ANALYTICAL METHOD VALIDATION
AND INSTRUMENT PERFORMANCE
VERIFICATION

Edited by CHUNG CHOW CHAN, HERMAN LAM, Y. C. LEE, XUE-MING ZHANG
JOHN WILEY & SONS, INC.

CHUNG CHOW CHAN
Eli Lilly Canada, Inc.

HERMAN LAM
GlaxoSmithKline Canada, Inc.

Y. C. LEE
Patheon YM, Inc.

XUE-MING ZHANG
Novex Pharma
three concentration levels covering the specified range (e.g., three concentrations/three replicates).

For a drug substance, the common method of determining accuracy is to apply the analytical procedure to the drug substance and to quantify it against a reference standard of known purity. For the drug product, accuracy is usually determined by application of the analytical procedure to synthetic mixtures of the drug product components or placebo dosage form to which known quantities of drug substance of known purity have been added. The range for the accuracy limit should be within the linear range. Typical accuracy of the recovery of the drug substance in the mixture is expected to be about 98 to 102%. Values of accuracy of the recovery data beyond this range need to be investigated.

2.4.3 Precision

The precision of an analytical procedure expresses the closeness of agreement (degree of scatter) between a series of measurements obtained from multiple samples of the same homogeneous sample under prescribed conditions. Precision is usually investigated at three levels: repeatability, intermediate precision, and reproducibility. For simple formulation it is important that precision be determined using authentic homogeneous samples. A justification will be required if a homogeneous sample is not possible and artificially prepared samples or sample solutions are used.

Repeatability (Precision). Repeatability is a measure of the precision under the same operating conditions over a short interval of time. It is sometimes referred to as intraassay precision. Two assaying options are allowed by the ICH for investigating repeatability:

1. A minimum of nine determinations covering the specified range for the procedure (e.g., three concentrations/three replicates as in the accuracy experiment), or
2. A minimum of six determinations at 100% of the test concentration.

The standard deviation, relative standard deviation (coefficient of variation), and confidence interval should be reported as required by the ICH.

Tables 2.2 and 2.3 are examples of repeatability data. Table 2.2 shows good repeatability data. However, note that the data show a slight bias below 100% (all data between 97.5 and 99.1%). This may not be an issue, as the true value of the samples and the variation of the assay may be between 97.5 and 99.1%. Table 2.3 shows two sets of data for a formulation at two dose strengths that were performed using sets of six determinations at 100% test concentration. The data indicate a definite bias and high variability for the low-strength dose formulation. It may call into question the appropriateness of the low-dose samples for the validation experiment.
and method transfer between different laboratories). To validate this characteristic, similar studies need to be performed at other laboratories using the same homogeneous sample lot and the same experimental design. In the case of method transfer between two laboratories, different approaches may be taken to achieve the successful transfer of the procedure. However, the most common approach is the direct method transfer from the originating laboratory to the receiving laboratory. The originating laboratory is defined as the laboratory that has developed and validated the analytical method or a laboratory that has previously been certified to perform the procedure and will participate in the method transfer studies. The receiving laboratory is defined as the laboratory to which the analytical procedure will be transferred and that will participate in the method transfer studies. In direct method transfer it is recommended that a protocol be initiated with details of the experiments to be performed and acceptance criteria (in terms of the difference between the means of the two laboratories) for passing the method transfer. Table 2.4 gives a set of sample data where the average results obtained between two laboratories were within 0.5%.

### 2.4.4 Robustness

The robustness of an analytical procedure is a measure of its capacity to remain unaffected by small but deliberate variations in the analytical procedure parameters. The robustness of the analytical procedure provides an indication of its reliability during normal use. The evaluation of robustness should be considered during development of the analytical procedure. If measurements are susceptible to variations in analytical conditions, the analytical conditions should be suitably controlled or a precautionary statement should be included in the procedure. For example, if the resolution of a critical pair of peaks was very sensitive to the percentage of organic composition in the mobile phase, that observation would have been observed during method development and should be stressed in the procedure. Common variations that are investigated for robustness include filter effect, stability of analytical solutions, extraction time during sample preparation, pH variations in the mobile-phase composition, variations in mobile-phase composition, columns, temperature effect, and flow rate.

Table 2.5 shows examples of sample and standard stability performed on an analytical procedure. The two sets of data indicate that the sample and standard

| Table 2.4. Results from Method Transfer between Two Laboratories |
|------------------|------------------|
| **Runs** | **Average %** |
| Originating laboratory | 12 | 100.7 |
| Receiving laboratory | 4 | 100.2 |
at expiry), one can determine the repeatability by performing three replicate preparations for each sample. ICH guidelines require a minimum of three samples with three different levels of related substance.

Instead of using spike samples (as in accuracy determination), drug product lots that are representative of the commercial products should be used for precision (repeatability, intermediate precision). This is to ensure that the commercial drug product is used in at least one part of the method validation and that the repeatability results are representative of those that can be expected in the future.

**Intermediate Precision.** *ICH definition: Intermediate precision* expresses, within laboratories variations, different days, different analysts, different equipment, and so on.

Intermediate precision is to determine method precision in different experiments using different analysts and/or instrument setup. Similar to that of repeatability, one should evaluate the results of individual related substances, total related substances, and the consistency of related substance profiles in all experiments. The percent RSD and confidence level of these results are reported to illustrate the intermediate precision.

**Reproducibility.** *ICH definition: Reproducibility* expresses the precision between laboratories (collaborative studies are generally used, for standardization of methodology).

This is an optional validation parameter that requires demonstration of laboratory-to-laboratory variation only if multiple laboratories use the same procedure. The reproducibility data can be obtained during method transfer between laboratories.

**3.3.6 Range**

*ICH definition: The range of an analytical procedure is the interval between the upper and lower concentrations (amounts) of analytes in the sample (including these concentrations) for which it has been demonstrated that the analytical procedure has a suitable level of precision, accuracy, and linearity (Figure 3.13).*

Typically, linearity and accuracy determination covers a wide concentration range (e.g., 50% of the ICH reporting limit to 150% of specification). However, the concentration range for precision will be limited by the availability of sample of different related substance levels. Therefore, to ensure an appropriate method validation range with respect to precision, it is critical to use samples of low and high levels of related substance in precision experiments (e.g., fresh and stressed samples).

**3.3.7 Robustness**

*ICH definition: The robustness of an analytical procedure is a measure of its capacity to remain unaffected by small but deliberate variations in method parameters and provides an indication of its reliability during normal use.*
Table 5.1. Testing for Repeatability

<table>
<thead>
<tr>
<th>Method</th>
<th>Testing</th>
<th>Acceptance Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assay</td>
<td>Six sample determinations at the nominal concentration.</td>
<td>The variability of the automated method should not be more than 2.0% (RSD) or less than the manual method.</td>
</tr>
<tr>
<td>Degradation and impurity</td>
<td>Six determinations of a sample spiked at specification limit with available impurities.</td>
<td>The variability of the individual impurities is not more than 15.0% (RSD), and total impurities is not more than 10.0% (RSD) or less than the existing manual method, whichever is greater.</td>
</tr>
<tr>
<td>Content uniformity</td>
<td>Ten sample determinations at the nominal concentration.</td>
<td>The variability of the automated method should be not more than 6.0% (RSD).</td>
</tr>
<tr>
<td>Dissolution</td>
<td>Six sample determinations at the nominal concentration.</td>
<td>The variability of the automated method should be not more than 6.0% (RSD) at the Q point.</td>
</tr>
</tbody>
</table>

Repeatability. Repeatability expresses the precision under the same operating conditions over a short interval of time. The recommended testing for automated content uniformity, assays, degradation and impurity methods, and dissolution methods are listed in Table 5.1.

Intermediate Precision. Intermediate precision expresses within-laboratory variation and is generally performed on different days using different analysts, equipment, and sample preparations. This test may not be applicable if the laboratory has only one workstation. Additionally, this test may not be appropriate for automated workstations that are operating under the same environment and controls within a laboratory. This assumption is made on the basis that the automated workstations are identical (i.e., same configuration, same software and hardware) and that they have been suitably qualified and maintained to a consistent standard and operate under a similar climatic environment. The influence of the analyst is reduced to the preparation of solvents, and this should be covered by the robustness studies.

Reproducibility. Reproducibility expresses the precision between laboratories and would usually involve technical transfer of methods to laboratories in different geographical locations. The recommended testing for content uniformity, assays, degradation and impurity methods, and dissolution methods are listed in Table 5.2.
8

BIOANALYTICAL METHOD VALIDATION

Fabio Garofolo, Ph.D.
Vicuron Pharmaceuticals, Inc.

8.1 INTRODUCTION

8.1.1 Definition of Bioanalytical Method Validation

Bioanalytical method validation is a procedure employed to demonstrate that an analytical method used for quantification of analytes in a biological matrix is reliable and reproducible to achieve its purpose: to quantify the analyte with a degree of accuracy and precision appropriate to the task. Validation data, through specific laboratory investigations, demonstrate that the performance of a method is suitable and reliable for the analytical applications intended. The quantitative approach used in bioanalytical methods involves the use of a standard curve method with internal standard. In this approach the analyte concentration can be assigned by referring the response to other samples, called calibrators or calibration standards. In addition to the samples of unknown concentration, the bioanalytical set includes the calibration standards, and samples containing no analyte, called blanks, to assure that there are no interferences in the matrix. Accuracy and precision of the method are calculated using the back-calculated concentrations of samples of known composition called quality control samples (QCs). The calibrator standards and quality control samples should be prepared in the same matrix as the actual samples.

All these checks should be performed to guarantee the reliability of selective and sensitive bioanalytical method before applying them for the quantitative
8

BIOANALYTICAL METHOD VALIDATION

FABIO GAROFOLO, PH.D.
Vicuron Pharmaceuticals, Inc.

8.1 INTRODUCTION

8.1.1 Definition of Bioanalytical Method Validation

Bioanalytical method validation is a procedure employed to demonstrate that an analytical method used for quantification of analytes in a biological matrix is reliable and reproducible to achieve its purpose: to quantify the analyte with a degree of accuracy and precision appropriate to the task. Validation data, through specific laboratory investigations, demonstrate that the performance of a method is suitable and reliable for the analytical applications intended. The quantitative approach used in bioanalytical methods involves the use of a standard curve method with internal standard. In this approach the analyte concentration can be assigned by referring the response to other samples, called calibrators or calibration standards. In addition to the samples of unknown concentration, the bioanalytical set includes the calibration standards, and samples containing no analyte, called blanks, to assure that there are no interferences in the matrix. Accuracy and precision of the method are calculated using the back-calculated concentrations of samples of known composition called quality control samples (QCs). The calibrator standards and quality control samples should be prepared in the same matrix as the actual samples.

All these checks should be performed to guarantee the reliability of selective and sensitive bioanalytical method before applying them for the quantitative...
developed. This technical advancement leads to the use of commercial hyphenated mass spectrometric techniques and automation as preferred instrumentation for bioanalytical methods. LC-MS/MS assays replaced the conventional LC and GC, and the use of multiwell plates, automated robotic sample processing (Multiprobe, Tomtec), and online extraction techniques (Prospeck, Turbulent Flow Chromatography) took over manual sample preparation procedures.

The worldwide use of these new powerful bioanalytical techniques, characterized by more rapid throughput and increase in sensitivity [12,16] required a review of the original 1990 Washington report. The Guidelines of the 1990 Conference were initially reviewed during a meeting in June 1994 in Munich, Germany [2]. This meeting focused primarily on the critical and statistical evaluation of the acceptance criteria defined in the guidelines and on formulating recommendations to improve the guidelines.

A conference titled “Bioanalytical Methods Validation: A Revisit with a Decade of Progress” was held on January 2000, again in Washington, DC. The objective of this conference was to reach a consensus on what should be required in bioanalytical methods validation, and which procedures should be used to perform the validation [3]. The FDA 2001 Guidance for Industry on Bioanalytical Method Validation [4] is based on the final report of both the 1990 and 2000 Washington conferences. At the beginning of this document the FDA states very clearly that its guidance for bioanalytical method validation represents its current thinking on this topic and that an alternative approach may be used if such an approach satisfies the requirements of applicable statutes and regulations [4]. This statement allows bioanalytical laboratories to adjust or modify the FDA recommendations depending on the specific type of bioanalytical method used.

Compliance with the FDA guidance can be considered a minimum requirement to test the performance of a bioanalytical method. Due to the fact that the validation process should simulate closely sample analysis, the real and decisive final test for a “validated” method will always be the sample analysis itself. It is possible that even if it passes all the validation criteria, a bioanalytical method may not be reliable for the analysis of actual samples. This undesirable situation could happen when actual samples (in vivo samples) contain new interferences not present in the spiked samples (in vitro samples) due to a metabolic process and/or other biotransformations. For this reason, bioanalytical laboratories could decide to use more stringent criteria and procedures and/or use actual sample during the method development to further guarantee the validity of the validated methods.

In the following section we summarize the current general recommendations for bioanalytical method validation practices according to the FDA guidelines, with other alternative approaches to be discussed later.

8.2 CURRENT VALIDATION PRACTICE

The validation procedures for bioanalytical methods are in continuous evolution since bioanalytical methods are constantly undergoing changes in improvements,
and in many instances they are at the cutting edge of the technology. An overview of the FDA Guidance for Industry, Bioanalytical Methods Validation, May 2001 [4], is reported here as a reference for current validation practice.

8.2.1 Definitions

As the first step in understanding the procedure used for the validation of bioanalytical methods, it is important to have clearly in mind definitions of the analytical terms used.

- **Accuracy**: the degree of closeness of the determined value to the nominal or known true value under prescribed conditions. This is sometimes termed *trueness*.
- **Analyte**: a specific chemical moiety being measured, which can be intact drug, biomolecule, or its derivative, metabolite, and/or degradation product in a biological matrix.
- **Analytical run (or batch)**: a complete set of analytical and study samples with the appropriate number of standards and QCs for their validation. Several runs (or batches) may be completed in one day, or one run (or batch) may take several days to complete.
- **Biological matrix**: a discrete material of biological origin that can be sampled and processed in a reproducible manner. Examples are blood, serum, plasma, urine, feces, saliva, sputum, and various discrete tissues.
- **Stock solutions**: the original solutions prepared directly by weighing the reference standard of the analyte and dissolving it in the appropriate solvent. Usually, stock solutions are prepared at a concentration of 1 mg/mL in methanol and kept refrigerated at −20°C if there are no problems of stability or solubility.
- **Calibration standard**: a biological matrix to which a known amount of analyte has been added or *spiked*. Calibration standards are used to construct calibration curves from which the concentrations of analytes in QCs and in unknown study samples are determined.
- **Internal standard**: test compound(s) (e.g., structurally similar analog, stable labeled compound) added to both calibration standards and samples at known and constant concentration to facilitate quantification of the target analyte(s).
- **Limit of detection (LOD)**: the lowest concentration of an analyte that the bioanalytical procedure can reliably differentiate from background noise.
- **Lower limit of quantification (LLOQ)**: the lowest amount of an analyte in a sample that can be determined quantitatively with suitable precision and accuracy.
- **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.
- **Method**: a comprehensive description of all procedures used in sample analysis.

- **Precision**: the closeness of agreement (degree of scatter) between a series of measurements obtained from multiple sampling of the same homogeneous sample under the prescribed conditions.

- **Processed sample**: the final extract (prior to instrumental analysis) of a sample that has been subjected to various manipulations (e.g., extraction, dilution, concentration).

- **Quantification range**: the range of concentration, including ULOQ and LLOQ, that can be quantified reliably and reproducibly with accuracy and precision through the use of a concentration–response relationship.

- **Recovery**: the extraction efficiency of an analytical process, reported as a percentage of the known amount of an analyte carried through the sample extraction and processing steps of the method.

- **Reproducibility**: the precision between two laboratories. It also represents precision of the method under the same operating conditions over a short period of time.

- **Sample**: a generic term encompassing controls, blanks, unknowns, and processed samples, as described below:
  - **Blank**: a sample of a biological matrix to which no analytes have been added that is used to assess the specificity of the bioanalytical method.
  - **Quality control sample (QC)**: 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 analyzed in an individual batch.
  - **Unknown**: a biological sample that is the subject of the analysis.
  - **Selectivity**: the ability of the bioanalytical method to measure and differentiate the analytes in the presence of components that may be expected to be present. These could include metabolites, impurities, degradants, or matrix components.
  - **Stability**: the chemical stability of an analyte in a given matrix under specific conditions for given time intervals.
  - **Standard curve**: the relationship between the experimental response value and the analytical concentration (also called a *calibration curve*).
  - **System suitability**: determination of instrument performance (e.g., sensitivity and chromatographic retention) by analysis of a reference standard prior to running the analytical batch.
  - **Upper limit of quantification (ULOQ)**: the highest amount of an analyte in a sample that can be determined quantitatively with precision and accuracy.

- **Validation**
  - **Full validation**: establishment of all validation parameters to apply to sample analysis for the bioanalytical method for each analyte.
  - **Partial validation**: modification of validated bioanalytical methods that do not necessarily call for full revalidation.
8.2.2 Selectivity

For selectivity, there should be evidence that the substance being quantified is the intended analyte. Therefore, analyses of blank samples of the appropriate biological matrix (plasma, urine, or other matrix) should be obtained from at least six sources. Each blank sample should be tested for interference, and selectivity should be ensured at the lower limit of quantification (LLOQ).

Potential interfering substances in a biological matrix include endogenous matrix components, metabolites, decomposition products, and in the actual study, concomitant medication. Whenever possible, the same biological matrix as the matrix in the intended samples should be used for validation purposes. For tissues of limited availability, such as bone marrow, physiologically appropriate proxy matrices can be substituted. Method selectivity should be evaluated during method development and method validation and can continue during the analysis of actual study samples.

As with chromatographic methods, microbiological and ligand-binding assays should be shown to be selective for the analyte. The following recommendations for dealing with two selectivity issues should be considered:

1. Specific interference from substances physiochemically similar to the analyte:
   a. Cross-reactivity of metabolites, concomitant medications, or endogenous compounds should be evaluated individually and in combination with the analyte of interest.
   b. Whenever possible, the immunoassay should be compared with a validated reference method (such as LC-MS) using incurred samples and predetermined criteria for agreement of accuracy of immunoassay and reference method.
   c. The dilutional linearity to the reference standard should be assessed using study (incurred) samples.
   d. Selectivity may be improved for some analytes by incorporation of separation steps prior to immunoassay.

2. Nonspecific matrix effects:
   a. The standard curve in biological fluids should be compared with standard in buffer to detect matrix effects.
   b. Parallelism of diluted study samples should be evaluated with diluted standards to detect matrix effects.
   c. Nonspecific binding should be determined.
In consideration of high throughput analyses, including but not limited to multiplexing, multicolumn, and parallel systems, sufficient QC samples should be used to ensure control of the assay. The number of QC samples to ensure proper control of the assay should be determined based on the run size. The placement of QC samples should be considered judiciously in the run. At a minimum, three concentrations representing the entire range of the standard curve should be studied: one within three times the lower limit of quantification (LLOQ) (low QC sample), one near the center (middle QC), and one near the upper boundary of the standard curve (high QC). Reported method validation data and the determination of accuracy and precision should include all outliers; however, calculations of accuracy and precision, excluding values that are statistically determined as outliers, can also be reported.

8.2.6 Dilutions
The ability to dilute samples originally above the upper limit of the standard curve should be demonstrated by accuracy and precision parameters in the validation.

8.2.7 Recovery
Recovery of the analyte need not be 100%, but the extent of recovery of an analyte and of the internal standard should be consistent, precise, and reproducible. Recovery experiments should be performed by comparing the analytical results for extracted samples at three concentrations (low, medium, and high) with unextracted standards that represent 100% recovery.

It may be important to consider the variability of the matrix due to the physiological nature of the sample. In the case of LC-M/MS-based procedures, appropriate steps should be taken to ensure the lack of matrix effects throughout application of the method, especially if the nature of the matrix changes from the matrix used during method validation. For Microbiological and immunoassay, if separation is used prior to assay for study samples but not for standards, it is important to establish recovery and use it in determining results. In this case, possible approaches to assess efficiency and reproducibility of recovery are:

- The use of radiolabeled tracer analyte (quantity too small to affect the assay)
- The advance establishment of reproducible recovery
- The use of an internal standard that is not recognized by the antibody but can be measured by another technique

8.2.8 Stability
The stability of an analyte in a particular matrix and container system is relevant only to that matrix and container system and should not be extrapolated to other matrices and container systems. Stability procedures should evaluate the stability of the analytes during sample collection and handling, after long-term (frozen at
the intended storage temperature) and short-term (benchtop, room-temperature) storage, and after going through freeze–thaw cycles and the analytical process. Conditions used in stability experiments should reflect situations likely to be encountered during actual sample handling and analysis. The procedure should include an evaluation of analyte stability in stock solution. For compounds with potentially labile metabolites, the stability of analyte in matrix from dosed subjects (or species) should be confirmed. All stability determinations should use a set of samples prepared from a freshly made stock solution of the analyte in the appropriate analyte-free, interference-free biological matrix.

**Freeze and Thaw Stability.** Analyte stability should be determined after three freeze–thaw cycles. At least three aliquots at each of the low and high concentrations should be stored at the intended storage temperature for 24 h and thawed unassisted at room temperature. When thawed completely, the samples should be refrozen for 12 to 24 h under the same conditions. The freeze–thaw cycle should be repeated twice more, then analyzed on the third cycle. If an analyte is unstable at the intended storage temperature, the stability sample should be frozen at −70°C during the three freeze–thaw cycles.

**Short-Term Temperature Stability.** Three aliquots of each of the low and high concentrations should be thawed at room temperature and kept at this temperature from 4 to 24 h (based on the expected duration that samples will be maintained at room temperature in the intended study) and analyzed.

**Long-Term Stability.** The storage time in a long-term stability evaluation should exceed the time between the date of first sample collection and the date of last sample analysis. Long-term stability should be determined by storing at least three aliquots of each of the low and high concentrations under the same conditions as the study samples. The volume of samples should be sufficient for analysis on three separate occasions. The concentrations of all the stability samples should be compared to the mean of back-calculated values for the standards at the appropriate concentrations from the first day of long-term stability testing.

**Stock Solution Stability.** The stability of stock solutions of drug and the internal standard should be evaluated at room temperature for at least 6 h. If the stock solutions are refrigerated or frozen for the relevant period, the stability should be documented. After completion of the desired storage time, the stability should be tested by comparing the instrument response with that of freshly prepared solutions.

**Postpreparative Stability.** The stability of processed samples, including the resident time in the autosampler, should be determined. The stability of the drug and the internal standard should be assessed over the anticipated run time for the batch size in validation samples by determining concentrations on the basis of original calibration standards. Reinjection reproducibility should be evaluated to determine if an analytical run could be reanalyzed in the case of instrument failure.
to establish interlaboratory reliability. All modifications should be assessed to
determine the degree of validation recommended. Immunoassay reoptimization or
validation may be important when there are changes in key reagents.

8.2.10 Documentation
A specific, detailed description of the bioanalytical method should be written.
This can be in the form of a protocol, study plan, report, and/or standard operat-
ing procedure (SOP). All experiments used to make claims or draw conclusions
about the validity of the method should be presented in a report (method valida-
tion report).

8.3 COMMON PROBLEMS AND SOLUTIONS
The 2000 Washington Conference on Bioanalytical Method Validation [3] revi-
wed the progress, impact, and advances made during the last decade of bioanalyti-
cal methods validation since the first Washington conference in 1990. Hyphenated
mass spectrometric–based assays, ligand-based assay, and high-throughput sys-
tems were discussed in depth. However, there are still some controversies on
some scientific approaches and criteria used during the validation process. Some
of the most interesting issues discussed during the last 10 years are discussed in
the following paragraphs.

8.3.1 Definitions
The glossaries in the 1990 and 2000 Washington conference final reports [1,3]
define most of the analytical terms used in the validation of a method. How-
ever, internationally accepted definitions such as those by ISO or IUPAC already
exist and have been elaborated carefully over many years [2,6]. The definitions
reported in the 1990 and 2000 Washington conference final reports sometime
agree only partially with the ISO and IUPAC. Following are some examples
for comparison.

<table>
<thead>
<tr>
<th><strong>Limit of Detection</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2000 Conference:</strong> the lowest concentration of an analyte that a bioanalytical procedure can reliably differentiate from background noise.</td>
</tr>
<tr>
<td><strong>1990 Conference:</strong> the lowest concentration of an analyte that an analytical process can reliably differentiate from background levels.</td>
</tr>
<tr>
<td><strong>United States Pharmacopoeia:</strong> “thus, limit tests merely substantiate that the analyte concentration is above or below a certain level . . .”</td>
</tr>
<tr>
<td><strong>IUPAC:</strong> the concentration giving a signal three times the standard deviation of the blank.</td>
</tr>
</tbody>
</table>
**Lower Limit of Quantification**

2000 Conference: the lowest amount of an analyte in a sample that can be determined quantitatively with suitable precision and accuracy.

Comment:

- This definition is connected with the definition of sensitivity of the method as the concentration of the lowest standard with a coefficient of variance (CV) ≤ 20%.

1990 Conference: the lowest concentration of an analyte that can be measured with a stated level of confidence.

IUPAC: the concentration that gives rise to a signal 10 times the standard deviation of the blank.

**Upper Limit of Quantification**

2000 Conference: the highest amount of an analyte in a sample that can be determined quantitatively with precision and accuracy.

Comment:

- From a practical point of view this definition can be interpreted as being imposed by the linear boundary of the calibration curve (quadratic behavior) due to saturation of the detector or/and ion suppression effect and/or contamination for low-level samples (carryover) (see Section 8.3.7).

**Accuracy**

2000 Conference: the degree of closeness of the determined value to the nominal or known true value under prescribed conditions. This is sometimes termed trueness.

Comments:

- This is expressed as percent relative error (% RE).
- RE may be positive, negative, or zero.
- \( \% \text{ RE} = \left( \frac{\text{mean value}}{\text{theoretical value}} - 1 \right) \times 100. \)
- In general, the measured concentration of sample of known composition compared to the theoretical concentration over an appropriate range of concentration is considered an indication of accuracy.
- Intraassay accuracy is the RE of the mean of the replicate analysis of a validation sample during a single validation batch.
- Interassay accuracy is the RE of the overall mean of the replicate analyses of a validation sample for all validation batches.

1990 Conference: the closeness of the determined value to the true value. Generally, recovery of added analyte over an appropriate range of concentrations is
### Lower Limit of Quantification

**2000 Conference:** the lowest amount of an analyte in a sample that can be determined quantitatively with suitable precision and accuracy.

**Comment:**

- This definition is connected with the definition of sensitivity of the method as the concentration of the lowest standard with a coefficient of variance (CV) ≤ 20%.

**1990 Conference:** the lowest concentration of an analyte that can be measured with a stated level of confidence.

**IUPAC:** the concentration that gives rise to a signal 10 times the standard deviation of the blank

### Upper Limit of Quantification

**2000 Conference:** the highest amount of an analyte in a sample that can be determined quantitatively with precision and accuracy.

**Comment:**

- From a practical point of view this definition can be interpreted as being imposed by the linear boundary of the calibration curve (quadratic behavior) due to saturation of the detector or ion suppression effect and/or contamination for low-level samples (carryover) (see Section 8.3.7).

### Accuracy

**2000 Conference:** the degree of closeness of the determined value to the nominal or known true value under prescribed conditions. This is sometimes termed *trueness*.

**Comments:**

- This is expressed as percent relative error (% RE).
- RE may be positive, negative, or zero.
- \[ % \text{RE} = \left(\frac{\text{mean value}}{\text{theoretical value}} - 1\right) \times 100. \]
- In general, the measured concentration of sample of known composition compared to the theoretical concentration over an appropriate range of concentration is considered an indication of accuracy.
- Intraassay accuracy is the RE of the mean of the replicate analysis of a validation sample during a single validation batch.
- Interassay accuracy is the RE of the overall mean of the replicate analyses of a validation sample for all validation batches.

**1990 Conference:** the closeness of the determined value to the true value. Generally, recovery of added analyte over an appropriate range of concentrations is
**Lower Limit of Quantification**

*2000 Conference:* the lowest amount of an analyte in a sample that can be determined quantitatively with suitable precision and accuracy.

*Comment:*

- This definition is connected with the definition of sensitivity of the method as the concentration of the lowest standard with a coefficient of variance (CV) $\leq 20\%$.

*1990 Conference:* the lowest concentration of an analyte that can be measured with a stated level of confidence.

*IUPAC:* the concentration that gives rise to a signal 10 times the standard deviation of the blank

**Upper Limit of Quantification**

*2000 Conference:* the highest amount of an analyte in a sample that can be determined quantitatively with precision and accuracy.

*Comment:*

- From a practical point of view this definition can be interpreted as being imposed by the linear boundary of the calibration curve (quadratic behavior) due to saturation of the detector or/and ion suppression effect and/or contamination for low-level samples (carryover) (see Section 8.3.7).

**Accuracy**

*2000 Conference:* the degree of closeness of the determined value to the nominal or known true value under prescribed conditions. This is sometimes termed *trueness*.

*Comments:*

- This is expressed as percent relative error (% RE).
- RE may be positive, negative, or zero.
- $% \text{ RE} = \left[\frac{\text{mean value}}{\text{theoretical value}} - 1\right] \times 100$.
- In general, the measured concentration of sample of known composition compared to the theoretical concentration over an appropriate range of concentration is considered an indication of accuracy.
- Intraassay accuracy is the RE of the mean of the replicate analysis of a validation sample during a single validation batch.
- Interassay accuracy is the RE of the overall mean of the replicate analyses of a validation sample for all validation batches.

*1990 Conference:* the closeness of the determined value to the true value. Generally, recovery of added analyte over an appropriate range of concentrations is
taken as an indication of accuracy. Whenever possible, the concentration range chosen should bracket the concentration of interest.

*ISO:* the closeness of agreement between the test result and the accepted reference value.

**Comment:**

- The definition of accuracy reported in the 1990 Washington conference glossary was partially reformulated in the 2000 conference. The first sentence of both definitions is close to the ISO definition. Some disagreements were raised [2] on the second sentence of the 1990 definition during the 1994 meeting in Germany: “It is correct that recovery can be taken as an indication that a method is accurate but it is no more than an indication. Inclusion of recovery in a definition of accuracy may lead some analysts to conclude that adequate recovery always means that a method is accurate and that, of course, is not true. Suppose that a method is not selective and that some interference is also measured. The result will then be a certain (approximately the same) amount too high in both the unspiked and the spiked sample. However, the difference between the two results from which the recovery is calculated, will be correct, leading to the false conclusion that the method is accurate.”

### 8.3.2 Selectivity/Specificity

**Selectivity** is the ability of the bioanalytical method to measure and differentiate the analytes in the presence of components that may be expected to be present. **Specificity** is the ability to assess unequivocally the analyte in the presence of components that may be expected to be present. In general, analytical methods are selective, and only in some cases also specific (e.g., an LC-MS/MS bioanalytical method is highly selective but not always also specific because it could be possible to find in the complex biological matrix an interference with the same retention time, molecular weight, and main fragment of the analyte of interest). Even if the 2000 Washington conference focuses only on selectivity, it is up to bioanalytical laboratories to differentiate in their documentation between selectivity and specificity or consider them equivalent and use them interchangeably.

A general approach to prove the selectivity (specificity) of the method is to verify that:

- The response of interfering peaks at the retention time of the analyte is less than 20% of the response of an LLOQ standard, or the response at the LLOQ concentration is at least five times greater than any interference in blanks at the retention time of the analyte.
- The responses of interfering peaks at the retention time of the internal standard are $\leq 5\%$ of the response of the concentration of the internal standard used in the studies.
An analysis of the SFSTP guide on validation of chromatographic bioanalytical methods: progresses and limitations

B. Boulanger a, P. Chiap b,*, W. Dewe a, J. Crommen b, Ph. Hubert b

a Lilly Development Centre, Statistical and Mathematical Sciences, rue Granbompré, 11, B-1348 Mont-Saint-Guibert, Belgium
b Department of Analytical Pharmaceutical Chemistry, Institute of Pharmacy, University of Liège, CHU, B36, B-4000 Liège 1, Belgium

Received 22 May 2002; received in revised form 23 September 2002; accepted 26 September 2002

Abstract

The Société Française des Sciences et Techniques Pharmaceutiques (SFSTP) published in 1997 a guide on the validation of chromatographic bio-analytical methods, which introduces new concepts in three different areas: stages of the validation, test of acceptability of a method and design of experiments to perform. In ‘stages of validation’, the SFSTP guide requires two phases to validate a method. The first phase, called ‘pre-validation’, is intended to (1) identify the model to use for the calibration curve; (2) evaluate the limits of quantitation; and (3) provide good estimates of the precision and bias of the method before designing the ‘validation’ phase per se. In the ‘test of acceptability’, the use of the interval hypotheses is envisaged by the SFSTP guide, not on the parameters of bias and precision, but on individual results by mixing mean bias and intermediate precision in a single test. The SFSTP guide also avoids the use of Satterthwaite’s df for testing the acceptability. The reasons for those choices are discussed extensively. In ‘design of experiments’, much effort has been devoted to improving the quality of results by optimally designing and sizing the experiments to perform. The rationale for using near D-optimal designs for the calibration curve is demonstrated and sample sizes are proposed to correctly size the validation experiments.

© 2003 Elsevier Science B.V. All rights reserved.

Keywords: Bioanalysis; Method validation; Statistics

1. Introduction

Before using an analytical method for quantitative determinations of drugs and their metabolites, an applicant laboratory must first demonstrate that the envisaged method fulfils a number of performance criteria. Since the publications of the ‘Washington Conference’ [1] and the ICH Guidelines on Validation of Analytical Methods Q2A and Q2B [2,3], which list the performance criteria to reach from a regulatory point of view, many laboratories have started to redesign their processes by involving analysts and statisticians, in order to define strategies that will allow the fulfilment of the regulatory requirements, while

* Corresponding author.
E-mail address: p.chiap@ulg.ac.be (P. Chiap).
An analysis of the SFSTP guide on validation of chromatographic bioanalytical methods: progresses and limitations

B. Boulanger\textsuperscript{a}, P. Chiap\textsuperscript{b, *}, W. Dewe\textsuperscript{a}, J. Crommen\textsuperscript{b}, Ph. Hubert\textsuperscript{b}

\textsuperscript{a} Lilly Development Centre, Statistical and Mathematical Sciences, rue Granbompré, 11, B-1348 Mont-Saint-Guibert, Belgium
\textsuperscript{b} Department of Analytical Pharmaceutical Chemistry, Institute of Pharmacy, University of Liège, CHU, B36, B-4000 Liège 1, Belgium

Received 22 May 2002; received in revised form 23 September 2002; accepted 26 September 2002

Abstract

The Société Française des Sciences et Techniques Pharmaceutiques (SFSTP) published in 1997 a guide on the validation of chromatographic bio-analytical methods, which introduces new concepts in three different areas: stages of the validation, test of acceptability of a method and design of experiments to perform. In ‘stages of validation’, the SFSTP guide requires two phases to validate a method. The first phase, called ‘prevalidation’, is intended to (1) identify the model to use for the calibration curve; (2) evaluate the limits of quantitation; and (3) provide good estimates of the precision and bias of the method before designing the ‘validation’ phase per se. In the ‘test of acceptability’, the use of the interval hypotheses is envisaged by the SFSTP guide, not on the parameters of bias and precision, but on individual results by mixing mean bias and intermediate precision in a single test. The SFSTP guide also avoids the use of Satterthwaite’s df for testing the acceptability. The reasons for those choices are discussed extensively. In ‘design of experiments’, much effort has been devoted to improving the quality of results by optimally designing and sizing the experiments to perform in validation. The rationale for using near D-optimal designs for the calibration curve is demonstrated and sample sizes are proposed to correctly size the validation experiments.

Keywords: Bioanalysis; Method validation; Statistics

1. Introduction

Before using an analytical method for quantitative determinations of drugs and their metabolites, an applicant laboratory must first demonstrate that the envisaged method fulfils a number of performance criteria. Since the publications of the ‘Washington Conference’ [1] and the ICH Guidelines on Validation of Analytical Methods Q2A and Q2B [2,3], which list the performance criteria to reach from a regulatory point of view, many laboratories have started to redesign their processes by involving analysts and statisticians, in order to define strategies that will allow the fulfilment of the regulatory requirements, while...
An analysis of the SFSTP guide on validation of chromatographic bioanalytical methods: progresses and limitations

B. Boulanger, P. Chiap, W. Dewe, J. Crommen, Ph. Hubert

Lilly Development Centre, Statistical and Mathematical Sciences, rue Granbompré, 11, B-1348 Mont-Saint-Guibert, Belgium
Department of Analytical Pharmaceutical Chemistry, Institute of Pharmacy, University of Liège, CHU, B36, B-4000 Liège 1, Belgium

Received 22 May 2002; received in revised form 23 September 2002; accepted 26 September 2002

Abstract

The Société Française des Sciences et Techniques Pharmaceutiques (SFSTP) published in 1997 a guide on the validation of chromatographic bio-analytical methods, which introduces new concepts in three different areas: stages of the validation, test of acceptability of a method and design of experiments to perform. In ‘stages of validation’, the SFSTP guide requires two phases to validate a method. The first phase, called ‘prevalidation’, is intended to (1) identify the model to use for the calibration curve; (2) evaluate the limits of quantitation; and (3) provide good estimates of the precision and bias of the method before designing the ‘validation’ phase per se. In the ‘test of acceptability’, the use of the interval hypotheses is envisaged by the SFSTP guide, not on the parameters of bias and precision, but on individual results by mixing mean bias and intermediate precision in a single test. The SFSTP guide also avoids the use of Satterthwaite’s df for testing the acceptability. The reasons for those choices are discussed extensively. In ‘design of experiments’, much effort has been devoted to improving the quality of results by optimally designing and sizing the experiments to perform in validation. The rationale for using near D-optimal designs for the calibration curve is demonstrated and sample sizes are proposed to correctly size the validation experiments.

1. Introduction

Before using an analytical method for quantitative determinations of drugs and their metabolites, an applicant laboratory must first demonstrate that the envisaged method fulfils a number of performance criteria. Since the publications of the ‘Washington Conference’ [1] and the ICH Guidelines on Validation of Analytical Methods Q2A and Q2B [2,3], which list the performance criteria to reach from a regulatory point of view, many laboratories have started to redesign their processes by involving analysts and statisticians, in order to define strategies that will allow the fulfilment of the regulatory requirements, while

* Corresponding author.
E-mail address: p.chiap@ulg.ac.be (P. Chiap).
being practicable and scientifically consistent. Some laboratories have probably been lucky in finding an easy way to reach the goals, most have certainly experienced, as we did, some frustrations while trying to cope with contradictory, sometimes scientifically irrelevant, requirements and definitions. As an indication of this difficulty to define reasonable practicable strategies to satisfy global regulatory requirements, laboratory constraints and scientific consistency, no guide has been published that entirely addresses that issue. For this reason, the ‘Société Française des Sciences et Techniques Pharmaceutiques (SFSTP)’ created in 1995 a Commission involving analysts and statisticians from the industry and the regulatory agencies with the objective of publishing a guide [4] that could be used by laboratories. The proposed guide has been validated in several real cases before being published and practical applications are now available [5,6] that provide the analyst, on the one hand, with a better understanding on the way to proceed and on the other hand, real data for qualifying his own computations that he could perform using a commercial spreadsheet.

The SFSTP guide does not constitute a final end point, but on the contrary, was envisaged as a large basis to pave the way for developments that are expected from readers and analysts that will practice the guide. On one hand, since the publication of the guide in 1997, members of the SFSTP Commission already have some modifications or warnings to propose in order to initiate a continuous process of improvements. On the other hand, many choices and decisions that have been taken in this guide constitute disruptive progresses compared to traditional ways to proceed in this area. Those choices must be clearly justified and understood because the guide is consistent as a whole and cannot be applied part by part. Finally, the SFSTP guide [4,7] does not cover all the topics or performance criteria imposed by the ICH, such as stability and robustness.

The objectives of the present article are precisely to identify and explain the progress permitted by the SFSTP guide, point out some of the limitations and suggest ways to overcome them.

2. Stages of validation

As pointed out by Smith and Sittampalam [8], the validation process involves four stages that are called by the authors ‘Concept’, ‘Performance’, ‘Operational’ and possibly ‘Cross Validation’. Behind the new words proposed, it is of initial importance to understand that the validation is a permanent process that starts from the very beginning of the life of the method until its retirement. In the Concept or development phase, the analyst must identify and evaluate the impact of potential sources of variability that could later alter the global quality of the results. The objective today in development is no more to find a method that ‘works’, nor to elaborate smartly an analytical method whose quality will have to be evaluated in a later stage; the objective becomes to build results of quality by means of an analytical method. In other words, questions about the bias, precision and robustness must conduct the actions of the analyst developing a new method and no more focus its efforts only on some performance criteria, such as minimal resolution or maximal retention (migration) time in the case of chromatographic or electrophoretic methods. The ability of an analytical method to provide individual determinations of high quality, i.e. measurements close from the true content of a sample, should be the very endpoint every developer has to focus on.

The SFSTP guide unfortunately does not explicitly put a great emphasis on the development phase and might give the impression that the ‘validation’ is only seen as a sequence of experiments and calculations to perform to successfully reach an endpoint that is the documentation step. The SFSTP guide indirectly addresses the issue of the development since, as clearly stated, preliminary knowledge or a priori on the performance of the method must be available before properly starting the characterisation stage. This formal validation stage must be seen as a set of experiments that will confirm the regulatory agencies and the analyst himself that the method can indeed be used for its intended purpose. The validation phase can absolutely not be envisaged as a mean to estimate the performance of the method. If nothing or very little is known about the bias, the
being practicable and scientifically consistent. Some laboratories have probably been lucky in finding an easy way to reach the goals, most have certainly experienced, as we did, some frustrations while trying to cope with contradictory, sometimes scientifically irrelevant, requirements and definitions. As an indication of this difficulty to define reasonable practicable strategies to satisfy global regulatory requirements, laboratory constraints and scientific consistency, no guide has been published that entirely addresses that issue. For this reason, the ‘Société Française des Sciences et Techniques Pharmaceutiques (SFSTP)’ created in 1995 a Commission involving analysts and statisticians from the industry and the regulatory agencies with the objective of publishing a guide [4] that could be used by laboratories. The proposed guide has been validated in several real cases before being published and practical applications are now available [5,6] that provide the analyst, on the one hand, with a better understanding on the way to proceed and on the other hand, real data for qualifying his own computations that he could perform using a commercial spreadsheet.

The SFSTP guide does not constitute a final end point, but on the contrary, was envisaged as a large basis to pave the way for developments that are expected from readers and analysts that will practice the guide. On one hand, since the publication of the guide in 1997, members of the SFSTP Commission already have some modifications or warnings to propose in order to initiate a continuous process of improvements. On the other hand, many choices and decisions that have been taken in this guide constitute disruptive progresses compared to traditional ways to proceed in this area. Those choices must be clearly justified and understood because the guide is consistent as a whole and cannot be applied part by part. Finally, the SFSTP guide [4,7] does not cover all the topics or performance criteria imposed by the ICH, such as stability and robustness.

The objectives of the present article are precisely to identify and explain the progress permitted by the SFSTP guide, point out some of the limitations and suggest ways to overcome them.

2. Stages of validation

As pointed out by Smith and Sittampalam [8], the validation process involves four stages that are called by the authors ‘Concept’, ‘Performance’, ‘Operational’ and possibly ‘Cross Validation’. Behind the new words proposed, it is of initial importance to understand that the validation is a permanent process that starts from the very beginning of the life of the method until its retirement. In the concept or development phase, the analyst must identify and evaluate the impact of potential sources of variability that could later alter the global quality of the results. The objective today in development is no more to find a method that ‘works’, nor to elaborate smartly an analytical method whose quality will have to be evaluated in a later stage; the objective becomes to build results of quality by means of an analytical method. In other words, questions about the bias, precision and robustness must conduct the actions of the analyst developing a new method and no more focus its efforts only on some performance criteria, such as minimal resolution or maximal retention (migration) time in the case of chromatographic or electrophoretic methods. The ability of an analytical method to provide individual determinations of high quality, i.e. measurements close from the true content of a sample, should be the very endpoint every developer has to focus on.

The SFSTP guide unfortunately does not explicitly put a great emphasis on the development phase and might give the impression that the ‘validation’ is only seen as a sequence of experiments and calculations to perform to successfully reach an endpoint that is the documentation step. The SFSTP guide indirectly addresses the issue of the development since, as clearly stated, preliminary knowledge or a priori on the performance of the method must be available before properly starting the characterisation stage. This formal validation stage must be seen as a set of experiments that will confirm the regulatory agencies and the analyst himself that the method can indeed be used for its intended purpose. The validation phase can absolutely not be envisaged as a mean to estimate the performance of the method. If nothing or very little is known about the bias, the
ABSTRACT
Method validation is a process that demonstrates that a method will successfully meet or exceed the minimum standards recommended in the Food and Drug Administration (FDA) guidance for accuracy, precision, selectivity, sensitivity, reproducibility, and stability. This article discusses the validation of bioanalytical methods for small molecules with emphasis on chromatographic techniques. We present current thinking on validation requirements as described in the current FDA Guidance and subsequent 2006 Bioanalytical Methods Validation Workshop white paper.

KEYWORDS: bioanalytical, validation, precision, accuracy, sensitivity, selectivity, reproducibility

INTRODUCTION
Bioanalytical methods are used for the quantitation of drugs and their metabolites in biological matrices. In today’s drug development environment, highly sensitive and selective methods are required to quantify drugs in matrices such as blood, plasma, serum, or urine. Chromatographic methods (high-performance liquid chromatography [HPLC] or gas chromatography [GC]) have been widely used for the bioanalysis of small molecules, with liquid chromatography coupled to triple quadrupole mass spectrometry (LC/MS/MS) being the single most commonly used technology. After developing a method with desired attributes, the method is validated to establish that it will continue to provide accurate, precise, and reproducible data during study-sample analysis. Method validation is a process that demonstrates that the method will successfully meet or exceed the minimum standards recommended in the Food and Drug Administration (FDA) Guidance for accuracy, precision, selectivity, sensitivity, reproducibility, and stability. The validation is performed using a control matrix spiked with the compounds to be quantified. This article discusses the validation of bioanalytical methods for small molecules with emphasis on chromatographic techniques. We present current thinking on validation requirements as described in the current FDA Guidance and subsequent 2006 Bioanalytical Methods Validation Workshop white paper.

VALIDATION PARAMETERS
Bioanalytical methods can be developed in the laboratory conducting the validation or obtained from another laboratory or literature. The results from a method validation can be no better than the quality of the method that was developed. Thus, before beginning the method validation, it is important that the method is set up and tested in the laboratory. For methods obtained externally, modifications may be necessary to achieve the desired performance of the method relative to how it was developed originally. This process will help ensure that when validation begins, chances for its successful completion (and more important, successful sample analysis) are high. During method validation, values for validation parameters are obtained. The essential parameters required according to the FDA Guidance are selectivity, sensitivity, accuracy, precision, reproducibility, and stability. While obtaining these parameters, other parameters are also determined during validation (eg, extraction efficiency, calibration range and response function [linear or nonlinear], positional differences within an analytical run, and dilution integrity for analyzing above limit of quantitation [ALQ] samples). These validation parameters are described below in detail and are summarized in Table 1.

Selectivity
Selectivity or specificity should be assessed to show that the intended analytes are measured and that their quantitation is not affected by the presence of the biological matrix, known metabolites, degradation products, or co-administered drugs. Specificity should be determined for each analyte in the assay. Selectivity determination depends on the type of the assay as discussed below.

In assays wherein the intrinsic selectivity is low (eg, HPLC or GC with detection other than MS), it is necessary to confirm, using blank matrices from at least 6 independent sources, that the biological matrix will not interfere significantly with the assay. The same matrix as in samples should be used whenever possible. A proxy matrix is allowed if the sample matrix
Key Elements of Bioanalytical Method Validation for Small Molecules

Surendra Bansal\(^1\) and Anthony DeStefano\(^2\)

\(^1\)Hoffman-LaRoche Inc, Nutley, NJ
\(^2\)Procter & Gamble Pharmaceuticals, Mason, OH

**ABSTRACT**

Method validation is a process that demonstrates that a method will successfully meet or exceed the minimum standards recommended in the Food and Drug Administration (FDA) guidance for accuracy, precision, selectivity, sensitivity, reproducibility, and stability. This article discusses the validation of bioanalytical methods for small molecules with emphasis on chromatographic techniques. We present current thinking on validation requirements as described in the current FDA Guidance and subsequent 2006 Bioanalytical Methods Validation Workshop white paper.

**KEYWORDS:** bioanalytical, validation, precision, accuracy, sensitivity, selectivity, reproducibility

**INTRODUCTION**

Bioanalytical methods are used for the quantitation of drugs and their metabolites in biological matrices. In today’s drug development environment, highly sensitive and selective methods are required to quantify drugs in matrices such as blood, plasma, serum, or urine. Chromatographic methods (high-performance liquid chromatography [HPLC] or gas chromatography [GC]) have been widely used for the bioanalysis of small molecules, with liquid chromatography coupled to triple quadrupole mass spectrometry (LC/MS/MS) being the single most commonly used technology. After developing a method with desired attributes, the method is validated to establish that it will continue to provide accurate, precise, and reproducible data during study-sample analysis. Method validation is a process that demonstrates that the method will successfully meet or exceed the minimum standards recommended in the Food and Drug Administration (FDA) Guidance\(^1\) for accuracy, precision, selectivity, sensitivity, reproducibility, and stability. The validation is performed using a control matrix spiked with the compounds to be quantified. This article discusses the validation of bioanalytical methods for small molecules with emphasis on chromatographic techniques. We present current thinking on validation requirements as described in the current FDA Guidance\(^1\) and subsequent 2006 Bioanalytical Methods Validation Workshop white paper.\(^2\)

**VALIDATION PARAMETERS**

Bioanalytical methods can be developed in the laboratory conducting the validation or obtained from another laboratory or literature. The results from a method validation can be no better than the quality of the method that was developed. Thus, before beginning the method validation, it is important that the method is set up and tested in the laboratory. For methods obtained externally, modifications may be necessary to achieve the desired performance of the method relative to how it was developed originally. This process will help ensure that when validation begins, chances for its successful completion (and more important, successful sample analysis) are high. During method validation, values for validation parameters are obtained. The essential parameters required according to the FDA Guidance\(^1\) are selectivity, sensitivity, accuracy, precision, reproducibility, and stability. While obtaining these parameters, other parameters are also determined during validation (eg, extraction efficiency, calibration range and response function [linear or nonlinear], positional differences within an analytical run, and dilution integrity for analyzing above limit of quantitation [ALQ] samples). These validation parameters are described below in detail and are summarized in Table 1.

**Selectivity**

Selectivity or specificity should be assessed to show that the intended analytes are measured and that their quantitation is not affected by the presence of the biological matrix, known metabolites, degradation products, or co-administered drugs. Specificity should be determined for each analyte in the assay. Selectivity determination depends on the type of the assay as discussed below.

In assays wherein the intrinsic selectivity is low (eg, HPLC or GC with detection other than MS), it is necessary to confirm, using blank matrices from at least 6 independent sources, that the biological matrix will not interfere significantly with the assay. The same matrix as in samples should be used whenever possible. A proxy matrix is allowed if the sample matrix
of precision and accuracy samples is analyzed on a different column or instrument on one of the days of validation. This method is a good practice but is not required for all validations.

• Reproducibility using incurred samples: Reproducibility using incurred samples should be shown if samples are available. This test can be postponed and performed during sample analysis, where it is more important to prove the reproducibility of incurred samples analysis.

Stability

Several types of stability should be evaluated during the validation. Suggested experiments to determine stability are provided below. Alternate experiments that evaluate equivalent aspects of stability may be performed.

• Stock solution stability: The stability of the stock solutions of drug and internal standards should be evaluated at room temperature for at least 6 hours. If the stock solutions are kept refrigerated or frozen over a period of time, the stability over that period should be evaluated by comparing the response of the aged stock solution to that of a freshly prepared stock solution. Stock solution stability should be performed at one concentration in at least duplicate.

• Postpreparative (extracted samples/autosampler tray) stability: This stability is determined for ~48 to 96 hours to cover the anticipated run time for the analytical batch and to allow for delayed injection owing to unforeseen circumstances (eg, an instrument malfunction or the need to store samples over a weekend prior to analysis). The extracted QC samples (ready to inject) are kept at autosampler temperature for the established time and analyzed with fresh standards.

• Benchtop stability: Replicate (eg, triplicate) QC samples in matrix at a minimum of 2 concentrations are analyzed after keeping them at ambient temperature for 4 to 24 hours to cover at least the duration of time it takes to extract the samples. The observed sample concentrations are compared with their nominal values. This experiment can be combined with that for the extracted samples/autosampler tray stability above to demonstrate overall process stability, if desired.

• Freeze-thaw stability: QC samples in matrix at a minimum of 2 concentrations (eg, low and high QC concentrations) are frozen overnight, at normal storage temperature (eg, −20°C or −70°C) and thawed unassisted at room temperature. When completely thawed, the samples are frozen again at the same temperature for 12 to 24 hours and thawed. This freeze-thaw cycle is repeated 2 more times. After the third cycle, the samples are analyzed. The observed concentrations are compared with their nominal values. If an unacceptable level of degradation is observed, cycles 1 and 2 are repeated to determine where the instability occurs. The number of freeze-thaw cycles can be extended if needed.

• Freezer storage stability: During validation, stability at the nominal freezer storage temperature should be determined to the extent possible. However, longer term stability should be determined and appropriately documented, as discussed below.

• Postvalidation long-term stability: After validation is complete, long-term stability of the analyte(s) in the matrix should be determined by storing a sufficient number of QC samples at the required long-term storage temperature and analyzing them in at least triplicate at a minimum of 2 QC concentrations (eg, low and high QC concentrations). The long-term stability should be determined at several time points (eg, 1, 3, 6, 9, and 12 months) depending on the length of stability required. If possible, it is recommended that some stored in vivo samples are analyzed to assess the long-term stability of incurred samples at storage temperature. Upon obtaining the long-term stability data, the validation report can be amended to include the stability results or a separate report can be written to describe the long-term stability.

Extraction Efficiency (Recovery)

The extraction efficiency is a ratio of the detector response of an analyte from an extracted sample to the detector response of the analyte from an unextracted sample containing the same amount of analyte that was added to the extracted sample. The unextracted sample can be made up in solvents and is not taken through the extraction process. Alternatively, blank samples can be extracted and the extracts fortified with the analytes after extraction. These preparations represent 100% recovery during extraction. Extraction efficiency need not be very high, but it should be consistent, precise, and reproducible. Extraction efficiency can also be determined for the IS, and the ratio of the extraction efficiencies of the analyte and IS provides an IS-normalized extraction efficiency.

Calibration Range and Response Function

The relationship between the detector response and concentration should be demonstrated to be well defined and reproducible. A calibration curve should consist of a blank sample (matrix sample processed without the IS), a zero standard (matrix sample processed with internal standard), and 6 to 8
of precision and accuracy samples is analyzed on a different column or instrument on one of the days of validation. This method is a good practice but is not required for all validations.

- Reproducibility using incurred samples: Reproducibility using incurred samples should be shown if samples are available. This test can be postponed and performed during sample analysis, where it is more important to prove the reproducibility of incurred samples analysis.

**Stability**

Several types of stability should be evaluated during the validation. Suggested experiments to determine stability are provided below. Alternate experiments that evaluate equivalent aspects of stability may be performed.

- Stock solution stability: The stability of the stock solutions of drug and internal standards should be evaluated at room temperature for at least 6 hours. If the stock solutions are kept refrigerated or frozen over a period of time, the stability over that period should be evaluated by comparing the response of the aged stock solution to that of a freshly prepared stock solution. Stock solution stability should be performed at one concentration in at least duplicate.

- Postpreparative (extracted samples/autosampler tray) stability: This stability is determined for ~48 to 96 hours to cover the anticipated run time for the analytical batch and to allow for delayed injection owing to unforeseen circumstances (eg, an instrument malfunction or the need to store samples over a weekend prior to analysis). The extracted QC samples (ready to inject) are kept at autosampler temperature for the established time and analyzed with fresh standards.

- Benchtop stability: Replicate (eg, triplicate) QC samples in matrix at a minimum of 2 concentrations are analyzed after keeping them at ambient temperature for 4 to 24 hours to cover at least the duration of time it takes to extract the samples. The observed sample concentrations are compared with their nominal values. This experiment can be combined with that for the extracted samples/autosampler tray stability above to demonstrate overall process stability, if desired.

- Freeze-thaw stability: QC samples in matrix at a minimum of 2 concentrations (eg, low and high QC concentrations) are frozen overnight, at normal storage temperature (eg, −20°C or −70°C) and thawed unassisted at room temperature. When completely thawed, the samples are frozen again at the same temperature for 12 to 24 hours and thawed. This freeze-thaw cycle is repeated 2 more times. After the third cycle, the samples are analyzed. The observed concentrations are compared with their nominal values. If an unacceptable level of degradation is observed, cycles 1 and 2 are repeated to determine where the instability occurs. The number of freeze-thaw cycles can be extended if needed.

- Freezer storage stability: During validation, stability at the nominal freezer storage temperature should be determined to the extent possible. However, longer term stability should be determined and appropriately documented, as discussed below.

- Postvalidation long-term stability: After validation is complete, long-term stability of the analyte(s) in the matrix should be determined by storing a sufficient number of QC samples at the required long-term storage temperature and analyzing them in at least triplicate at a minimum of 2 QC concentrations (eg, low and high QC concentrations). The long-term stability should be determined at several time points (eg, 1, 3, 6, 9, and 12 months) depending on the length of stability required. If possible, it is recommended that some stored in vivo samples are analyzed to assess the long-term stability of incurred samples at storage temperature. Upon obtaining the long-term stability data, the validation report can be amended to include the stability results or a separate report can be written to describe the long-term stability.

**Extraction Efficiency (Recovery)**

The extraction efficiency is a ratio of the detector response of an analyte from an extracted sample to the detector response of the analyte from an unextracted sample containing the same amount of analyte that was added to the extracted sample. The unextracted sample can be made up in solvents and is not taken through the extraction process. Alternatively, blank samples can be extracted and the extracts fortified with the analytes after extraction. These preparations represent 100% recovery during extraction. Extraction efficiency need not be very high, but it should be consistent, precise, and reproducible. Extraction efficiency can also be determined for the IS, and the ratio of the extraction efficiencies of the analyte and IS provides an IS-normalized extraction efficiency.

**Calibration Range and Response Function**

The relationship between the detector response and concentration should be demonstrated to be well defined and reproducible. A calibration curve should consist of a blank sample (matrix sample processed without the IS), a zero standard (matrix sample processed with internal standard), and 6 to 8
nonzero standards. The number of standards can be increased for a complex curve or a curve covering a very large range. The simplest relationship that provides acceptable back-calculated concentrations for the standards should be used to fit the calibration curve. If a weighting factor is used, it should be defined during validation. The concentrations of calibration standards are back-calculated, and the residuals (difference between the back-calculated concentration of the calibration standard and its nominal concentration) determined. The residuals should be no more than ±15% at all concentrations except at the LLOQ level, where they can be up to ±20% of the nominal value. To accept an analytical run, at least 75% of the calibration standards should meet the stated acceptance criteria. Calibration standards not meeting the acceptance criteria should be eliminated from the calibration curve calculations. No extrapolation from the calibration curves is allowed, therefore the range of the calibration curve will be truncated if the end points on the calibration curve are eliminated.

**Positional Differences**

During a chromatographic analysis, samples are injected in sequence over several hours. Therefore, it is important to determine if the sample position in the chromatographic run sequence has an influence on the observed response (eg, if there is response change over the course of the run or any carryover is observed from previous samples). An evaluation of the situation should be done during the validation of the method and monitored during sample analysis. Procedures
Method Validation in Pharmaceutical Analysis

Edited by

J. Ermer and J. H. McB. Miller

WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, 2005
The following sections discuss parameters and calculations, which describe the performance of analytical procedures according to the ICH validation characteristics. The selection and discussion of these parameters and calculations reflect the experience of the authors and is primarily based on practical considerations. Their relevance will vary with the individual analytical application; some are also suitable for addressing questions other than validation itself. It is not intended to replace statistical textbooks, but the authors have tried to provide sufficient background information – always with the practical analytical application in mind – in order to make it easier for the reader to decide which parameters and tests are relevant and useful in his/her specific case. Precision is discussed first, because many other performance parameters are linked to analytical variability.

2.1 Precision

Joachim Ermer

“Precision is the closeness of agreement (degree of scatter) between a series of measurements obtained from multiple sampling of the same homogeneous sample under the prescribed conditions. Precision may be considered at three levels: repeatability, intermediate precision and reproducibility.” [1a]

Precision should be obtained preferably using authentic samples. As parameters, the standard deviation, the relative standard deviation (coefficient of variation) and the confidence interval should be calculated for each level of precision.

Repeatability expresses the analytical variability under the same operating conditions over a short interval of time (within-assay, intra-assay). At least nine determinations covering the specified range or six determinations at 100% test concentration should be performed.
Intermediate precision includes the influence of additional random effects within laboratories, according to the intended use of the procedure, for example, different days, analysts or equipment, etc.

Reproducibility, i.e., the precision between laboratories (collaborative or inter-laboratory studies), is not required for submission, but can be taken into account for standardisation of analytical procedures.

Before discussing the precision levels in detail, some fundamentals concerning the distribution of data are recalled. This is deemed to be very important for a correct understanding and evaluation of the following sections. For practical applications, a good understanding of the acceptable and achievable precision ranges is crucial. The section concludes with the description of some approaches used to obtain precision results.

### 2.1.1 Parameters Describing the Distribution of Analytical Data

#### 2.1.1.1 Normal Distribution

‘Measurements are inherently variable’ [16], i.e., the analytical data obtained scatter around the true value. The distribution of data can be visualised by histograms, i.e., plotting the frequency of the data within constant intervals (classes) throughout the whole data range observed. Such histograms can be generated using Microsoft Excel® (Tools/Data Analysis/Histogram; the Analysis ToolPak can be installed by means of Tools/Add-Ins). Usually, the number of classes corresponds approximately to the square root of the number of data. Figure 2.1-1 shows clearly that a large number of data is required to obtain a clear picture. The data were obtained by recording the absorbance of a drug substance test solution at 291 nm for 60 minutes with a sampling rate of 10/s. Of course, over such a long time, an instrumental drift cannot be avoided. From 15 minutes onwards, the drift in the absorbance values was constant. Various time segments were further investigated and for the drift between 35 and 60 minutes the lowest residual standard deviation of the regression line was observed. The data were corrected accordingly, i.e., the corrected data represent the scattering of the absorbance values around the regression line of the drift. The mean and standard deviation of these 15 000 data were calculated to be 692 and 0.1774 mAU, respectively. The very small relative standard deviation of 0.026 % represents only the detection variability of the spectrophotometer.

The usually assumed normal distribution, in physico-chemical analysis, could be confirmed for the data sets in the example, but even with 15 000 data the theoretical distribution cannot be achieved (Fig. 2.1-2). The normal distribution or Gaussian curve is bell-shaped and symmetrically centred around the mean (true value) for which the highest frequency is expected. The probability of measured data decreases with the distance from the true value and can be calculated with the probability density function (Eq. 2.1-1).
(as described in the control test), rather merely to injecting the same sample solution six times. This is also the reason for using **authentic samples** [1b], because only then can the analytical procedure be performed exactly as in the routine application. There may be exceptions, but these should be demonstrated or cautiously justified. For example, analysing degradants near the quantitation limit, where the variance contribution of the sample preparation can be neglected, injection precision and repeatability are identical (Figs. 2.1-7 and 2.1-8). For some applications, where precision can be regarded as less critical, such as in early development (see Chapter 5), or if the variability demands only a small part of specification range (less than approximately 10%), or if the expected content of impurities is far away from the specification limit, artificially prepared (spiked) samples may be used, allowing several validation characteristics (linearity, precision and accuracy) to be addressed simultaneously.

Repeatability can be calculated using Eq.( 2.1-2) from a larger number of repeatedly prepared samples (at least 6), or according to Eq.( 2.1-8) from a sufficient number of duplicate sample preparations. Calculations should not be performed with smaller number of data due to the large uncertainty involved (Fig. 2.1-4B). The true standard deviation may be up to 4.4 times greater than a result obtained from three determinations!

### 2.1.2.3 Intermediate Precision and Reproducibility

**Intermediate precision** includes the influence of additional random effects according to the intended use of the procedure in the same laboratory and can be regarded as an (initial) estimate for the long-term variability. Relevant factors, such as operator, instrument, and days should be varied. Intermediate precision is obtained from several independent series of applications of the (whole) analytical procedure to (preferably) authentic, identical samples. In case of relative techniques, the preparation and analysis of the reference standard is an important variability contribution. Therefore, it is not appropriate to determine intermediate precision from the peak area of the sample alone (analysed on different days or even several concentrations only), as is sometimes reported in validation literature. Apart from ignoring the contribution of the reference standard, any signal shift of the instrument will be falsely interpreted as random variability. In order to reflect the expected routine variability properly, the calibration must be performed exactly as described in the control test.

**Reproducibility**, according to the ICH definition is obtained varying further factors between laboratories and is particularly important in the assessment of ‘official’ compendial methods or if the method is applied at different sites. However, understood in the long-term perspective, both intermediate precision and reproducibility approach each other, at least in the same company. Reproducibility from collaborative trials can be expected to include additional contributions due to a probably larger difference of knowledge, experience, equipment, etc. among the participating laboratories.

**Analysis of variances**

It is very important to address intermediate precision/reproducibility appropriately as it is an estimate for the variability (and robustness) to be expected in long-term applications, such as in stability testing. According to ICH, standard deviations
"Specificity is the ability to assess unequivocally the analyte in the presence of components which may be expected to be present. Typically these might include impurities, degradants, matrix, etc. Lack of specificity of an individual procedure may be compensated by other supporting analytical procedure(s)" [1a].

With respect to identification, discrimination between closely related compounds likely to be present should be demonstrated by positive and negative samples. In the case of chromatographic assay and impurity tests, available impurities/degradants can be spiked at appropriate levels to the corresponding matrix or else degraded samples can be used. For assay, it can be demonstrated that the result is unaffected by the spiked material. Impurities should be separated individually and/or from other matrix components. Specificity can also be demonstrated by verification of the result with an independent analytical procedure.

In the case of chromatographic separation, resolution factors should be obtained for critical separation. Tests for peak homogeneity, for example, by diode array detection (DAD) or mass spectrometry (MS) are recommended.

There has been some controversial discussion about the terminology for this validation characteristic. In contrast to the ICH, most other analytical organisations define this as selectivity, whereas specificity is regarded in an absolute sense, as the “ultimate degree of selectivity” (IUPAC) [68]. Despite this controversy, there is a broad agreement that specificity/selectivity is the critical basis of each analytical procedure. Without a sufficient selectivity, the other performance parameters are meaningless. In order to maintain a consistent terminology, in the following ‘specificity’ is used as the general term for the validation characteristic, whereas ‘selective’ and ‘selectivity’ describe its qualitative grade. The latter is important to realise, because there is no absolute measure of selectivity, there is only an absence of evidence, no evidence of absence.

In contrast to chemical analysis, where each analytical procedure is regarded (and evaluated) separately, in pharmaceutical analysis, a whole range of control tests is used to evaluate a batch. Therefore, the performance of these individual analytical procedures can complement each other in order to achieve the required overall level of selectivity. For example, an assay by means of a less selective titration that will include impurities with the same functional groups, can be confirmed (or corrected) by a selective impurity determination by LC [1b].

Specificity is to be considered from the beginning of the method development, taking into account the properties of both analyte and sample (matrix). The (sufficiently) selective determination of the analyte can be achieved by appropriate sample
2.7

Robustness

Gerd Kleinschmidt

Although robustness of analytical procedures is generally noticed least of all, it is one of the most important validation parameters. Fortunately, in pharmaceutical analysis more and more attention is paid to it. Basically, robustness testing means to evaluate the ability of a method to perform effectively in a typical laboratory environment and with acceptable variations. Robustness definitions have been widely harmonised among international drug authorities, which is mainly the merit of the International Conference on Harmonisation (ICH).

2.7.1

Terminology and Definitions

Definitions provided by regulatory bodies, which play a significant role in the pharmaceutical world are itemised below.

2.7.1.1 International Conference on Harmonisation (ICH)

According to ICH Q2A [1a] “the robustness of an analytical procedure is a measure of its capacity to remain unaffected by small, but deliberate variations in method parameters and provides an indication of its reliability during normal usage”.

Furthermore, it is stated in ICH Q2B [1b], “The evaluation of robustness should be considered during the development phase and depends on the type of procedure under study. It should show the reliability of an analysis with respect to deliberate variations in method parameters. If measurements are susceptible to variations in analytical conditions, the analytical conditions should be suitably controlled or a precautionary statement should be included in the procedure. One consequence of the evaluation of robustness should be that a series of system suitability parameters (e.g., resolution test) is established to ensure that the validity of the analytical procedure is maintained whenever used”.

Additionally, the ICH guideline Q2B lists examples of typical variations such as extraction time or in case of liquid chromatography the mobile phase pH, the mobile phase composition and flow rate etc.

Even though these explanations are not very detailed, they guide an analyst on when and how to evaluate robustness. To decide what is small, but deliberate depends on the method and is the responsibility of the analyst.

2.7.1.2 Food and Drug Administration (FDA)

The FDA utilises the ICH definition for robustness and remarks that “data obtained from studies for robustness, though not usually submitted, are recommended to be included as part of method validation”. This is stated in the Reviewer Guidance “Validation of Chromatographic Methods” [3].

Corresponding to ICH, robustness testing “should be performed during development of the analytical procedure and the data discussed and / or submitted. In cases where an
Guidance for Industry
U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Veterinary Medicine (CVM)
May 2001

Bioanalytical Method Validation
A calibration curve should be generated for each analyte to assay samples in each analytical run and should be used to calculate the concentration of the analyte in the unknown samples in the run. The spiked samples can contain more than one analyte. An analytical run can consist of QC samples, calibration standards, and either (1) all the processed samples to be analyzed as one batch or (2) a batch composed of processed unknown samples of one or more volunteers in a study. The calibration (standard) curve should cover the expected unknown sample concentration range in addition to a calibrator sample at LLOQ. Estimation of concentration in unknown samples by extrapolation of standard curves below LLOQ or above the highest standard is not recommended. Instead, the standard curve should be redefined or samples with higher concentration should be diluted and reassayed. It is preferable to analyze all study samples from a subject in a single run.

Once the analytical method has been validated for routine use, its accuracy and precision should be monitored regularly to ensure that the method continues to perform satisfactorily. To achieve this objective, a number of QC samples prepared separately should be analyzed with processed test samples at intervals based on the total number of samples. The QC samples in duplicate at three concentrations (one near the LLOQ (i.e., #3 x LLOQ), one in midrange, and one close to the high end of the range) should be incorporated in each assay run. The number of QC samples (in multiples of three) will depend on the total number of samples in the run. The results of the QC samples provide the basis of accepting or rejecting the run. At least four of every six QC samples should be within "15% of their respective nominal value." Two of the six QC samples may be outside the "15% of their respective nominal value," but not both at the same concentration.

The following recommendations should be noted in applying a bioanalytical method to routine drug analysis:

- A matrix-based standard curve should consist of a minimum of six standard points, excluding blanks (either single or replicate), covering the entire range.

- Response Function: Typically, the same curve fitting, weighting, and goodness of fit determined during prestudy validation should be used for the standard curve within the study. Response function is determined by appropriate statistical tests based on the actual standard points during each run in the validation. Changes in the response function relationship between prestudy validation and routine run validation indicate potential problems.

- The QC samples should be used to accept or reject the run. These QC samples are matrix spiked with analyte.

- System suitability: Based on the analyte and technique, a specific SOP (or sample) should be identified to ensure optimum operation of the system used.
QC), and one approaching the high end of the range (high QC) should be incorporated into each run. The results of the QC samples provide the basis of accepting or rejecting the run. At least 67% (four out of six) of the QC samples should be within 15% of their respective nominal (theoretical) values; 33% of the QC samples (not all replicates at the same concentration) can be outside the ±15% of the nominal value. A confidence interval approach yielding comparable accuracy and precision is an appropriate alternative.

The minimum number of samples (in multiples of three) should be at least 5% of the number of unknown samples or six total QCs, whichever is greater.

- Samples involving multiple analytes should not be rejected based on the data from one analyte failing the acceptance criteria.
- The data from rejected runs need not be documented, but the fact that a run was rejected and the reason for failure should be recorded.

VII. DOCUMENTATION

The validity of an analytical method should be established and verified by laboratory studies, and documentation of successful completion of such studies should be provided in the assay validation report. General and specific SOPs and good record keeping are an essential part of a validated analytical method. The data generated for bioanalytical method establishment and the QCs should be documented and available for data audit and inspection. Documentation for submission to the Agency should include (1) summary information, (2) method development and establishment, (3) bioanalytical reports of the application of any methods to routine sample analysis, and (4) other information applicable to method development and establishment and/or to routine sample analysis.

A. Summary Information

Summary information should include:

- Summary table of validation reports, including analytical method validation, partial revalidation, and cross-validation reports. The table should be in chronological sequence, and include assay method identification code, type of assay, and the reason for the new method or additional validation (e.g., to lower the limit of quantitation).
- Summary table with a list, by protocol, of assay methods used. The protocol number, protocol title, assay type, assay method identification code, and bioanalytic report code should be provided.
- A summary table allowing cross-referencing of multiple identification codes should be provided (e.g., when an assay has different codes for the assay method, validation reports,
Bioanalytical method validation and its implications for forensic and clinical toxicology – A review

Abstract The reliability of analytical data is very important to forensic and clinical toxicologists for the correct interpretation of toxicological findings. This makes (bio)analytical method validation an integral part of quality management and accreditation in analytical toxicology. Therefore, consensus should be reached in this field on the kind and extent of validation experiments as well as on acceptance criteria for validation parameters. In this review, the most important papers published on this topic since 1991 have been reviewed. Terminology, theoretical and practical aspects as well as implications for forensic and clinical toxicology of the following validation parameters are discussed: selectivity (specificity), calibration model (linearity), accuracy, precision, limits, stability, recovery and ruggedness (robustness).

Keywords Method · Validation · Bioanalysis · Toxicology

Introduction

The reliability of analytical findings is a matter of great importance in forensic and clinical toxicology, as it is a prerequisite for correct interpretation of toxicological findings. Unreliable results might not only be contested in court, but could also lead to unjustified legal consequences for the defendant or to wrong treatment of the patient. The importance of validation, at least of routine analytical methods, can therefore hardly be overestimated. This is especially true in the context of quality management and accreditation, which have become matters of increasing importance in analytical toxicology in recent years. This is also reflected in the increasing requirements of peer-reviewed scientific journals concerning method validation. Therefore, this topic should be extensively discussed on an international level to reach a consensus on the extent of validation experiments and on acceptance criteria for validation parameters of bioanalytical methods in forensic and clinical toxicology.

Over the last decade, similar discussions have been going on in the closely related field of pharmacokinetic studies for registration of pharmaceuticals. This is reflected by the number of publications on this topic since 1991 have been reviewed. Terminology, theoretical and practical aspects as well as implications for forensic and clinical toxicology of the following validation parameters are discussed: selectivity (specificity), calibration model (linearity), accuracy, precision, limits, stability, recovery and ruggedness (robustness).

Important publications on validation (1991 to present)

A review on validation of bioanalytical methods was published by Karnes et al. in 1991, which was intended to provide guidance for bioanalytical chemists [1]. One
Abstract

The reliability of analytical data is very important to forensic and clinical toxicologists for the correct interpretation of toxicological findings. This makes (bio)analytical method validation an integral part of quality management and accreditation in analytical toxicology. Therefore, consensus should be reached in this field on the kind and extent of validation experiments as well as on acceptance criteria for validation parameters. In this review, the most important papers published on this topic since 1991 have been reviewed. Terminology, theoretical and practical aspects as well as implications for forensic and clinical toxicology of the following validation parameters are discussed: selectivity (specificity), calibration model (linearity), accuracy, precision, limits, stability, recovery and ruggedness (robustness).

Keywords

Method · Validation · Bioanalysis · Toxicology

Introduction

The reliability of analytical findings is a matter of great importance in forensic and clinical toxicology, as it is a prerequisite for correct interpretation of toxicological findings. Unreliable results might not only be contested in court, but could also lead to unjustified legal consequences for the defendant or to wrong treatment of the patient. The importance of validation, at least of routine analytical methods, can therefore hardly be overestimated. This is especially true in the context of quality management and accreditation, which have become matters of increasing importance in analytical toxicology in recent years. This is also reflected in the increasing requirements of peer-reviewed scientific journals concerning method validation. Therefore, this topic should be extensively discussed on an international level to reach a consensus on the extent of validation experiments and on acceptance criteria for validation parameters of bioanalytical methods in forensic and clinical toxicology.

Over the last decade, similar discussions have been going on in the closely related field of pharmacokinetic studies for registration of pharmaceuticals. This is reflected by the number of publications on this topic since 1991 have been reviewed. Terminology, theoretical and practical aspects as well as implications for forensic and clinical toxicology of the following validation parameters are discussed: selectivity (specificity), calibration model (linearity), accuracy, precision, limits, stability, recovery and ruggedness (robustness).

Keywords

Method · Validation · Bioanalysis · Toxicology

Important publications on validation (1991 to present)

A review on validation of bioanalytical methods was published by Karnes et al. in 1991, which was intended to provide guidance for bioanalytical chemists. One
The aim of our review is to present and compare the contents of the above mentioned publications on (bio)analytical method validation, and to discuss possible implications for forensic and clinical toxicology.

Terminology

The first problem encountered when studying literature on method validation are the different sets of terminology employed by different authors. A detailed discussion of this problem can be found in the review of Hartmann et al. [9]. Therein, it was proposed to adhere, in principle, to the terminology established by the ICH [13], except for accuracy, for which the use of a more detailed definition was recommended (cf. Accuracy). However, the ICH terminology lacked a definition for stability, which is an important parameter in bioanalytical method validation. Furthermore, the ICH definition of selectivity did not take into account interferences that might occur in bioanalysis (e.g. from metabolites). For both parameters, however, reasonable definitions were provided by Conference Report II [10].

Validation parameters

There is a general agreement that at least the following validation parameters should be evaluated for quantitative procedures: selectivity, calibration model (linearity), stability, accuracy (bias, precision) and limit of quantification. Additional parameters which might have to be evaluated include limit of detection, recovery, reproducibility and ruggedness (robustness) [2, 4–10, 12].

Selectivity (specificity)

In Conference Report II, selectivity was defined as follows: “Selectivity is the ability of the bioanalytical method to measure unequivocally and to differentiate the analyte(s) in the presence of components, which may be expected to be present”. Typically, these might include metabolites, impurities, degradants, matrix components, etc. [10]. This definition is very similar to the one established by the ICH [13], but takes into account the possible presence of metabolites, and thus is more applicable for bioanalytical methods.

There are two points of view on when a method should be regarded as selective. One way to establish method selectivity is to prove the lack of response in blank matrix [1, 2, 4–10, 12, 14]. The requirement established by the Conference Report [2] to analyse at least six different sources of blank matrix has become state of the art. However, this approach has been subject to criticism in the review of Hartmann et al., who stated from statistical considerations, that relatively rare interferences will remain undetected with a rather high probability [9]. For the same reason, Dadgar et al. proposed to evaluate at least 10–20 sources of blank samples [4]. However, in Conference Report II [10], even analysis of only one source of blank matrix was deemed acceptable, if
measurement uncertainty has been published by EURACHEM/CITAC [20].

Conclusion

There are only a few principle differences concerning validation of bioanalytical methods in the fields of pharmacokinetic studies and forensic and clinical toxicology. Therefore, it seems reasonable to base the discussion on validation in the field of toxicology on the experiences and consensus already existing in the closely related field of pharmacokinetic studies for registration of pharmaceuticals and focus the discussion on those parameters, which are of special importance for toxicologists, i.e. selectivity, LOD, LLOQ and stability.

References

13. International Conference on Harmonization (ICH) (1994) Validation of analytical methods: Definitions and terminology ICH Q1A
18. NCCLS (1999) Evaluation of precision performance of clinical chemistry devices; Approved guideline. NCCLS, Wayne, Pa., USA
Determination of loratadine and its active metabolite in human plasma by high-performance liquid chromatography with mass spectrometry detection

Laurian Vlase a, Silvia Imre b,* Dana Muntean a, Sorin E. Leucuta a

a Faculty of Pharmacy, University of Medicine and Pharmacy “Iuliu Hatieganu”, Emil Isac 13, RO-400023, Cluj-Napoca, Romania
b Faculty of Pharmacy, University of Medicine and Pharmacy, Gh. Marinescu 38, RO-540139, Tg. Mures, Romania

Received 5 June 2006; received in revised form 8 August 2006; accepted 10 August 2006
Available online 8 September 2006

Abstract

A new sensitive and selective liquid chromatography coupled with mass spectrometry (LC/MS/MS) method for quantification of loratadine (LOR) and its active metabolite descarboethoxyloratadine (DSL) in human plasma was validated. After addition of the internal standard, metoclopramide, the human plasma samples (0.3 ml) were precipitated using acetonitrile (0.75 ml) and the centrifuged supernatants were partially evaporated under nitrogen at 37°C at approximately 0.3 ml volume. The LOR, DSL and internal standard were separated on a reversed phase column (Zorbax SB-C18, 100 mm × 3.0 mm i.d., 3.5 μm) under isocratic conditions using a mobile phase of an 8:92 (v/v) mixture of acetonitrile and 0.4% (v/v) formic acid in water. The flow rate was 1 ml/min and the column temperature 45°C. The detection of LOR, DSL and internal standard was in MRM mode using an ion trap mass spectrometer with electrospray positive ionisation. The ion transitions were monitored as follows: 383 → 337 for LOR, 311 → (259 + 294 + 282) for DSL and 300 → 226.8 for internal standard. Calibration curves were generated over the range of 0.52–52.3 ng/ml for both LOR and DSL with values for coefficient of determination greater than 0.994 by using a weighted (1/y) quadratic regression. The lower limits of quantification were established at 0.52 ng/ml LOR and DSL, respectively, with an accuracy and precision less than 20%. Both analytes demonstrated good short-term, long-term, post-preparative and freeze-thaw stability. Besides its simplicity, the sample treatment allows obtaining a very good recovery of both analytes, around 100%. The validated LC/MS/MS method has been applied to a pharmacokinetic study of loratadine tablets on healthy volunteers.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Loratadine; Bioequivalence; Pharmacokinetics; LC–MS/MS

1. Introduction

4-(8-Chloro-5,6-dihydro-11H-benzo[5,6]-cycloheptal[1,2-b]-pyridin-11-yldine)-1-piperidinecarboxylic acid ethyl ester (loratadine) is a long acting tricyclic antihistamine with selective peripheral histamine H1-receptor antagonist activity that is used for relief of symptoms of seasonal allergies and skin rash (Fig. 1). Among the second-generation antihistamines, loratadine is free from sedation at recommended dosages. Following an oral administration of 10 mg tablet, loratadine (LOR) is rapidly absorbed and reaches peak concentration ($T_{\text{max}}$) at 1.3 h. For its major active metabolite, descarboethoxyloratadine, the $T_{\text{max}}$ is 2.5 h [1]. Descarboethoxyloratadine or desloratadine (DSL) would be expected to produce results similar to LOR and other non-sedating antihistamines. The elimination half-life of LOR is 8–14 h, and that of DSL 17–24 h. An oral dose of loratadine (20 mg) leads to maximum plasma concentrations of only 11 and 10 ng/ml for LOR and DSL, respectively [2].

In view of these facts, the analytical method for LOR and DSL determination in human plasma has to be very sensitive and in the case of a large number of samples, the development of a chromatographic method suitable for this kind of analysis must take into account not only a sensitive procedure, but also a fast one, and as much as possible a simple sample preparation. An HPLC–MS or GC–MS method offers the solution from these points of view, in many cases.

Loratadine and desloratadine in plasma were studied by GC [3,4] or HPLC methods with UV [5] or fluorescence detec-
tion [6–9], with a sufficient lower limit of quantification for the purpose of study, by applying liquid–liquid or solid phase extraction. HPLC with MS detection was extensively used in the past years for the sensitive quantification of LOR and DSL [10–18], with a very low limit of quantification obtained mainly after liquid–liquid extraction of analytes.

Taking into account these facts, the aim of the present study was to develop a fast HPLC/MS/MS method able to quantify loratadine and desloratadine in human plasma after oral administration of a therapeutic dose of loratadine after a simple step of extraction. Finally, the developed and validated method was applied for bioequivalence investigation of two medicinal products containing 10 mg loratadine.

2. Experimental

2.1. Reagents

Loratadine and desloratadine were reference standards from Morepen Lab. Limited, India. Metoclopramide hydrochloride (MTC) (Fig. 1) was the internal standard (European Pharmacopoeia standard). Acetonitrile, methanol and formic acid were Merck products (Merck KgAa, Darmstadt, Germany). Distilled, deionised water was produced by a Direct Q-5 Millipore (Molsheim, France) water system. The human blank plasma was supplied by the Local Bleeding Centre Cluj-Napoca, Romania.

2.2. Standard solutions

Two stock solutions of loratadine and desloratadine, respectively, with concentration of 2.5 mg/ml were prepared by dissolving appropriate quantities of reference substances (weighed on an Analytical Plus balance from Ohaus, USA) in 10 ml methanol. Two working solutions were then obtained for each substance by diluting specific volumes of stock solution with plasma. Then these were used to spike different volumes of plasma blank, providing finally eight plasma standards with the concentrations ranged between 0.52 and 52.3 ng/ml, equally for loratadine and desloratadine. Accuracy and precision of the method was verified using plasma standards with concentrations of 0.52, 1.68, 10.47 and 20.94 ng/ml loratadine and desloratadine, respectively. Quality control samples (QC) of 1.68, 10.47 and 20.94 ng/ml analytes were used during clinical samples analysis. The internal standard solution was prepared by sequential dilution of a stock solution of metoclopramide in acetonitrile (1 mg/ml) to reach a concentration of 3.65 ng/ml. This solution was used for precipitation of plasma proteins.

2.3. Chromatographic and mass spectrometry systems and conditions

The HPLC system was an 1100 series model (Agilent Technologies) consisted of a binary pump, an in-line degasser, an autosampler, a column thermostat and an Ion Trap VL mass spectrometer detector (Brucker Daltonics GmbH, Germany). Chromatograms were processed using QuantAnalysis software. The detection of LOR, DSL and internal standard was in MRM mode using an ion trap mass spectrometer with electrospray positive ionisation. The ion transitions were monitored as follows: 383 → 337 for LOR, 311 → (259 + 294 + 282) for DSL and 300 → 226.8 for internal standard. Chromatographic separation was performed at 45 °C on a Zorbax SB-C18 100 mm × 3 mm, 3.5 µm column (Agilent Technologies), protected by an in-line filter.

2.4. Mobile phase

The mobile phase consisted of a mixture of water containing 0.4% formic acid and acetonitrile (92:8, v/v), each component being degassed, before elution, for 10 min in an Elma Transsonic 700/H (Singen, Germany) ultrasonic bath. The pump delivered the mobile phase at 1 ml/min.

2.5. Sample preparation

Standard and test plasma samples were prepared as follows in order to be chromatographically analyzed. In a test tube of 1.5 ml, 0.3 ml plasma and 0.75 ml acetonitrile containing internal standard (3.65 ng/ml metoclopramide as base) were added. The tube was vortex-mixed for 10 s (Vortex Genie 2, Scientific Industries) and then centrifuged for 6 min at 6000 rpm (204 Sigma centrifuge, Osterode am Harz, Germany). The supernatant was transferred in a glass centrifuge tube and evaporated at 37 °C under nitrogen to approximate 0.3 ml. The final solution was transferred to an autosampler vial and 10 µl were injected into the HPLC system.

2.6. Validation

As a first step of method validation [19–21], specificity was verified using six different plasma blanks obtained from...
healthy human volunteers who had not previously taken any medication.

The concentration of analytes was determined automatically by the instrument data system using the internal standard method. Calibration was performed using singlicate calibration standards on five different occasions. The calibration curve model was determined by the least squares analysis. The applied calibration model was \( y = c + bx + ax^2 \), weight 1/(1/\( y \)) quadratic response, where \( y \), area ratio and \( x \), concentration ratio. Distribution of the residuals (%difference of the back-calculated concentration from the nominal concentration) was investigated. The calibration model was accepted, if the residuals were within ±20% at the lower limit of quantification (LLOQ) and within ±15% at all other calibration levels and at least two-thirds of the standards meet this criterion, including highest and lowest calibration levels.

The lower limit of quantification was established as the lowest calibration standard with an accuracy and precision less than 20%.

The within- and between-run precision (expressed as coefficient of variation, CV%) and accuracy (expressed as relative difference between obtained and theoretical concentration; bias%) of the assay procedures were determined by analysis on the same day of five different samples at each of the lower \( 1.68 \text{ ng/ml} \), medium \( 10.47 \text{ ng/ml} \) and higher \( 20.94 \text{ ng/ml} \) levels of the considered concentration range and one different sample of each on five different occasions, respectively. The selected concentration values are relevant in practice taking in account to the fact that the reported maximum concentration levels of either loratadine or its metabolite are not greater that 8–10 ng/ml at usual oral doses (20 mg loratadine) [1,2].

The relative recoveries at each of the previously three levels of concentration and limit of quantification were measured by comparing the response, relative to the internal standard, of the treated plasma standards with the response of standards in solvent with the same concentration of analytes and internal standard as the plasma sample.

The stability of the analytes in human plasma was investigated in three ways, in order to characterize each operation during the process of bioequivalence studies: room-temperature stability (RTS), post-preparative stability (PPS) in the autosampler, freeze-thaw stability (FTS) and long-term stability (LTS), in the first validation day, there were injected and analyzed four samples at each of low and high concentrations, and values were calculated against calibration curve of the day. Other two sets with the same plasma concentrations were stored in freezer below –20 °C and analyzed together with calibration samples after 5 months. The values were calculated against calibration curve of the day and the mean values for the stored samples and nominal concentrations were compared. The requirement for stable analytes was that the difference between mean concentrations of the tested samples in various conditions and nominal concentrations had to be in ±15% range.

2.7. Clinical application and in-study validation

The validated method was applied in a bioequivalence study of two dosage forms—tablets containing 10 mg loratadine. The collecting times were 0, 0.25, 0.5, 1, 1.5, 2, 2.5, 3, 4, 6, 8, 10, 12, 24, 48 h after oral administration of 40 mg loratadine. The accuracy and precision of the validated method was monitored to ensure that it continued to perform satisfactorily during analysis of volunteer samples. To achieve this objective, a number of QC samples prepared in duplicate at three concentration levels were analyzed in each assay run and the results compared with the corresponding calibration curve. At least 67% (four out of six) of the QC samples should be within 15% of their respective nominal values; 33% of the QC samples (not all replicates at the same concentration) can be outside ±15% of the nominal value.

3. Results and discussions

No significant interference at the retention time of LOR (7 min) and DSL (3.2 min) was observed in different plasma blank samples chromatograms (Figs. 2 and 3) due to the specificity of selected signals (Fig. 4). It is well known that an advantage of the ion trap over the triple quadrupole is the sensitivity in scan mode. This allows adding multiple fragments from an MS spectrum in order to improve the overall signal. In the case of desloradine, the sum of ions from MS spectrum (\( m/z \) 259, 294, 282) was chosen for quantification because the detection is about 40% more sensitive that the case based only on ion \( m/z \) 259.
healthy human volunteers who had not previously taken any medication.

The concentration of analytes was determined automatically by the instrument data system using the internal standard method. Calibration was performed using singlicate calibration standards on five different occasions. The calibration curve model was determined by the least squares analysis. The applied calibration model was $y = c + bx + ax^2$, where $y$, area ratio and $x$, concentration ratio. Distribution of the residuals (%difference of the back-calculated concentration from the nominal concentration) was investigated. The calibration model was accepted, if the residuals were within $\pm 20\%$ at the lower limit of quantification (LLOQ) and within $\pm 15\%$ at all other calibration levels and at least two-third of the standards meet this criterion, including highest and lowest calibration levels.

The lower limit of quantification was established as the lowest calibration standard with an accuracy and precision less than 20%.

The within- and between-run precision (expressed as coefficient of variation, CV%) and accuracy (expressed as relative difference between obtained and theoretical concentration, bias%) of the assay procedure were determined by analysis on the same day of five different samples at each of the lower (1.68 ng/ml), medium (10.47 ng/ml) and higher (20.94 ng/ml) levels of the considered concentration range and one different sample of each on five different occasions, respectively. The selected concentration values are relevant in practice taking in account to the fact that the reported maximum concentration levels of either loratadine or its metabolite are not greater that 8–10 ng/ml at usual oral doses (20 mg loratadine) [1,2].

The relative recoveries at each of the previously three levels of concentration and limit of quantification were measured by comparing the response, relative to the internal standard, of the treated plasma standards with the response of standards in solvent with the same concentration of analytes and internal standard as the plasma sample.

The stability of the analytes in human plasma was investigated in three ways, in order to characterize each operation during the process of bioequivalence studies: room-temperature stability (RTS), post-preparative stability (PPS) in the autosampler, freeze-thaw stability (FTS) and long-term stability (LTS) below $-20\,^\circ\mathrm{C}$. For all stability studies plasma standards at low (1.68 ng/ml both LOR and DSL) and high concentrations (20.94 ng/ml both LOR and DSL) were used. Four plasma standards at each of the two levels were prepared and let at room temperature 4 h before processing (RTS study). Other four pairs were prepared, immediately processed and stored in the autosampler (25 °C) of the HPLC system (PPS study). The samples were injected after 10 h, the expected longest storage time of the samples in autosampler before injection. For the freeze-thaw stability aliquots at the same low and high concentrations were prepared. These samples were subjected to three cycles of freeze-thaw operations in 3 consecutive days. After the third cycle, the samples were analyzed against calibration curve of the day. The mean concentration calculated for the samples subjected to the cycles and the nominal ones were compared. For long-term stability, in the first validation day, there were injected and analyzed four samples at each of low and high concentrations, and values were calculated against calibration curve of the day. Other two sets with the same plasma concentrations were stored in freezer below $-20\,^\circ\mathrm{C}$ and analyzed together with calibration samples after 5 months. The values were calculated against calibration curve of the day and the mean values for the stored samples and nominal concentrations were compared. The requirement for stable analytes was that the difference between mean concentrations of the tested samples in various conditions and nominal concentrations had to be in $\pm 15\%$ range.

2.7. Clinical application and in-study validation

The validated method was applied in a bioequivalence study of two dosage forms—tablets containing 10 mg loratadine. The collecting times were 0, 0.25, 0.5, 1, 1.5, 2, 2.5, 3, 4, 6, 8, 10, 12, 24, 48 h after oral administration of 40 mg loratadine. The accuracy and precision of the validated method was monitored to ensure that it continued to perform satisfactorily during analysis of volunteer samples. To achieve this objective, a number of QC samples prepared in duplicate at three concentration levels were analyzed in each assay run and the results compared with the corresponding calibration curve. At least 67% (four out of six) of the QC samples should be within 15% of their respective nominal values; 33% of the QC samples (not all replicates at the same concentration) can be outside $\pm 15\%$ of the nominal value.

3. Results and discussions

No significant interference at the retention time of LOR (7 min) and DSL (3.2 min) was observed in different plasma blank samples chromatograms (Figs. 2 and 3) due to the specificity of selected signals (Fig. 4). It is well known that an advantage of the ion trap over the triple quadrupole is the sensitivity in scan mode. This allows adding multiple fragments from an MS spectrum in order to improve the overall signal. In the case of desloratadine, the sum of ions from MS spectrum (m/z 259, 294, 282) was chosen for quantification because the detection is about 40% more sensitive that the case based only on ion m/z 259.
healthy human volunteers who had not previously taken any medication.

The concentration of analytes was determined automatically by the instrument data system using the internal standard method. Calibration was performed using singlicate calibration standards on five different occasions. The calibration curve model was determined by the least squares analysis. The applied calibration model was \( y = c + bx + ax^2 \), weight \( 1/(1/y) \) quadratic response, where \( y \), area ratio and \( x \), concentration ratio. Distribution of the residuals (%difference of the back-calculated concentration from the nominal concentration) was investigated. The calibration model was accepted, if the residuals were within ±20% at the lower limit of quantification (LLOQ) and within ±15% at all other calibration levels and at least two-third of the standards meet this criterion, including highest and lowest calibration levels.

The lower limit of quantification was established as the lowest calibration standard with an accuracy and precision less than 20%.

The within- and between-run precision (expressed as coefficient of variation, CV%) and accuracy (expressed as relative difference between obtained and theoretical concentration, bias%) of the assay procedure were determined by analysis on the same day of five different samples at each of the lower (1.68 ng/ml), medium (10.47 ng/ml) and higher (20.94 ng/ml) levels of the considered concentration range and one different sample of each on five different occasions, respectively. The selected concentration values are relevant in practice taking into account to the fact that the reported maximum concentration levels of either loratadine or its metabolite are not greater than 8–10 ng/ml at usual oral doses (20 mg loratadine) [1,2].

The relative recoveries at each of the previously three levels of concentration and limit of quantification were measured by comparing the response, relative to the internal standard, of the treated plasma standards with the response of standards in solvent with the same concentration of analytes and internal standard as the plasma sample.

The stability of the analytes in human plasma was investigated in three ways, in order to characterize each operation during the process of bioequivalence studies: room-temperature stability (RTS), post-preparative stability (PPS) in the autosampler, freeze-thaw stability (FTS) and long-term stability (LTS), in the first validation day, there were injected and analyzed four samples at each of low and high concentrations, and values were calculated against calibration curve of the day. Other two sets with the same plasma concentrations were stored in freezer below −20 °C and analyzed together with calibration samples after 5 months. The values were calculated against calibration curve of the day and the mean values for the stored samples and nominal concentrations were compared. The requirement for stable analytes was that the difference between mean concentrations of the tested samples in various conditions and nominal concentrations had to be in ±15% range.

2.7. Clinical application and in-study validation

The validated method was applied in a bioequivalence study of two dosage forms—tablets containing 10 mg loratadine. The collecting times were 0, 0.25, 0.5, 1, 1.5, 2, 2.5, 3, 4, 6, 8, 10, 12, 24, 48 h after oral administration of 40 mg loratadine. The accuracy and precision of the validated method was monitored to ensure that it continued to perform satisfactorily during analysis of volunteer samples. To achieve this objective, a number of QC samples prepared in duplicate at three concentration levels were analyzed in each assay run and the results compared with the corresponding calibration curve. At least 67% (four out of six) of the QC samples should be within 15% of their respective nominal values; 33% of the QC samples (not all replicates at the same concentration) can be outside ±15% of the nominal value.

3. Results and discussions

No significant interference at the retention time of LOR (7 min) and DSL (3.2 min) was observed in different plasma blank samples chromatograms (Figs. 2 and 3) due to the specificity of selected signals (Fig. 4). It is well known that an advantage of the ion trap over the triple quadrupole is the sensitivity in scan mode. This allows adding multiple fragments from an MS spectrum in order to improve the overall signal. In the case of desloratadine, the sum of ions from MS spectrum (m/z 259, 294, 282) was chosen for quantification because the detection is about 40% more sensitive that the case based on ion m/z 259.
healthy human volunteers who had not previously taken any medication.

The concentration of analytes was determined automatically by the instrument data system using the internal standard method. Calibration was performed using singlicate calibration standards on five different occasions. The calibration curve model was determined by the least squares analysis. The applied calibration model was $y = c + bx + ax^2$, weight 1/y (1/y) quadratic response, where $y$, area ratio and $x$, concentration ratio. Distribution of the residuals (%difference of the back-calculated concentration from the nominal concentration) was investigated. The calibration model was accepted, if the residuals were within ±20% at the lower limit of quantification (LLOQ) and within ±15% at all other calibration levels and at least two-third of the standards meet this criterion, including highest and lowest calibration levels.

The lower limit of quantification was established as the lowest calibration standard with an accuracy and precision less than 20%.

The within- and between-run precision (expressed as coefficient of variation, CV%) and accuracy (expressed as relative difference between obtained and theoretical concentration, bias%) of the assay procedure were determined by analysis on the same day of five different samples at each of the lower (1.68 ng/ml), medium (10.47 ng/ml) and higher (20.94 ng/ml) concentrations of the considered concentration range and one different sample of each on five different occasions, respectively. (The selected concentration values are relevant in practice taking in account to the fact that the reported maximum concentration levels of either loratadine or its metabolite are not greater that 8–10 ng/ml at usual oral doses (20 mg loratadine) [1,2].

The relative recoveries at each of the previously three levels of concentration and limit of quantification were measured by comparing the response, relative to the internal standard, of the treated plasma standards with the response of standards in solvent with the same concentration of analytes and internal standard as the plasma sample.

The stability of the analytes in human plasma was investigated in three ways, in order to characterize each operation during the process of bioequivalence studies: room-temperature stability (RTS), post-preparative stability (PPS) in the autosampler, freeze-thaw stability (FTS) and long-term stability (LTS), in the first validation day, there were injected and analyzed four samples at each of low and high concentrations, and values were calculated against calibration curve of the day. Other two sets with the same plasma concentrations were stored in freezer below −20 °C and analyzed together with calibration samples after 5 months. The values were calculated against calibration curve of the day and the mean values for the stored samples and nominal concentrations were compared. The requirement for stable analytes was that the difference between mean concentrations of the tested samples in various conditions and nominal concentrations had to be in ±15% range.

### 2.7. Clinical application and in-study validation

The validated method was applied in a bioequivalence study of two dosage forms—tablets containing 10 mg loratadine. The collecting times were 0, 0.25, 0.5, 1, 1.5, 2, 2.5, 3, 4, 6, 8, 10, 12, 24, 48 h after oral administration of 40 mg loratadine. The accuracy and precision of the validated method was monitored to ensure that it continued to perform satisfactorily during analysis of volunteer samples. To achieve this objective, a number of QC samples prepared in duplicate at three concentration levels were analyzed in each assay run and the results compared with the corresponding calibration curve. At least 67% (four out of six) of the QC samples should be within 15% of their respective nominal values; 33% of the QC samples (not all replicates at the same concentration) can be outside ±15% of the nominal value.

### 3. Results and discussions

No significant interference at the retention time of LOR (7 min) and DSL (3.2 min) was observed in different plasma blank samples chromatograms (Figs. 2 and 3) due to the specificity of selected signals (Fig. 4). It is well known that an advantage of the ion trap over the triple quadrupole is the sensitivity in scan mode. This allows adding multiple fragments from an MS spectrum in order to improve the overall signal. In the case of desloratadine, the sum of ions from MS spectrum (m/z 259, 294, 282) was chosen for quantification because the detection is about 40% more sensitive that the case based only on ion m/z 259.

![Fig. 2. Chromatograms of a plasma blank containing metoclopramide (MTC, internal standard).](image-url)
New validated method for piracetam HPLC determination in human plasma

Augustin Curticapean *, Silvia Imre

University of Medicine and Pharmacy Tg-Mures, Faculty of Pharmacy, Gh. Marinescu Str. No.38,
Tg-Mures RO-540139, Romania

Received 19 January 2006; received in revised form 3 May 2006; accepted 6 June 2006

Abstract

The new method for HPLC determination of piracetam in human plasma was developed and validated by a new approach. The simple determination by UV detection was performed on supernatant, obtained from plasma, after proteins precipitation with perchloric acid. The chromatographic separation of piracetam under a gradient elution was achieved at room temperature with a RP-18 LiChroSpher 100 column and aqueous mobile phase containing acetonitrile and methanol. The quantitative determination of piracetam was performed at 200 nm with a lower limit of quantification LLQ=2 μg/ml. For this limit, the calculated values of the coefficient of variation and difference between mean and the nominal concentration are CV% = 9.7 and bias% = 0.9 for the intra-day assay, and CV% = 19.1 and bias% = −7.45 for the between-days assay. For precision, the range was CV% = 1.8 ÷ 11.6 in the intra-day and between-days assay, and for accuracy, the range was bias% = 2.3 ÷ 14.9 in the intra-day and between-days assay. In addition, the stability of piracetam in different conditions was verified. Piracetam proved to be stable in plasma during 4 weeks at −20 °C and for 36 h at 20 °C in the supernatant after protein precipitation. The new proposed method was used for a bioequivalence study of two medicines containing 800 mg piracetam.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Piracetam; HPLC determination; Method validation

Abbreviations: SC, Stock Solution; WC, Working Solution; STD, Standard Plasma Solutions; QC, Quality Control; LLQ, Lower Limit of Quantification; PPS, Post-Preparative Stability; FTS, Freeze-Thaw Stability; LTS, Long-Term Stability; SD, Standard Deviation; CV%, Coefficient of Variation; bias%, ratio of the difference between mean found and nominal concentration and nominal concentration.

* Corresponding author. Tel.: +40 265 215551x171; fax: +40 265 210407.

E-mail address: augustin@umftgm.ro (A. Curticapean).
2.3. Method validation

Specificity of the method was investigated by analyzing blank plasma samples of six individual volunteers, in order to determine any interference of endogenous compounds with the analyte. Investigation of the selectivity was carried out together with the calibration curve assay.

The six calibration standard replications were prepared and measured in a single run in the first validation day. Then, during the next four consecutive days there were measured standards for only one calibration curve. Percentage residuals (% difference of the back-calculated concentration from the nominal concentration) were calculated for each calibration standard. The applied requirements for a valid calibration model were: a regression coefficient higher than 0.99 ($R^2$), and the residuals and coefficient of variation (CV%) to be within ±20% at the lower limit of quantification (LLQ) and ±15% for the rest of the concentrations tested. For the first validation day there were analyzed in a single validation batch five replicates of QC samples, at each of the concentration levels (5, 20 and 80 μg/ml), and reported to the calibration curve of the batch (intra-day assay). On each of the four remaining consecutive days, a single QC sample from each of the concentration levels (5, 20 and 80 μg/ml) was analyzed against daily calibration (between-days assay). The requirement for precision was acceptable if the coefficient of variation (CV%) of the determined QC concentrations did not exceed ±15%, either for intra-day or between-days assay. The accuracy was acceptable if the ratio of the difference between the mean measured concentration and the nominal values against nominal concentration (bias%) did not exceed ±15%, either for intra-day or between-days assay.

The lower limit of quantification LLQ was analyzed against calibration curve (intra-day assay), by determining in the first day, five couples of STD replicates with the concentrations of 1 μg/ml and 2 μg/ml, respectively. Then, on each of the four consecutive days a single couple of STD samples were analyzed, against daily calibration curve (between-days assay). The requirement for precision was acceptable if the coefficient of variation (CV%) of the measured STD samples (at lowest concentration of the calibration curve) did not exceed ±20%, either for intra-day or between-days assay. The accuracy was acceptable if the bias% did not exceed ±20%, either for intra-day or between-days assay.

The stability of the piracetam in human plasma was investigated by three ways, in order to characterize each operation during the process of bioequivalence studies (post-preparative stability in the autosampler, freeze-thaw stability and long-term plasma stability below −20 °C). For post-preparative stability (PPS), prepared samples at low (10 μg/ml) and high concentrations (100 μg/ml) were investigated. All aliquots were stored at 20 °C in the autosampler of the HPLC system and were injected in time, after the following timetable: 2, 5, 7, 10, 24, 36 and 48 h. For the freeze-thaw stability (FTS), aliquots at the same low (10 μg/ml) and high (100 μg/ml) concentrations were prepared. These samples were subjected to 3 cycles of freeze-thaw operations in three consecutive days. After the third cycle the samples were analyzed against calibration curve of the day. The mean concentration calculated for the samples subjected to the cycles and the nominal ones were compared. For long-term stability (LTS), in the first validation day, there were injected and analyzed four samples at each of low (10 μg/ml) and high (100 μg/ml) concentrations, and values were calculated against calibration curve of the day. Other three sets with the same plasma concentrations were stored in freezer below −20 °C and analyzed together with calibration samples, after two, three and four weeks, each time being used four pairs of replicates for low and high concentrations. The values were calculated against calibration curve of the day and the mean values for the stored samples and nominal concentrations were compared.
The requirement for a stable analyte in the frozen matrix is that the difference between mean concentration of the stored samples and nominal concentration is between ±15%.

The recovery was determined at four levels of concentration as follows: one for the LLQ concentration, and the rest for the QC concentrations. Five replications of each level of concentration were prepared and analyzed. The efficiency was calculated by height ratio of the mean values of the samples against height of the reference samples with the same concentrations prepared in supernatant.

3. Results

The developed method showed that piracetam is well separated by the joint action of the mobile phase and the solid one.

For specificity determination, the requirements are fulfilled, as one can observe in the chromatograms recorded for six different volunteers (Fig. 2a and b). No endogenous peaks interfered at the retention time of piracetam.

The chromatograms of the high QC (80 μg/ml) sample and overlay of the standard plasma samples, spiked with piracetam, are shown in Fig. 3a and b, respectively.

The retention time at RT = 5.1(±1%) min, having an unretained solute time of \( t_0 = 1.6 \) min, offers a good separation factor \( k' = 2.19(±1\%) \).

The regression coefficient was higher than 0.99 (\( R^2 \)) with the values between 0.9986÷ 0.9999 and the residuals and coefficient of variation (CV%) are within ±20% at the LLQ and within ±15% for the rest of the tested concentrations.

The method proved to be precise, with CV% = 3.2÷11.6 for the intra-day assay and CV% = 1.8÷11.6 for the between-days assay. The accuracy (bias%) ranged between 8.4÷14.9 for intra-day assay and between 2.3÷3.9 for the between-days assay (Table 2).

<table>
<thead>
<tr>
<th>Stability</th>
<th>Post-preparative</th>
<th>Freeze-thaw</th>
<th>Long-term (4 weeks)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample code</td>
<td>PPS</td>
<td>FTS</td>
<td>LTS</td>
</tr>
<tr>
<td>( c_{nominal} ) (μg/ml)</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>( c_{mean} ) (μg/ml)</td>
<td>9.62</td>
<td>10.52</td>
<td>10.03</td>
</tr>
<tr>
<td>bias%</td>
<td>-3.8</td>
<td>5.2</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Table 3

Stability results

<table>
<thead>
<tr>
<th>Stability</th>
<th>Post-preparative</th>
<th>Freeze-thaw</th>
<th>Long-term (4 weeks)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample code</td>
<td>PPS</td>
<td>FTS</td>
<td>LTS</td>
</tr>
<tr>
<td>( c_{nominal} ) (μg/ml)</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>( c_{mean} ) (μg/ml)</td>
<td>9.62</td>
<td>10.52</td>
<td>10.03</td>
</tr>
<tr>
<td>bias%</td>
<td>-3.8</td>
<td>5.2</td>
<td>0.3</td>
</tr>
</tbody>
</table>