ELSEVIER

Journal of Pharmaceutical and Biomedical Analysis
24 (2001) 755–767

JOURNAL OF
PHARMACEUTICAL
AND BIOMEDICAL
ANALYSIS

www.elsevier.com/locate/jpba

Validation in pharmaceutical analysis. Part I: An integrated approach

Joachim Ermer *

Aventis Pharma AG, Global Analytical Development, QOTSS, Poseidonhaus, D-65926 Frankfurt am Main, Frankfurt, Germany

Received 16 May 2000; received in revised form 16 November 2000; accepted 17 November 2000

Abstract

The ICH guidelines achieved a great deal in harmonising the definitions of the required validation characteristics and their basic requirements. However, they provide only a basis for a general discussion of the validation parameters, their calculation and interpretation. It is the responsibility of the analyst to identify parameters which are relevant to the performance of the given analytical procedure as well as to design proper validation protocols including acceptance criteria and to perform an appropriate evaluation. In order to fulfil this resposibility properly, the background of the validation parameters and their consequences must be understood. In this part, the general concept of an integrated validation is discussed. The interdependencies to other ICH guidelines and topics during drug development (e.g. impurities and degradants, stability and specification design) must be taken into account to define the required acceptance criteria. Evaluation of the results in order to prove the suitability of the analytical procedure must be based on the specification limits. Important parameters and aspects are discussed for the individual validation characteristics. In the following parts, these parameters will be discussed in detail. Examples will be given for their interpretation in order to facilitate the selection of parameters which are relevant to the performance and suitability of the given analytical procedure.

Keywords: Pharmaceutical analysis; Validation; Quality assurance; Acceptance limits; Specification limits

1. Introduction

The validation of analytical procedures, i.e. the proof of its suitability for the intended purpose, is an important part of the registration application for a new drug [1–8]. The International Conference on the Harmonization of the Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) has harmonised the requirements in two guidelines [7,8]. The first one summarizes and defines the validation characteristics needed for various types of test procedures, the second one extends the previous text to include the experimental data required and some statistical interpretation. These guidelines serve as a basis worldwide both for regulatory authorities and industry and bring the importance of a

* Tel.: +49-69-30584890; fax: +49-69-30525538.
E-mail address: joachim.ermer@aventis.com (J. Ermer).

PII: S0731-7085(00)00530-6

proper validation to the attention of all those involved in the process of submission.

Nowadays, the validation characteristics needed for the various test procedures and their general requirements (see Table 1) are well understood. However, in spite of recommendations not to do so [8], there is an increasing tendency to misuse the guidelines as a kind of checklist which is automatically applied. Such a 'checklist mentality' may be caused by an incorrect understanding of standardisation and improving efficiency. Due to the integration of analytics in all aspects of drug development and quality control, inappropriate analytical procedures may, however, lead to, for example, wrong decisions, work which has to be repeated (out-of specification results) and delays. Consequently, both the design of the validation studies and the evaluation of the results must be adjusted to the individual analytical procedure in order to achieve an understanding of its real performance. Only on this basis can proof be obtained that the procedure 'is suitable for its intended purpose' [14].

2. Integration and interdependencies of analytical validation

Analytical procedures are used throughout drug development and the manufacturing of drug substances and drug products. Important decisions such as the establishment of the shelf-life from stability studies, the need for additional toxicological trials if new impurities appear or if known impurities exceed the qualified levels, the reworking of batches and batch release or rejection are based on analytical results. In order to make the right decisions and to avoid additional work, an appropriate performance of the analytical procedures is essential.

But what does 'suitability for its intended purpose' mean? For some applications, the requirements are defined in the ICH guidelines, for example, for impurity testing [9–11]. If the reporting level for unknown impurities in drug substances is set to 0.05 or 0.03% [9], the corresponding test procedure must be able to quantify impurities at this concentration with an appropriate level of precision and accuracy.

Table 1
Validation characteristics normally evaluated for the different types of test procedures [7] and the minimum number of determinations required (if applicable) [8]

Validation characteristics	Minimum number	Test procedure			
		Identity	Impurities		Assay[a]
			Quantitative	Limit	
Specificity[b]	–	Yes	Yes	Yes	Yes
Linearity	5 concentrations	No	Yes	No	Yes
Range	–	No	Yes	No	Yes
Accuracy	9 determinations over 3 concentration levels (e.g. 3×3)	No	Yes	No	Yes
Precision					
Repeatability	6 determinations at 100% or 9 determinations over 3 concentration levels (e.g. 3×3)	No	Yes	No	Yes
Intermediate precision/ reproducibility[c]	2 series	No	Yes	No	Yes
Detection limit	–	No	No[d]	Yes	No
Quantitation limit	–	No	Yes	No	No

[a] Including dissolution, content/potency.
[b] Lack of specificity of one analytical procedure could be compensated by other supporting analytical procedure(s).
[c] Intermediate precision sufficient for submission.
[d] May be needed in some cases.

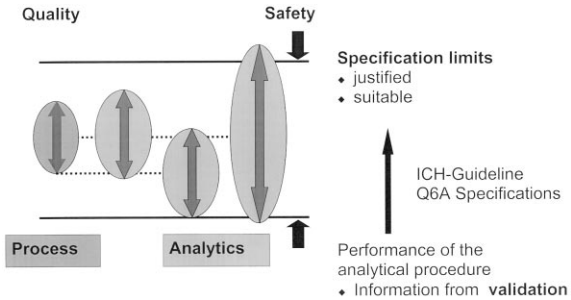

Fig. 1. Suitability of the analytical procedure with respect to specification limits.

With respect to assay determinations, the variability of the analytical procedure is often larger than the variability of the manufacturing. This must be taken into account in the establishment of specification limits [12,13], i.e. the variability of the procedure and the acceptance limits must be compatible (Fig. 1). Of course, safety requirements are of primary importance but if they are satisfied, specification limits can also be defined based on the analytical variability [15,16].

The analytical state of the art should be taken into account although it is not the ultimate goal to optimise an individual analytical procedure as well as possible. It is also very important to recognise that the release of a given batch is based on a whole set of test procedures which complement and supplement each other. Their selection in the specification design [12,13] has, therefore, considerable influence on the required performance of the individual control test and, hence, on its validation.

Besides this 'horizontal' integration, there are also 'vertical' connections. There is a strong feedback between method development and validation [17]. Due to the important aspect of time, it is advisable (if possible) to perform a 'progressive' validation starting from a basic data set which is supplemented, for instance, with respect to (intermediate) precision and robustness. The validation results can also have a feedback effect on details of the analytical procedure such as on the number of replicate determinations or the mode of calibration (see next chapter).

During method development and validation the critical performance parameters of the analytical procedure should be identified in order to design discriminating system suitability tests.

If these interdependencies are ignored and if parameters are determined during the validation which do not describe the critical performance of the analytical procedure, severe consequences can be expected. For example, if the procedure is not sufficiently robust, problems are likely to occur in a method transfer and repeated adjustments in the system suitability test will, for example, be required. If the acceptance limits for, for example, assay are too narrow, out-of specification results will require extensive investigations [18]. Consequently, the time and effort (perhaps) saved in a 'checklist' validation approach will most likely result in problems at a later date which could be much more expensive.

3. Acceptance criteria for validation parameters

Using the ICH guidelines as a basis [7,8], it is the responsibility of the analyst to select for the given individual test procedure relevant parameters and appropriate acceptance criteria and to design the experimental studies accordingly. These acceptance criteria can often be derived from specification limits.

As a general rule, the standard deviation of the analytical procedure should be lower than 1/6 of the specification range. A detailed approach taking the number of repeated determinations into account is based on confidence intervals [15,16] (Fig. 2). Basic specification limits (BL) include the variability of the manufacturing process and represent the final limits (SL) if an error-free analytical procedure is used. Describing the analytical variability as a normal distribution of the experimental results, confidence intervals can be constructed as a representation of the result probability for a given number of replicates. The combination of the basic limit and the limit (upper or lower) of the 95% confidence interval (one-sided) then represents the overall specification limit (Eq. (1)). It must be taken into account that this calculation is based on the true standard deviation whereas the standard deviation obtained in a validation study is only a random estimate

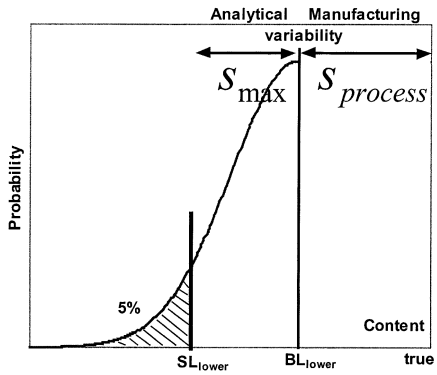

Fig. 2. Construction of specification limits from 95% confidence intervals. The procedure is shown for the lower limits. BL, basic limits, imposed by the variability of the manufacturing process; SL, overall specification limit, combination of BL and the (lower) limit of the one-sided 95% confidence interval of the analytical variability.

which also displays a variability. A reliable experimental determination of the true value requires repeated intermediate precision studies, but it can be estimated from the statistical (χ^2) distribution of standard deviations. The upper 95% confidence limit of this distribution (UL) will represent the maximum value for the true standard deviation (Eq. (2)). For six experimental values, it can be approximated by twice the experimental standard deviation [16].

$$\mathrm{SL} = \mathrm{BL} \pm t_{n-1,0.95} \times \frac{s_{\mathrm{true}}}{\sqrt{n}} \quad (1)$$

$$\mathrm{UL}(s_{\mathrm{true}}) = s_{\mathrm{exp}} \times \sqrt{\frac{(n-1)}{\chi^2_{n-1,0.95}}} \quad (2)$$

Rearranging Eqs. (1) and (2) gives the maximum permitted experimental standard deviation for the given specification limits, i.e. the acceptance limit for the validation. For example, a drug substance LC-assay (performed with four repetitions) with a lower specification limit of 98.0% and a limit for the sum of impurities of 0.5% (i.e. a lower basic limit of 99.5%) would require an experimental standard deviation in validation below 0.64%. It should be noted that BL and SL in Eq. (3) refer to the 'critical' half of the specification range. For drug substances, due to the presence of impurities, these are the lower limits. As the required standard deviation is dependent on the number of repetitions in the assay, adjustments are possible (Table 2). Thus, the number of repeated determinations can also be fine-tuned according to the results of the validation. In case of sufficiently wide specification limits compared to the analytical variability, this allows an efficiency optimisation of the analytical procedure.

$$s_{\mathrm{max}} = \frac{|(\mathrm{BL} - \mathrm{SL})| \times \sqrt{n}}{2 \times t_{n-1,0.95}} \quad (3)$$

While in the case of a drug substance assay the basic limits are defined mainly by the sum of impurities, in the case of drug products, often only an estimation is possible. For simple dosage forms, the same variability may be expected for analytics and manufacturing.

It should be noted that the estimation of the true standard deviation will result in a maximum value and, therefore, reduce the required acceptance limit. Consequently, in critical cases an ex-

Table 2
Maximum permitted standard deviation for assay determinations in dependence on the number of repetitions and the specification range

	Drug product (%)	Drug product (%)	Drug substance (%)
Specification range	95–105	95–105	98–102
Basic limit (lower)	97.5 (estimated)	99.0 (estimated)	99.5 (sum of impurities)
Number of repetitions	Acceptance limit for experimental standard deviation in validation ($n = 6$)		
2	0.28	0.45	0.17
3	0.74	1.19	0.45
4	1.06	1.7	0.64
6	1.44	2.3	0.86

perimental determination (using at least four to six series) might be reasonable to obtain an estimate for the true standard deviation. Then, Eq. (1) can be directly rearranged and the factor of two in Eq. (3) can be omitted.

If specification limits are not yet defined or implied due to safety requirements, Eq. (1) can be used directly to calculate the limits. For the true standard deviation, the upper limit of the experimental determination (Eq. (2)) can be calculated or twice this value as an approximation. Alternatively, a soundly based intermediate precision may be used as an estimation for the true standard deviation.

The validation acceptance limit should then be derived from previous experiences of comparable analytical procedures (analytical state of the art). For example, from a well based as well as acceptable intermediate precision of 1.0% and a basic limit of 99.0%, the specification range for a drug substance to be determined with four repetitions can be calculated to 97.8–101.2%. If the experimental repeatability was determined with 1.0%, limits from 96.4–102.4% would result.

With respect to impurity determinations, the ICH reporting threshold of 0.05% [9] can be regarded as the required quantitation limit for unknown impurities which would guarantee a reliable quantitation at the specification limit of 0.1%. If other limits are required due, for example, to safety considerations, the above-mentioned approach can be applied.

Statistical tests such as the Student's t-test or the evaluation of 95% confidence intervals should only be carefully (directly) applied as acceptance criteria because they test for statistical differences. Due to sometimes abnormally small variabilities in the analytical series, differences are identified as significant which are of no practical relevance [19]. In addition, when comparing independent methods for the proof of accuracy, different specificities can be expected which add a systematic bias, thus increasing the risk of the aforementioned danger.

The analyst must decide, if detected statistical differences are of practical relevance. For example, if (e.g. due to abnormal small variability in one series) for a mean t-test or in linearity a (statistical) significant difference to another series or of the intercept to zero is detected, the absolute magnitude of this difference should be taken into account. On the other hand, a large variability can also obscure differences which are not acceptable.

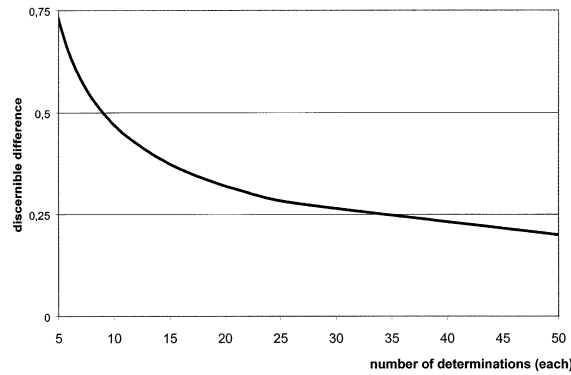

Fig. 3. Dependence of the test power on the number of determinations. Mean t-test at 95% level of statistical significance, assuming a true standard deviation of 0.5 and the same number of determinations for each mean.

In addition, the power of these tests increases with the number of determinations. As shown in Fig. 3, with six determinations, a difference of 0.64 between two means can be detected, whereas a difference of 0.25 can be distinguished with 32 repetitions. Although the latter difference is statistically significant, in most cases (such as LC-assay determination) it has no practical relevance.

For practical purposes, a sufficient agreement between two results (two means or a mean and a nominal value, e.g. recovery) is completely adequate. The acceptance criteria can be derived from previous experience or calculated based on specification limits and statistical considerations (see next part). For example, assuming a LC-assay with specification limits from 95 to 105%, a recovery range from 98 to 102% would be acceptable.

The acceptance criteria and limits should be defined before starting the validation and included in a protocol. After the validation studies, they will serve as a basis for the evaluation.

4. Validation characteristics

When performing validation studies, the whole analytical procedure including all the steps of the sample preparation should be applied, as far as possible. In contrast, the term 'method' should be restricted to the mode of analytical determination alone (e.g. capillary electrophoresis, (reversed-phase) chromatography and spectrometry).

Allowed exceptions to the written procedure concern the number of repetitions as the number of determinations for the various validation characteristics is described in the ICH guideline (Table 2) and the repetitions in the procedure may be fine-tuned based on the validation results (see previous chapter). Such an adjustment may also be used for the final calibration mode (Section 4.2).

As far as possible, the analytical procedure should be independent of the actual equipment used provided that the equipment has been appropriately qualified. This must be taken into account for the validation studies.

4.1. Specificity

There has been some controversy regarding the technical term for this validation characteristic, i.e. specificity vs. selectivity [20]. In contrast to an isolated test procedure, in pharmaceutical analysis the sum of various control tests and hence their combined specificity is used for the overall batch evaluation. A very pragmatic definition describes selectivity as the (physical) separation of substance mixtures with, for example, chromatography and electrophoresis, i.e. the determination of the analyte in addition to other substances. The individual determination of an analyte in the presence of other substances, (i.e. without significant influence of other substances or classes of substances) is defined as specific by, for instance, mass spectrometry, NMR, infrared, fluorescence or UV spectrometry, electrochemical detection and titration [21].

In spite of this discussion, there is a broad agreement that this validation characteristic is of crucial importance and is the critical basis for each analytical procedure. As no absolute and quantitative measure exists (at least for the overall specificity), the requirements depend on the individual analytical procedure as well as on its combination with others. For example, the overall specificity of a quality control can be obtained by securing (or correcting) a precise and efficient assay titration with a selective chromatographic impurity determination. Assuming a titration for the assay of a basic drug substance with a molecular weight of 300. Three impurities are specified, two of which are also basic compounds (A: 200 and B: 100 MW). The third impurity, a neutral molecule will have no titration response, for the other impurities titration response factors of $300/200 = 1.5$ and $300/100 = 3.0$ can be calculated. Therefore, the result of the batch titration must be corrected with the amounts of the basic impurities obtained by a selective LC-procedure. Of course, the amount of unknown impurities must be limited in order to prevent non predictable titration responses.

$$\text{drug}[\%\ \text{as is}] = \text{titration}[\%] \times \frac{100 - 1.5 \times A[\%,\ \text{LC}] - 3 \times B[\%,\ \text{LC}]}{100} \quad (4)$$

With respect to chromatographic techniques, specificity can be demonstrated by a sufficient separation of the substances present. For the assay, appropriate separation means an adequate resolution between the main peak and the impurity and placebo peaks which need not to be separated from each other. The same can be applied to individual impurity or degradant determination. In contrast, universal procedures for the determination of impurities require a sufficient separation of all relevant impurity peaks. The required resolution is strongly dependent on the difference in the size of the corresponding peaks as well as on their elution order [22]. In order to be able to detect the coelution of unknown substances, peak homogeneity investigations such as rechromatography, diode array detection or LC-MS coupling should be performed [next part].

If samples from stress testing are used to demonstrate appropriate separation power, care

should be taken to avoid overdegradation as this would result in secondary (or even higher order) degradants which are of no practical relevance. Therefore, degradation should be restricted to about 10%. Alternatively, samples from regular stability studies (accelerated storage conditions) may be used.

4.2. Linearity/range

A linear dependence of the signal and the analyte concentration is certainly the most convenient case and widely used in pharmaceutical analysis. However, there are analytical procedures with a nonlinear response such as TLC, fluorescence detection and atom absorption spectrometry. Therefore, the term 'analytical response' would have been more appropriate for this validation characteristic.

The essential question to be answered here is on the suitability of the calibration mode to be used in the test procedure. The requirements and relevant parameters for the various calibrations are given in Table 3. A detailed discussion will follow in the next part.

It should be noted that in most cases only a qualitative statement is needed. For example, if a single-point calibration (external standard) is

Table 3
Requirements for different calibration modes with relevant parameters

Quantitation	Requirements	Relevant parameters
Single-point calibration		
External standard	Linear function	Standard error of slope (residual standard deviation), sensitivities (relative standard deviation, graph), residual analysis, statistical tests (vs. quadratic regression)
	Non-significant ordinate intercept	Inclusion of zero in confidence interval of the ordinate intercept, magnitude of the intercept (as percent of the signal at 100% test concentration)
	Homogeneity of variances[a]	*F*-test of the variances at the lower and upper limit of the range
Multiple-point calibration		
Linear, unweighted	Linear function	Standard error of slope (residual standard deviation), sensitivities (relative standard deviation, graph), residual analysis, statistical tests (vs. quadratic regression)
	Homogeneity of variances[a]	*F*-test of the variances at the lower and upper limit of the range
Linear, weighted	Linear function	Standard error of slope (residual standard deviation), sensitivities (relative standard deviation, graph), residual analysis, statistical tests (vs. quadratic regression)
Non-linear	Continuous function	Appropriate equation
100%-method (area normalisation for impurities):	For main peak: linear function	Standard error of slope (residual standard deviation), sensitivities (relative standard deviation, graph), residual analysis, statistical tests (vs. quadratic regression)
	Non-significant ordinate intercept	Inclusion of zero in confidence interval of the ordinate intercept, magnitude of the intercept (as percent of the signal at 100% test concentration)
	Homogeneity of variances[a]	*F*-test of the variances at the lower and upper limit of the range
	For impurities: linear function	Standard error of slope (residual standard deviation), sensitivities (relative standard deviation, graph), residual analysis, statistical tests (vs. quadratic regression)

[a] May be presumed for a limited range (factor 10–20).

Table 4
Quantitative approaches to demonstrate accuracy according to ICH [8]

Drug substance	Application of the analytical procedure to a reference material Comparison of the results with those of a second, well characterised procedure
Drug product	Application of the analytical procedure to synthetic mixtures of drug product components Spiking of analyt to drug product Comparison of the results with those of a second, well characterised procedure
Impurities (quantitative)	Spiking of the impurity to drug substance or product Comparison of the results with those of a second, well characterised procedure

aimed at, the requirements are a linear response function and the zero intercept. If these prerequisites are fulfilled, the actual figures for the standard error of slope or the confidence interval of the intercept, for example, are not used further or referred to. Therefore, it makes no sense to repeat linearity investigations on other days or with other operators [23]. However, care should be taken to remain within the linear range of the individual detector used but this information can be obtained from the equipment qualification.

4.3. Accuracy

The ICH guideline recommends the demonstration of accuracy over the whole working range (see Table 1). However, if only a narrow range is required (e.g. assay or impurities with a low specification limit), a six-fold determination at a 100% test concentration as described for the precision studies may also be used.

Several approaches discussed in the ICH guideline are given in Table 4.

If the analytical test to be validated is compared with another procedure or applied to a reference substance, the probably different specificities must be taken into account. Therefore, statistical tests should be performed only if the systematic bias based on these differences can be quantified and thus corrected or are negligible. Otherwise, the comparison should be performed as a qualitative verification of plausibility or an acceptable maximum difference should be defined (e.g. 2.0% for an LC assay).

Spiking experiments for recovery investigations should be performed as closely to the authentic conditions as possible so that possible interferences between the analyte and matrix can be recognised. This ranges, for example, from the direct preparation of a drug product with various contents of active ingredient to which the whole analytical procedure is applied to the addition of a drug substance stock solution to a placebo solution.

For the quantitation of the analyte, the same calibration mode as described in the final test procedure must be used. Again, statistical tests should be used carefully, especially with complex matrices and low concentrations of impurities. Alternatively, acceptable deviations from the theoretical recovery of 100% can be defined based on the application, experiences or general statistical considerations (next part).

Using UV detection, response factors for known and available impurities and degradants can be obtained from linearity or recovery studies (ratio of the slopes or of specific peak area of analyte and active ingredient). However, this is not possible for unknown impurities or impurities which are not available. In these cases, safety checks should be performed to identify possible problems. As a first step, the peak area percent can be obtained additionally at a ‘check’ wavelength in the lower UV range (e.g. 210 nm) [24] where absorbance coefficients often differ less. Comparable results for the two wavelengths indicate similar response factors, wheras large differences in the peak area may indicate response factors different to unity (Fig. 4). However, the latter belongs rather to the analytical development and the design of the analytical procedure.

4.4. Precision

In addition to the ICH precision levels (Table 1), it is advisable to determine the system precision (injection repeatability) either by repeated

determinations of the same test solution or from double determinations of each test solution used for repeatability (Eq. (5)).

$$s_d = \sqrt{\frac{1}{2m}\sum_{i=1}^{m}(x_{i,1} - x_{i,2})^2} \quad (5)$$

Repeatability, also termed intra-assay precision, refers to the precision obtained under the same operating conditions over a short interval of time by applying the whole analytical procedure to the sample. Intermediate precision refers to within-laboratory variations. The extent of investigations will depend on the intended use of the analytical procedure. A typical investigation might include analysts, days, equipment, reagents, columns, etc. Preferably, the intermediate precision studies should be extended over a longer period of time, in order to obtain a measure of the analytical variability which is representative for the long-term routine use. A basic validation study can also be supplemented by incorporating investigations into the routine application of the analytical procedure.

Repeatability and intermediate precision can be calculated by an analysis of variances [25,26]. The former figure represents the overall variability within the performed series, the latter also includes the variability between the series (Table 5). The difference between the precision levels as well as their absolute magnitude indicate the robustness of the analytical procedure. For the evaluation of the suitability (compatibility with specification limits, see chapter 'Acceptance Criteria for Validation Parameters'), the intermediate precision can be regarded as the relevant parameter, especially if the analytical investigations extend over several years such as during stability studies.

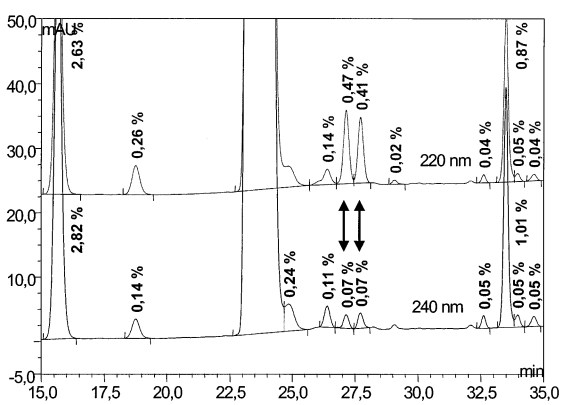

Fig. 4. Safety check for deviating response factors of unknown impurities in UV detection. In addition to the test procedure wavelength of 240 nm, relative peak area are displayed at 220 nm. The ratio for the labelled peaks is about seven. In further investigations, a response factor of 14 was determined.

Table 5
Analysis of variances for the investigation of intermediate precision[a]

	Series 1	Series 2
Number of determinations	6	7
Relative standard deviation	0.32%	0.63%
Mean	10.07	10.14
95% confidence interval	0.034	0.060
Overall mean	10.11	
Repeatability	0.52%	
Intermediate precision	0.65%	

[a] Two operators performed an assay of a tablet formulation (10 mg) with different LC-systems, mobile phases, and columns.

From the standard deviation, the repeatability limit can be calculated (Eq. (6)) which represents the maximum permitted difference between two repeated measurements. In this way, a straightforward verification is possible if the degree of scattering under the actual conditions is comparable to the validation. By this parameter, the precision obtained during the validation studies using a larger number of data (thus increasing the reliability) is linked to the routine analyses without requiring a larger number of experimental data.

$$r = t_{n-1,0.95} \times \sqrt{2} \times s \approx 2.8 \times s \quad (6)$$

Of course, a reliable estimate of the standard deviation is required to calculate appropriate repeatability limits.

Most statistical tests and calculations are based on the assumption that the experimental values are only influenced by random variability (i.e. that they are normally distributed). Data, which do not fulfill these assumptions (e.g. due to so called 'gross errors', weighing, dilution, or by problems with the instrument, etc.) will affect the results.

Such values can be identified by statistical outlier tests (e.g. according to Dixon or Grubbs [25]) in order to eliminate them before performing further calculations. However, the problem is — especially with a small number of data where groupings could easily occur — avoiding the incorrect rejection of values belonging to the same distribution. When an outlier is identified, the absolute magnitude of the standard deviation must also be considered for evaluation. If this parameter (calculated including the suspected 'outlier') lies in a normally expected range, preferably all values should be retained. Outlier tests should be applied carefully and only obviously deviating values ('gross errors') should be eliminated. It is preferable to increase the reliability of the obtained standard deviation by repeated investigations, e.g. using the overall repeatability obtained with an analysis of variances of an intermediate precision study.

4.5. Detection and quantitation limit

Several approaches are given in the ICH guideline to determine the detection and quantitation limits (Table 6). Generally, they are based either on the analysis of blanks or on the scattering (variability) of the analytical signals in the low concentration range.

Using the blank procedures, the corresponding calculation value is multiplied by the factors of 3.3 and 10 for the detection and quantitation limits, respectively. The calculation value may represent the signal of the blank, the standard deviation of the blank or of the intercept of a calibration line (corresponding to an extrapolated blank). In the latter two cases, the analytical signal is transformed by the slope of the calibration line into a concentration [8]. Using the calibration line directly, the aforementioned factors (3.3 and 10) can be multiplied by the ratio from the residual standard deviation and the slope (corresponding to the standard error of slope) [8].

Limits calculated or extrapolated by these procedures should be verified by the analysis of samples in the corresponding concentration range [8]. This additional verification is already included in other procedures which make direct use of the scattering around the calibration line by means of the 95% prediction interval around the regression line [27,28].

The quantitation limit can also be obtained directly from precision studies. For this approach, decreasing analyte concentrations are analysed repeatedly. The coefficient of variation (relative standard deviation) is plotted against the corresponding concentration. If a predefined limit for the coefficient of variation (relative standard deviation) is exceeded (e.g. 10 or 20%), the corresponding concentration is established as the quantitation limit [6,29]. However, a sufficiently large number of analyte concentrations must be analysed because a large scattering of standard deviations occur in the low concentration range.

More than the other validation characteristics, detection and quantitation limits are dependent on the equipment used and the actual conditions as well as on the method of calculation. In Fig. 5, the results of a repeated determination using five LC-systems over a period of about 9 months are shown. A calibration using a model compound was performed with six concentrations in the range from 0.05–1 μg/ml. Three approaches were applied to obtain the quantitation limit: the calculations from the residual standard deviation and from the 95% prediction interval were performed using the linear regression parameters, the concentrations corresponding to a signal-to-noise ratio of 10 were interpolated from the experimental

Table 6
Approaches for determining the detection and quantitation limit [8][a]

Approach	Detection limit	Quantitation limit
Visual evaluation	Minimum level detectable	Minimum level quantifiable
Signal-to-noise	3:1 or 2:1	10:1
Standard deviation of the response (σ)[b] and the slope (S)	$3.3 \times \sigma/S$	$10 \times \sigma/S$

[a] Verification with a suitable number of samples.
[b] Standard deviation of the blank, residual standard deviation of the calibration line, or standard deviation of the intercept.

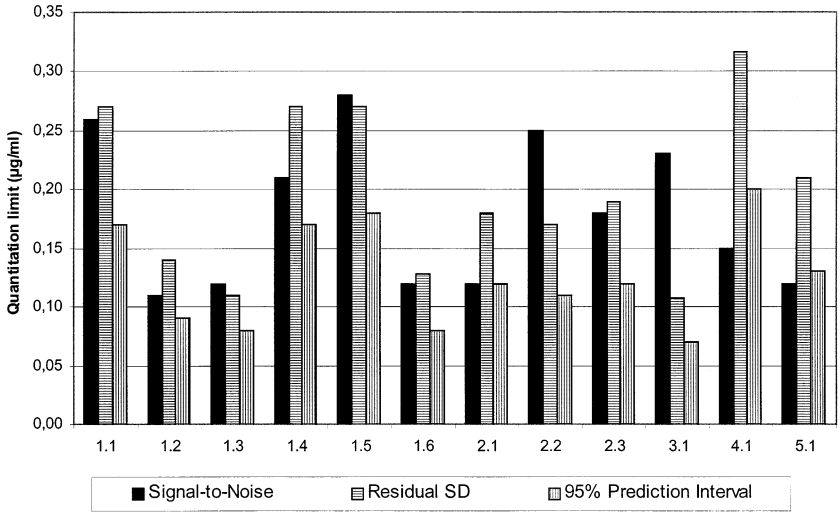

Fig. 5. Reproducibility of the quantitation limit ('intermediate QL'). Three different procedures were used to calculate the quantitation limit. 1.1–1.6, six determinations were performed on the same LC system; 2.1–2.3, three series were performed on a second LC-system; 3.1–5.1, three other LC-systems were used. The series were analysed over a time period of about 9 months.

signal-to-noise ratio of the concentrations below 0.5 μg/ml.

Large differences can be seen for the various calculation procedures (factor 2–3), but also in the case of repeated analysis for the same calculation (factor 2).

However, these limits are of special importance in the transfer of analytical procedures and for the reporting of impurities. Therefore, the (general) quantitation limit of the analytical procedure should be defined taking the requirements of the analytical procedure into account, rather than determined experimentally. In order to guarantee a reliable quantitation, it must be at least three standard deviations away from the specification limit. For orientation purposes — taking the higher variability in this concentration range into account — the quantitation limit may be defined as 50% of the specification limit. For example, with regard to an impurity specified as 0.1%, the quantitation limit can be established at 0.05%. This corresponds to the reporting thresholds of the ICH impurity guideline [9]. During validation, one of the described methods is used to verify (generally) whether this quantitation limit can be reliably achieved. Taking the random variability of the determination into account, it does not matter whether the 'actual' QL is determined to 0.02 or 0.04%, provided that the upper limit of the 'general' QL of 0.05% can be achieved. If impurities are present or can be spiked at the quantitation limit in batches used for assay precision studies, the standard deviation with respect to the impurities can be calculated from the same experimental data. The QL can be regarded as verified if an acceptable precision (e.g. below 10–20%) is achieved.

5. Conclusions

Beyond the regulatory requirements, the performance and reliability of the control test procedure are essential to the quality control of drugs. Results which reflect the quality of the analytical procedure more than the quality of the pharmaceutical to be tested may easily lead to drastic financial consequences. Therefore, validation should be regarded as part of an integrated concept to ensure the quality, safety, and efficacy of pharmaceuticals.

Based on the validation characteristics and requirements of the ICH guidelines, each analytical procedure must be validated with respect to parameters which are relevant to its performance.

It is the responsibility of the analyst to identify the critical performance parameters and design the validation study accordingly. Acceptance criteria should be defined in the validation protocol. They can be established from previous experiences (analytical state of the art) or calculated from specification limits. For the intended use of the test procedure acceptable absolute acceptance limits are preferred. Statistical tests should be used carefully and preferably for orientation purposes. The evaluation of the validation results is the responsibility of the analyst and must not be left or reduced to the outcome of a statistical test! However, statistical analysis and considerations are very helpful in verifying the compatibility of specification limits and analytical variability, calculating acceptance limits and performing simulations in order to predict future risks.

Relevant parameters for the evaluation of linearity are dependent on the intended calibration mode of the analytical procedure. For the evaluation and further calculations, the different levels of precision must be taken into account. Due to the large variability of experimentally obtained results, the 'general' quantitation limit should be defined according to the requirements, for example the reporting threshold for unknown impurities of 0.05%.

References

[1] Erläuterungen des BGA zum Antrag auf Zulassung eines Arzneimittels (Febr. 1988).

[2] The Rules Governing Medicinal Products in the European Community, volume 3 Addendum (1990).

[3] CDER Guideline on Validation of Chromatographic Methods, Reviewer Guidance of Chromatographic Methods, U.S. Food and Drug Administration, Center for Drugs and Biologics, Department of Health and Human Services (1994).

[4] Guidelines for Submitting Samples and Analytical Data for Methods Validation, U.S. Food and Drug Administration, Center for Drugs and Biologics, Department of Health and Human Services (1987).

[5] United States Pharmacopeia, Section ⟨1225⟩ 'Validation of Compendial Methods', United States Pharmacopeial Convention, Rockville (1995).

[6] A'cceptable Methods', Drug Directorate Guidelines, National Health and Welfare, Health Protection Branch, Health and Welfare Canada (1992).

[7] International Conference on the Harmonization of Technical Requirements for the Registration of Pharmaceuticals for Human Use (ICH) Q2A: Validation of Analytical Methods (Definitions and Terminology) (October 1994).

[8] ICH: Q2B Analytical Validation-Methodology (November 1996).

[9] ICH: Q3A, Impurities in New Drug Substances (March 1995, revised October 1999, Step 2).

[10] ICH: Q3B, Impurities in New Drug Products (November 1996, revised October 1999, Step 2).

[11] ICH: Q3C, Residual Solvents (July 1997).

[12] ICH: Q6A, Specifications: test procedures and acceptance criteria for new drug substances and new drug products, Chemical Substances (October 1999).

[13] ICH: Q6B, Specifications: test procedures and acceptance criteria for biotechnological/biological products (March 1999).

[14] L. Huber, LC-GC International, Febr. (1998) 96–105.

[15] A.G.J. Daas, J.H.McB. Miller, Pharmeuropa 9.1 (1997) 148–156.

[16] F.J. Van de Vaart, Pharmeuropa 9.1 (1997) 139–143.

[17] B. Renger, Analytical validation: formal requirements and practical approaches, solution for scientists symposium, 29–30 Nov. 1999, London.

[18] Draft guidance for industry on investigating out of specification (OOS) test results in pharmaceutical production. US Food and Drug Administration, Center for Drug Evaluation and Research (CDER) (1998).

[19] R.G. Whitfield, D.W. Hughes, T.P. Layloff, R.R. Cox, G.E. Gressett, P.J. Jimenez, J. Andersen, R.R. Reck, S. Schniepp, Pharmacopeial Forum 24 (1998) 7051–7056.

[20] B.-A. Persson, J. Vessman, R.D. McDowall. LC-GC International, March 1998, 160-164.

[21] C.M. Riley, Statistical parameters and analytical figures of merit, in: C.M. Riley, T.W. Rosanske (Eds.), Development and Validation of Analytical Methods, Elsevier, Oxford, 1996, p. 15.

[22] V.R. Meyer, Chromatographia 40 (1995) 15–22.

[23] J. Caporal-Gautier, J.M. Nivet, P. Algranti, M. Guilloteau, M. Histe, M. Lallier, J.J. N'Guyen-Huu, R. Russotto, S.T.P. Pharma Pratiques 2 (4) (1992) 205–226.

[24] Guidance for industry: analytical procedures and methods validation, chemistry, manufacturing, and controls documentation, Draft Guidance (Food and Drug Administration), August 2000.

[25] DIN ISO 5725-2: Accuracy (trueness and precision) of measurement methods and results; a basic method for the determination of repeatability and reproducibility of a standard measurement method (1990).

[26] MVA-Method validation in analytics (PC-software, Windows NT), NOVIA GmbH, Saarbrücken, Germany (http://www.novia.de).
[27] W. Funk, V. Dammann, C. Vonderheid, G. Oehlmann, Statistische Methoden in der Wasseranalytik, Verlag Chemie, Weinheim, 1985.
[28] DIN 32 645: Chemische Analytik: Nachweis-, Erfassungs- und Bestimmungsgrenze, Ermittlung unter Wiederholbedingungen. Begriff, Verfahren, Auswertung. Beuth Verlag GmbH, Berlin.
[29] EURACHEM Guidance Document No. WDG 2: Accreditation for chemical laboratories: Guidance on the interpretation of the EN 45000 series of standards and ISO/IEC Guide 25, 1993.