ensured by evaluation of data from calibration standards and QC samples during the period the reagent is used.