1
Analytical Validation within the Pharmaceutical Environment

Joachim Ermer

Validation is, of course, a basic requirement to ensure quality and reliability of the results for all analytical applications [8]. However, in comparison with analytical chemistry, in pharmaceutical analysis, some special aspects and conditions exist that need to be taken into consideration. For example, the analytical procedures (apart from pharmacopoeial monographs) are often in-house developments and applications. Therefore, the degree of knowledge and expertise is initially much larger compared with standard methods. The same can be assumed for the samples analysed. The matrix (placebo) in pharmaceutical analysis is usually constant and well known and the ranges where the sample under analysis can be expected are usually well defined and not very large. Evaluation (of batches, stability investigations, etc.) is based on the results of various procedures or control tests, thus their performances can complement each other. Acceptance limits of the specification are fixed values, often based on tradition, as in the case of assay of an active ingredient, or they may be based on specific toxicological studies, which take large safety factors into account, as for impurities. Last, but not least, validation in pharmaceutical analysis has its own regulations. These few – by far from exhaustive – remarks should make it obvious that these special considerations will have an impact on the way validation in pharmaceutical analysis is performed.

The first part of this book focusses on the fundamentals of validation in pharmaceutical analysis, the 'environmental' framework as well as the *implications* for experimental design and suitable calculations. Of course, the basic principles of validation are the same for any analytical procedure, regardless of its field of application. However, the discussions and recommendations focus on pharmaceutical applications, so the reader needs to adjust these to suit his or her purpose, if different. Nevertheless – as validation should never be regarded as simply working through a checklist – this is also required in the case of pharmaceutical analysis, but perhaps to a lesser extent, compared with other areas of application.

Method Validation in Pharmaceutical Analysis. A Guide to Best Practice. Joachim Ermer, John H. McB. Miller (Eds.)

ISBN: 3-527-31255-2

1.1 Regulatory Requirements

"The object of validation of an analytical procedure is to demonstrate that it is suitable for its intended purpose" [1a], determined by means of well-documented experimental studies. Accuracy and reliability of the analytical results is crucial for ensuring quality, safety and efficacy of pharmaceuticals. For this reason, regulatory requirements have been published for many years [1–7].

The International Conference on the Harmonisation of Technical Requirements for the Registration of Pharmaceuticals for Human Use (ICH) was initiated in 1990, as a forum for a constructive dialogue between regulatory authorities and industry, in order to harmonise the submission requirements for new pharmaceuticals between Europe, the United States of America and Japan. One of the first topics within the Quality section was analytical validation and the ICH was very helpful in harmonising terms and definitions [1a] as well as determining the basic requirements [1b]. Of course, due to the nature of the harmonisation process, there are some compromises and inconsistencies. In Table 1-1, the required validation characteristics for the various types of analytical procedures are shown.

Table 1-1: Validation characteristics normally evaluated for the different types of test procedures [1a] and the minimum number of determinations required [1b]

		Analytical procedure			
Validation characteristic	**Minimum number**	**Identity**	**Impurities**		**Assay**[1]
			Quantitative	**Limit**	
1. Specificity [2]	Not applicable	Yes	Yes	Yes	Yes
2. Linearity	5	No	Yes	No	Yes
3. Range	Not applicable	No	Yes	No	Yes
4. Accuracy	9 (e.g. 3 × 3)	No	Yes	No	Yes
5. Precision					
Repeatability	6 or 9 (e.g. 3 × 3)	No	Yes	No	Yes
Intermediate precision/ Reproducibility [3]	(2 series)[4]	No	Yes	No	Yes
6. Detection limit	Approach dependent	No	No [5]	Yes	No
7. Quantitation limit		No	Yes	No	No

Yes / No normally evaluated / not evaluated

1 including dissolution, content/potency

2 lack of specificity of one analytical procedure could be compensated by other supporting analytical procedure(s)

3 reproducibility not needed for submission

4 no number given in [1b], logical conclusion

5 may be needed in some cases

Two guidelines on validation were issued by the US Food and Drug Administration (FDA), one for the applicant [2], the other for inspectors and reviewers [3]. The first one is also intended to ensure that the analytical procedure can be applied in an FDA laboratory and therefore requires a detailed description of the procedure, reference materials, as well as a discussion of the potential impurities, etc. The second guideline focuses on reversed-phase chromatography and provides a lot of details with regard to critical methodological issues, as well as some indication of acceptability of results. A revised draft of the first guideline was published in 2000 [4]. According to the title "*Analytical procedures and methods validation*", it also includes the content and format of the analytical procedures, the requirements for reference standards and various types of analytical technique. Therefore, this guidance is more comprehensive than the ICH Guidelines, but is rather too focussed on providing 'instrument output/raw data'. As this is an inspection and documentation issue, it should be separated from the validation. A very detailed discussion is provided in the Canadian guideline [7] with respect to requirements and particularly acceptance criteria. Although this allows some orientation, the given acceptance criteria were sometimes rather too ambiguous, for example, the intermediate precision / reproducibility of less than 1% for drug substances (see Section 2.1.3.2 and Fig. 2.1-12).

So why is it still important to discuss validation?

First of all, the ICH guidelines should be regarded as the basis and philosophical background to analytical validation, not as a checklist. *"It is the responsibility of the applicant to choose the validation procedure and protocol most suitable for their product"* [1b]. It will be shown in the next sections that suitability is strongly connected with the requirements and design of the given analytical procedure. As this obviously varies, at least with the type of procedure, it must be reflected in the analytical validation. This includes the identification of the performance parameters relevant for the given procedure, the definition of appropriate acceptance criteria and the appropriate design of the validation studies. In order to achieve this, the analyst must be aware of the fundamental meaning of these performance parameters, as well as the calculations and tests and their relationship to the specific application. The former is discussed in detail in Chapter 2, the latter in the following sections. A lack of knowledge or (perhaps) a wrong understanding of 'efficiency' will lead to validation results that address the real performance of the analytical procedure only partly or insufficiently. This is, at the very least a waste of work, because the results are meaningless. Unfortunately, this can also be found rather too frequently in publications, although to a varying extent for the different validation characteristics. Such common insufficiencies are discussed in the respective sections of Chapter 2.

1.2 Integrated and Continuous Validation

Validation should not be regarded as a singular activity [4], but should always be understood with respect to the life cycle of the analytical procedure. Starting with the method development or optimisation, the performance of the analytical proce-

dure should be matched to the requirements in an *iterative process*. Some validation characteristics, such as specificity (selective separation) or robustness, are more important in this stage (see Section 2.7). However, this depends on the type of procedure. In the case of a complex sample preparation, or cleaning methods (see Section 2.3.4), precision and accuracy may play an important role in the optimisation process. One should also be aware that the validation requested for submission, i. e. a demonstration of the *general* suitability of the respective analytical procedure – can only be considered as a basis. The user of any method has to guarantee that it will stay consistently in a validated status, also referred to as the life-cycle concept of analytical validation [9]. In this process, an increasing amount of information can be compiled.

This does not necessarily mean that additional work always needs to be done. During the actual application of the methods, a lot of data is generated, but often left unused ('data graveyard'). In order to make rational and efficient use of these data, they must be transformed to information (i.e., processed and condensed into performance parameters). When enough reliable information is compiled, it can be further processed to gain knowledge that eventually enables us to achieve a better understanding and control of the analytical procedure (see also Section 2.1.4 and Chapter 9). The whole process is well known as an *'information pyramid'* (Fig. 1-1). This knowledge can also be used to improve analytical procedures, for example, by changing from the traditional 'daily' calibration in an LC assay to a quantitation using 'predetermined' calibration parameters (comparable to a specific absorbance in spectrophotometry), with advantages both in efficiency and reduced analytical variability [10].

Transfers of analytical procedures to another site of the company or to a contract laboratory – quite common nowadays – often result in a challenging robustness test, especially if not appropriately addressed in the validation. Acceptance criteria for a successful transfer may be derived from the validation itself, or from the same principles as for calculations and tests in validation, because here the performance of the analytical procedure is also addressed (see Chapter 7). On the other hand, comparative studies will provide quite reliable performance data of the analytical procedure (see Section 2.1.3.2).

Besides this 'horizontal' integration, analytical validation also needs to be included in the whole system of *Analytical Quality Assurance* (AQA) [8], i.e., 'vertical' integration. This involves all (internal and external) measures which will ensure the quality and reliability of the analytical data, such as an equipment qualification program (see Chapter 4), appropriate system suitability tests (see Section 2.8), good documentation and review practices, operator training, control charts (see Chapter 9), etc.

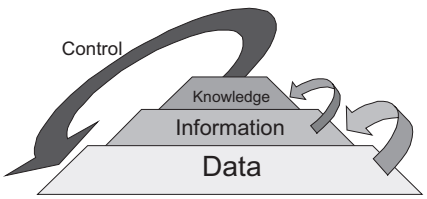

Figure 1-1 Information pyramid.

1.3 General Planning and Design of Validation Studies

Performance is strongly connected with the requirements and design of the given analytical procedure (see Section 1.4.1). As this obviously varies, it must be reflected in the planning and design of the analytical validation. Consequently, a checklist approach is not appropriate. In order to ensure thorough planning, i.e., to identify the relevant performance parameters, to define appropriate acceptance criteria and then to design the studies accordingly, *validation protocols* should be prepared. In addition to this 'good science' reason, protocols can also be regarded as a general GMP requirement and are common practice also the in case of process validation, cleaning validation, equipment qualification, transfer, etc.

The analyst may be faced with the problem of the iterative nature of the method development / validation process. However, here one may distinguish between performance parameters (and the corresponding validation characteristics) of the final analytical procedure and those obtained or derived from different method conditions, such as specificity and robustness. The former can be addressed (before starting the experimental studies, following usual practice) in the protocol, the latter can be referred to in the validation report and/or protocol (see Chapter 5).

Of course, the *extent and depth* of the validation studies, as well as acceptance criteria, should be defined in relation to the required performance ('importance') and the 'environment' of the respective analytical procedure, such as the stages of development (see Chapter 5), or the stages of manufacturing / synthesis. Important or critical procedures (within the context of validation) can be expected to have tighter specification limits. In these cases, such as the assay of active or of critical impurities, it is recommended to address the validation characteristics *separately* (for example, precision with authentic samples and accuracy with spiked samples), in order to increase the power of the results. In other cases, such as the determination of other ingredients or of impurities or water sufficiently below specification limits, several validation characteristics, for example, precision, linearity, and accuracy (quantitation) limit in dependence on the range, see Section 2.6.4) can be investigated simultaneously, using the same spiked samples.

The ICH Guidelines [1a,b] are mainly focused on chromatographic procedures, as can be seen in the methodology guideline [1b]. Therefore, they should be regarded more as a guide to the *philosophy of validation* – i.e., used to identify relevant performance parameters of the given analytical procedure – than as a 'holy grail'. If the special conditions or techniques are not covered in the ICH guideline, the validation approach must then be adapted accordingly (see Chapter 11). The FDA Guidance [4], and the Technical Guide of the European Pharmacopoeia (EP) [11], as well as Chapter 8 also provide details for specific analytical techniques.

1.3.1 Always Look on the 'Routine' Side of Validation

Curiously, one aspect often neglected during validation is its primary objective, i.e., to obtain the real *performance of the routine application* of the analytical procedure. As far as possible, all steps of the procedure should be performed as described in the control test. Of course, this cannot always achieved, but at least the analyst should always be aware of such differences, in order to evaluate the results properly.

What does this mean in practice?

For example, precision should preferably be investigated using *authentic* samples, because only in this case is the sample preparation identical to the routine application. It is also important to apply the intended calibration mode exactly as described in the analytical procedure. Sometimes the latter is not even mentioned in the literature. Precision is reported only from repeated injections of the same solution, ignoring the whole sample preparation. This is certainly not representative for the (routine) variability of the analytical procedure (see Section 2.1.2). Investigating pure solutions is usually of very limited practical use, for example, in the case of cleaning methods (see Section 2.3.4) or quantitation limit (see Section 2.6), or may even lead to wrong conclusions, as the following examples will show.

The minor (impurity) enantiomer of a chiral active ingredient was analysed by chiral LC using an immobilised enzyme column (Chiral-CBH 5 µm, 100 × 4 mm, ChromTech). The quantitation should be carried out by area normalisation (100%-method, 100%-standard), which would require a linear response function and a negligible intercept for both active and impurity enantiomer (see also Section 2.4.1). The experimental linearity investigation of dilutions of the active, revealed a clear deviation from a linear response function (Fig. 1-2). However, when the design was adjusted to simulate the conditions of the routine application, i.e., spiking the impurity enantiomer to the nominal concentration of the active, an acceptable linear relationship was found. Although a slight trend remained in the results, the recoveries between 99 and 105% can be regarded as acceptable for the intended purpose. A possible explanation for such behaviour might be that the interaction between the enantiomers and the binding centres of the immobilised enzyme (cellobiohydrolase, hydrolysing crystalline cellulose) is concentration dependent. Maintaining the nominal test concentration in the case of the spiked samples, the sum of both enantiomers is kept constant and consequently so are the conditions for interactions. In this case, the linearity of the active enantiomer cannot be investigated separately and the validity of the 100% method must be demonstrated by obtaining an acceptable recovery.

Stress samples

Another area where the primary focus of validation is often ignored is the use of stress test samples (see also Section 2.2). At least some of the applied conditions [1g] will result in degradation products without any relevance for the intended storage condition of the drug product. Therefore, such samples should be used with reasonable judgement for method development and validation. It is the primary objective of a suitable (impurity) procedure (and consequently its validation) to address degra-

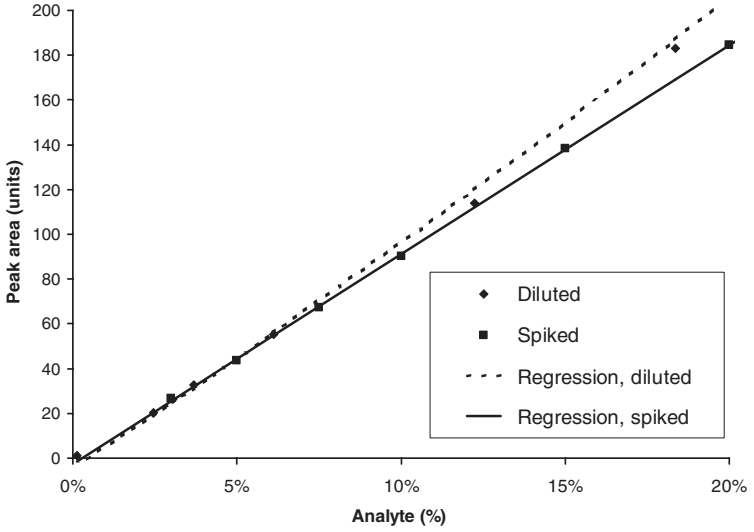

Figure 1-2 Linearity investigation of an enantiomeric LC determination. The diamonds and the squares represent dilutions of the active enantiomer and spikings of the impurity enantiomer to the active enantiomer, respectively. An obvious deviation from a linear function is observed in the case of the dilutions (broken line, polynomial to 3rd order), in contrast to the impurity in the presence of the active enantiomer (solid line, linear regression). The concentration on the x-axis is given with reference to the nominal test concentration of the active enantiomer.

dants "likely to be present" [1b], rather than a 'last resort'. However, it is also reasonable to allow for some 'buffer' [12].

Sometimes, applying artificial conditions cannot be avoided, in order to approach validation parameters, as in recovery investigations (see Section 2.3.2) or in dissolution, where no homogeneous samples are available. In the latter case, the assay part of the analytical procedure may be investigated separately. However, possible influences on the results due to the different application conditions need to be taken into account in the evaluation process as well as in the definition of acceptance criteria.

1.4 Evaluation and Acceptance Criteria

1.4.1 What does Suitability Mean?

The suitability of an analytical procedure is primarily determined by the requirements of the given test item, and secondly by its design (which is normally more flexible). Usually, the (minimum) requirements are defined by the acceptance limits of the specification (often termed traditionally as 'specification limits', but according to ICH [1e], the term 'specification' defines a *"list of tests, references to analytical proce-*

dures, and appropriate acceptance criteria"). For some applications, the requirements are explicitly defined in the ICH Guidelines. For example, the reporting level for unknown degradants in drug products is set to 0.1% and 0.05% for a maximum daily intake of less and more than 1 g active, respectively [1d] (Table 2.6-1). In the case of cleaning validation, the maximum acceptable amount of cross-contamination can be calculated based on the batch sizes and doses of the previous and subsequent product, the toxicological or pharmacological activity and/or the safety factors, and the so called specific residual cleaning limit (SRCL) [13]. Consequently, the corresponding test procedure must be able to quantify impurities or residual substance at this concentration with an appropriate level of precision and accuracy (see Section 2.3.4).

With respect to stability studies, the analytical variability must be appropriate to detect a (not acceptable) change in the tested property of the batch. This is illustrated in Figure 1-3 for determination of the content of active ingredient. The intrinsic degradation of 1.0% within 36 months can be reliably detected by an assay with a true variability of 0.5% (Fig. 1-3A), but not by one with 2.0% variability (Fig. 1-3B). Generally, acceptance limits of the specification (SL) have to enclose (at least) both the analytical and the manufacturing variability (see Chapter 6). Rearranging the equation describing this relationship (Eq. 6-12), the maximum permitted analytical variability can be calculated from the acceptance limits of the specification (Eq.1-1).

$$RSD_{\max}(\%) = \frac{|(BL-SL)| * \sqrt{n_{assay}}}{t(P,df)} \tag{1-1}$$

SL: Acceptance limits of the specification for active (% label claim).

BL: Basic limits, 100% – maximum variation of the manufacturing process (in %). In case of shelf-life limits, the lower basic limit will additionally include the maximum acceptable decrease in the content.

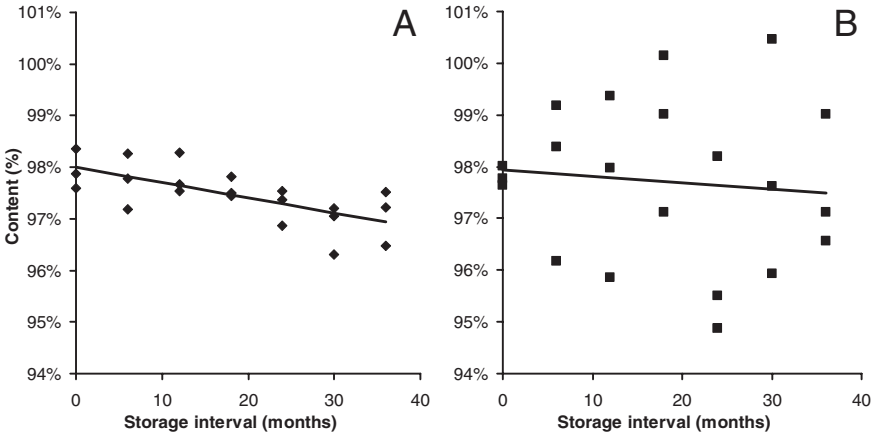

Figure 1-3 Illustration of the requirements for assay of an active ingredient during a stability study. The three individual results per storage interval were simulated based on a 1% decrease of content within 36 months and a normally distributed error of 0.5% (A) and 2.0% (B) using Eq. (2.1-3). The slope of the regression line in B is not significant.

n_{assay}: Number of repeated, independent determinations in routine analyses, insofar as the mean is the reportable result, i.e., is compared to the acceptance limits. If each individual determination is defined as the reportable result, n=1 has to be used.

$t(P,df)$: Student t-factor for the defined level of statistical confidence (usually 95%) and the degrees of freedom in the respective precision study.

The same basic considerations of the relationship between content limits and analytical variability [14] were applied to the system precision (injection repeatability) requirements of the EP [15] (see Section 2.8.3.8). The method capability index (see Section 10.5, Eq. 10-5) is based on similar considerations. However, here the normal distribution is used to describe the range required for the analytical variability (see Section 2.1.1). Consequently, the method capability index must be applied to single determinations (or to means if the standard deviation of means is used) and requires a very reliable standard deviation, whereas Eq.(1-1) can take a variable number of determinations directly into account, as well as the reliability of the experimental standard deviation (by means of the Student t-factor).

Of course, the precision acceptance limit thus obtained will be the minimum requirement. If a tighter control is needed, or if a lower variability is expected for the given type of method (analytical state of the art, see Section 2.1.3), the acceptance limits should be adjusted. A further adjustment may be required if there is a larger difference between repeatability and intermediate precision, i.e., if there is a larger inter-serial contribution (Eq. (2.1-10), Section 2.1.3.2). In such a case, an increased number of determinations in the assay will only reduce the repeatability variance, but not the variance between the series (s^2_g). Therefore, the term $\sqrt{n_{assay}}$ must be transferred to the left-hand side of Eq. (1-1) and RSD_{max}(%) rearranged to $\sqrt{s_g^2 + \frac{s_r^2}{n_{assay}}}$. This term corresponds to the standard deviation of the means from the routine assay determinations.

Many other performance parameters are linked with the analytical variability. Therefore, once an acceptable precision is defined, it can serve as an orientation for other acceptance criteria (for details, see Table 1-2 and Sections 2.1–2.6). As far as possible, *normalised (percentage) parameters* should be defined as validation acceptance limits, because they can be compared across methods and therefore more easily drawn from previous experience.

As can be seen from Eq. (1-1), the number of determinations also influences the acceptable performance, as well as the intended calibration mode (see Section 2.4). In principle, the analyst is rather flexible in his/her decision, provided that the minimum requirements are fulfilled. Often, the design of the calibration is more influenced by tradition or technical restrictions (for example the capabilities of the acquisition software) than by scientific reasons. Sometimes a 'check standard' is applied, i.e., the standard prepared and used for calibration is verified by a second standard preparation, the response of which needs to be within an acceptable range of the first one (e.g. ± 1.0%). This approach is not optimal. If the 'check standard' is only used for verification, 50% of the available data are ignored. Increasing the number of determi-

nations improves the reliability of the mean (see Fig. 2.1-4A). Therefore, it would be preferable to calculate the mean from all standard preparations (after verification of their agreement), in order to reduce the variability of the standard that will be included in the result for the sample (see discussion on repeatability and intermediate precision, Section 2.1.2). Of course, if the overall variability utilising only the first standard preparation is still acceptable, the procedure will be suitable. However, the analyst must be aware of the inter-relations and their consequences in order to make an appropriate decision and evaluation. This example also highlights the importance of applying the intended calibration, exactly as described in the control test for the intermediate precision study, otherwise the obtained result will not reflect the performance of the routine analytical procedure.

1.4.2 Statistical Tests

Significance Tests

Statistical significance tests should very cautiously be (directly) applied as acceptance criteria, because they can only test for a *statistical* significance (and with respect to the actual variability). On one hand, due to the small number of data normally used in pharmaceutical analysis, large confidence intervals (see Section 2.1.1) may obscure unacceptable differences (Fig. 1-4, scenario 3, S). On the other hand, because of sometimes abnormally small variabilities in (one of) the analytical series (that, however, pose no risk for routine application), differences are identified as significant which are of no *practical relevance* (Fig. 1-4, scenario 1, S) [16]. The analyst must decide whether or not detected statistical differences are of practical relevance. 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. Therefore, a statistical significance test should always be applied (as acceptance criteria) in a two-tiered manner, including a measure for practical relevance. For example, in the case of comparison of results with a target value, in addition to the nominal value *t*-test (see Section 2.3.1, Eq. 2.3-2), an upper limit for the precision and a maximum acceptable difference between the mean and the target value should be defined, in order to avoid the scenario 3 illustrated in Figure 1-4 (S).

Equivalence Tests

Such measures of practical relevance are an intrinsic part of the so-called *equivalence tests* [16, 28] (see also Section 7.3.1.3). In contrast to the significance tests, where the confidence intervals of the respective parameter(s) must include the target value (Fig. 1-4, scenario 2 and 3, S), equivalence tests, must be within an *acceptable range*. This measure of practical relevance is defined by the analyst. It is obvious in Figure 1-4, that such equivalence tests are robust with respect to small (scenario 1, E), but sensitive to large (scenario 3, E) variabilities.

Absoute Acceptance Limit

Another alternative is to use *absolute acceptance limits*, derived from experience (see Section 2.1.3) or from statistical considerations, as described in Section 1.4.1 for pre-

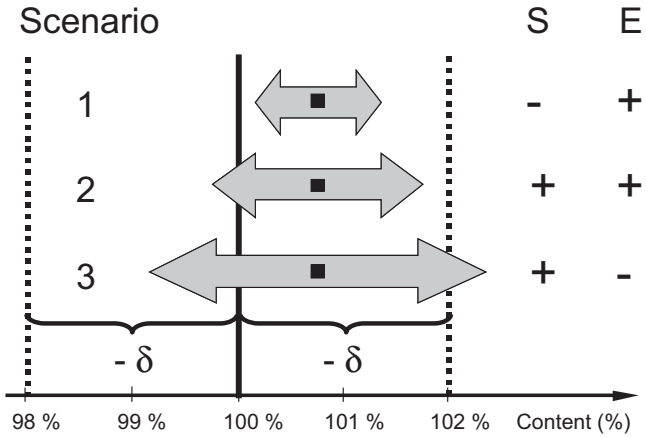

Figure 1-4 Illustration of statistical significance (S) and equivalence (E) tests for the example of a comparison between a mean and a target value of 100% (e.g., a reference or theoretical recovery). The acceptable deviation δ from the target (for the equivalence test) is symbolised by vertical dotted lines, the means, with confidence intervals indicated by double arrows. The outcome of the statistical tests for the three scenarios is indicated by '+' and '–' for 'pass' and 'fail' of the respective (H0) hypothesis, these are 'no statistical significant difference' and 'acceptable difference' for significance and equivalence test, respectively.

cision, and for a maximum acceptable difference in accuracy (see Section 2.3.5). In contrast to the equivalence tests, the actual variability of the data is neglected for the purpose of comparison (if means are used). However, usually the variability will be investigated separately.

If validation software is used, it must be flexible enough to meet these precautions [28].

Of course, statistical significance tests also have their merits, if properly applied. Even if a small variability does not pose a practical risk, when the suitability of a procedure is investigated, it may be assumed that such data are not representative for the usual (routine) application of the analytical procedure. This is an important consideration when the true parameter (standard deviation, mean) is the investigational objective, for example, the true precision of an analytical procedure, or if a reference standard is characterised. In collaborative trials, significance tests such as outlier tests are often defined as intermediary acceptance criteria for checking the quality of the data [17–19]. Deviating (i.e., unrepresentative) results (laboratories) are removed before proceeding to the next step, in which results are combined.

1.5 Key Points

- Validation should address the performance of the analytical procedure under conditions of routine use.
- Suitability is strongly connected with both the requirements and the design of the individual analytical procedure.
- Consequently, the analyst has to identify relevant parameters which reflect the routine performance of the given analytical procedure, to design the experimental studies accordingly and to define acceptance criteria for the results generated.
- Absolute, preferably normalised parameters should be selected as acceptance criteria. These can be defined from (regulatory) requirements, statistical considerations, or experience. Statistical significance tests should be applied with caution, they do not take into consideration the practical relevance.
- Validation must not be regarded as a singular event. The analyst is responsible for the continued maintenance of the validated status of an analytical procedure.

Acknowledgements

Some of the examples presented in my chapters, as well as the experience gained, are based on the work of, and discussion with, many colleagues in Aventis. Their important input is greatfully acknowledged, but I will abstain from an attempt to list them, both because of space as well as the danger of forgetting some of them. I would like to acknowledge in particular, John Landy, Heiko Meier, and Eva Piepenbrock.

Table 1-2 Examples of performance parameters in analytical validation. The acceptance criteria given are for *orientation purposes only*. They refer mainly to LC/GC procedures and have to be adjusted according to the requirements and the type of the individual test procedure. For details, see the respective sections in Chapter 2.

Validation characteristics parameter / calculations[1]	Type of analytical procedure / application	Acceptance criteria (orientational!)	Conditions / comment
Precision			
System precision (injection repeatability)	Assay	< 1%	Mainly reflection of the instrument (injection) variability, if sufficiently above QL
	Assay (DS)	According to EP [15]	Dependence on *n* and upper specification limit
	Assay (DP)	< 2% (USP)	Usually not sufficiently discriminative
	Impurities	< 2–5%	The smaller the concentration, the greater the influence of the detection/integration error
Repeatability			Preferably, authentic samples should be used
	Assay	Calculation from specification limits (Eq. 1-1)	Minimum requirement to achieve compatibility with specification limits
		< ≈ 1 – 2% (< 2 * TSD)	Dependent on type of DP (sample / preparation)
	Impurities	At QL: calculation from specification limits (Eq. 1-1, BL=QL)	Minimum requirement
		< ≈ 10 – 20%	Dependent on concentration level, preferably linked to QL
Intermediate precision / reproducibility			Analysis of variances (ANOVA)
Overall repeatability	Assay	< ≈ 1.5* TSD	More reliable due to increased number of determinations
Int. prec. / reproducibility		< ≈ 3–4* TSD	Dependent on type of DP (sample / preparation)
Specificity (quantitatively)			
Comparison with an independent procedure		see Accuracy	
Resolution factor	Chromatographic separations	> ≈ 2 (large difference in size)	For baseline-separated peaks, dependent on size difference, tailing and elution order
		> ≈ 1 (similar size)	
Peak-to-valley ratio	Chromatographic separations	> ≈ 0.25	For partly separated peaks

Table 1-2 Continued.

Validation characteristics parameter / calculations[1]	Type of analytical procedure / application	Acceptance criteria (orientational!)	Conditions / comment
Accuracy			
Comparison with an independent procedure or with a reference			Likely different specificities have to be taken into account
Difference between the means / to the reference	Assay (DS, n=6) Assay (DP, n=6) Impurities (n=6)	< ≈ 1 – 2% < ≈ 2% < ≈ 10 – 20%	Acceptable precision (of the most variable procedure) in the given concentration range
t-test		No significant difference between the means (95% level of significance)	Statistical significance test (see 1.4.2), only if specificities of the two procedures are the same or can be corrected.
Equivalence test	Assay (DS, n=6) Assay (DP, n=6) Impurities (n=6)	± ≈ 2% ± ≈ 3% ± ≈ 10 – 20%	Two-fold acceptable precision in the given concentration range; in contrast to simple comparison, the variability needs to be included (see 1.4.2)
Recovery			Spiking of known amounts of analyte into the respective matrix
Percent recovery			Concentration range < factor 10
Range of recovery mean	Assay (DP, n=9) Impurities (n=9)	≈ 98–102% ≈ 80/90 – 110/120%	Acceptable precision in the given concentration range
Statistical evaluation		95% confidence interval of the mean includes 100%	Statistical significance test (see 1.4.2)
		95% confidence interval within 96 – 104%	Statistical equivalence test, definition of a practical acceptable deviation (see 1.4.2)
Relative standard deviation	Assay (DP, n=9) Impurities (n=9)	< ≈ 2% < ≈ 10 – 20%	Weighing effects: small concentrations have a larger influence on the result
Individual recoveries		No systematic trend	Graphical presentation strongly recommended.
Range of individual recoveries	Assay (DP, n=9)	≈ 97 –103%	Corresponds to ≈ 6*TSD
	Impurities (n=9)	≈ 70 – 130%	Dependent on smallest concentration
Recovery function (unweighted linear regression)			
Amount added vs. amount found			

Table 1-2 Continued.

Validation characteristics parameter / calculations[1]	Type of analytical procedure / application	Acceptance criteria (orientational!)	Conditions / comment
Slope	Assay (DP, n=9)	≈ 0.98 – 1.02	Larger weight of higher concentrations
	Impurities (n=9)	≈ 0.9 – 1.1	
Confidence interval of the slope		95% CI includes 1	Statistical significance test (see 1.4.2)
	Assay (DP, *n*=9)	95% CI within 0.96 – 1.04	Statistical equivalence test, definition of a practical acceptable deviation (see 1.4.2)
	Impurities (*n*=9)	95% CI within 0.90 – 1.1	
Further parameter see Linearity			
Linearity			Verification of the intended calibration model
Unweighted linear regression	Single-point calibration		Concentration range < factor 10 (constant variability over the whole range required)
	Multiple point calibration		
Residual plot		Random scatter, no systematic trend	
	Assay	± ≈ 2% around zero	Corresponds to ± 3*TSD, at higher concentrations
	Impurities	± ≈ 10 – 20% around zero	
Sensitivity plot		No systematic trend	If intercept negligible Corresponds to ≈ ± 3*TSD, at lower concentrations (larger weight)
	Assay	± ≈ 3% around the mean	
	Impurities	± ≈ 10 – 20% around the mean	
Numerical parameters are only meaningful after verification/demonstration of a linear function			
Residual standard deviation	Assay (DS)	< ≈ 1 – 1.5%	Acceptable precision in the given concentration range
	Assay (DP, spiking)	< ≈ 2 – 3%	
	Impurities	< ≈ 10 – 20%	
Coefficient of correlation		No suitable for quantitative measure of linearity! Relation to the experimental variability depends on number of values and concentration range.	
Statistical linearity tests		Only recommended in case of indication or assumption of non-linearity (statistical significance vs. practical relevance).	

Table 1-2 Continued.

Validation characteristics parameter / calculations[1]	Type of analytical procedure / application	Acceptance criteria (orientational!)	Conditions / comment
Mandel test		No significant better fit by quadratic regression	Statistical significance test (see 1.4.2)
Significance of the quadratic coefficient		95% CI includes zero	Statistical significance test (see 1.4.2)
ANOVA lack of fit		Measurement variability (pure error) > than deviation from linear regression line	Statistical significance test (see 1.4.2), requires replicate determinations for each concentration level
Absence of a constant systematic error			Required for single point calibration (external standard) and 100% method
Intercept as % signal at working or target concentration	Assay Impurities	< ≈ 1 – 1.5% < ≈ 10 – 20%	Acceptable precision in the given concentration range, avoid large extrapolation
Statistical evaluation of the intercept		95% CI includes zero	Statistical significance test (see 1.4.2)
		95% CI within –2% and + 2%	Statistical equivalence test, definition of a practical acceptable deviation (see 1.4.2)
Deviation between single-point and multiple-point calibration line within the working range		Maximum deviation < ≈ 1%	Error from the calibration model should be less than TSD
Weighted linear regression	Multiple point calibration	If quantitation is required over a larger concentration range (> factor 10–20), when variances are not constant. In case of constant matrix and negligible intercept, a single-point calibration is also appropriate.	
Residual plot (absolute)		Random scatter around zero	Deviations are concentration dependent (wedge-shaped)
Residual plot (relative)		Random scatter around zero, no systematic trend	Deviations dependent on the precision of the respective concentration
Non-linear regression	Non-linear calibration		Non-linear response function
Residual plot	Assay Impurities	Random scatter, no systematic trend ± ≈ 1–2% around zero ± ≈ 10 – 20% around zero	Corresponds to ≈ ± 3*TSD, at higher concentrations

Table 1-2 Continued.

Validation characteristics parameter / calculations[1]	Type of analytical procedure / application	Acceptance criteria (orientational!)	Conditions / comment
Detection and Quantitation limit		If required, DL corresponds to QL/3	Be aware of the high variability of the actual QL. Usually a general QL is required, valid for any future application.
Establishment from reporting thresholds	Unknown impurities (DS)	0.03% or 0.05%	> 1g or < 1g daily dose
	Unknown degradants (DP)	0.05% or 0.1%	> 2g or < 2g daily dose
Establishment from specification limits	Impurities, cleaning methods	50% SL	
Calculation from specification limits	Impurities, cleaning methods	According to Eq. (2.6-3)	Minimum requirement
Acceptable precision	Impurities, cleaning methods	RSD < ≈ 10 – 20%	
'Intermediate QL'	Impurities, cleaning methods	QL_{max} or Eq. (2.6-2)	Repeated determinations of QL

Abbreviations:
TSD = target standard deviation, average repeatability of a sufficient number of determinations/series, estimation for the true repeatability of the given analytical procedure or type of drug product (sample complexity/preparation)
DS = drug substance
DP = drug product
CI = confidence interval
SL = acceptance limit of the specification
D/QL = detection / quantitation limit
1: Only parameters and calculations are listed for which general acceptance criteria can be given. As some of the parameters are alternative possibilities, the analyst has to choose the parameters/tests most suitable for his/her purposes.