Transfer of Analytical Procedures: Position Paper

M. Limberger [a, 1], J. Ermer [b, 1], K. Lis[a], T. Faust[d], I. Astner[e], D. Behrens[f], H. Höwer-Fritzen[g] and H. Wätzig [c]

[1]The authors contributed equally to this work
[a]Phast GmbH, Homburg/Saar, Germany
[b]Sanofi-Aventis Deutschland GmbH, Frankfurt, Germany
[c]University of Braunschweig, Germany
[d]Actavis Deutschland GmbH, München-Riem, Germany
[e]Gewerbeaufsichtsamt Braunschweig, Germany
[f]Piramal Enterprises, Eckernförde, Germany
[g]Dr. Willmar Schwabe GmbH & Co. KG, Ettlingen, Germany

Chapters:

Abstract

Analytical method transfers are certainly among the most discussed topics in the GMP regulated sector. However, they are surprisingly little regulated in detail. General information is provided by USP, WHO, and ISPE in particular. Most recently, the EU emphasised the importance of analytical transfer[1] by including it in their draft of the revised GMP guideline. In this position paper, further direction is given in order to facilitate individual transfer projects.

The key to success for method transfers is the excellent communication between sending and receiving unit. In order to facilitate this communication, procedures, flow charts and checklists for responsibilities, success factors, transfer categories, the transfer plan and report, strategies in case of failed transfers, tables with acceptance limits are provided here,

together with a comprehensive glossary. Potential pitfalls are described such that they can be avoided.

In order to assure an efficient and sustainable transfer of analytical procedures, a practically relevant and scientifically sound evaluation with corresponding acceptance criteria is crucial. Various strategies and statistical tools such as significance tests, absolute acceptance criteria, and equivalence tests are thoroughly described and compared in detail giving examples. Significance tests should be avoided. The success criterion is not statistical significance, but rather analytical relevance. Depending on a risk assessment of the analytical procedure in question, statistical equivalence tests are recommended, because they include both, a practically relevant acceptance limit and a direct control of the statistical risks. However, for lower risk procedures, a simple comparison of the transfer performance parameters to absolute limits is also regarded as sufficient.

1 INTRODUCTION

Method transfer is obviously an important aspect in the lifecycle of pharmaceuticals [1, 2] and certainly belongs to the most discussed and complex issues in the GMP regulated sector. It is regularly examined in audits and inspections. The need to properly address the topic of analytical method transfers is well illustrated by the adoption of the WHO guideline [3], the new USP General Information Chapter <1224> [4], and recently the draft of revision of Chapter 6 of the EU GMP Guideline [5].

These guides provide a good general orientation to organise, manage, and document the transfer process of analytical procedures. However, the lack of explicit details (with the exception of the recommendations of the ISPE Guideline [6], which are partly insufficient [7]) has led to a multitude of empirical procedures that differ very much in the validity of their

[1] In order to facilitate readability, the terms “method” and “analytical procedure” are used synonymously. All analytical steps are included, such as sample preparation, analytical methodology, calibration, reportable result,
etc.

results. The implementation of correct and efficient transfer processes is still far from being part of daily laboratory routine. An improved general concept for the implementation in daily laboratory practice is therefore urgently needed.

An analytical method is transferred from a sending unit (SU) to the receiving unit (RU). The sending unit is the laboratory, where the method was originally developed and validated and/or routinely applied. The receiving unit is another laboratory, which is close to an additional production site or a contract laboratory.

The goal of transfer validation is to demonstrate the ability of the RU to perform the relevant analytical procedures successfully. It has to be pointed out that the performance and ability of the sites is always the sum of the ability of the staff and the performance characteristics of their equipment and should not depend on the properties or quality of the samples. The basic aspects are defined by cGMP requirements, for example that the suitability of all employed test methods must be verified under actual conditions of use [8-10]. Nevertheless, general concept and details should be defined in an internal standard operating procedure [11]. General information about aspects of analytical method transfers can be found in [2].

In our position paper we outline how the individual circumstances can be considered best and how a detailed procedure for the individual company can be employed. Support is given by check lists, flow charts and spread sheets, which can be used as such or can be customized to one´s individual requirements.

2 PLANNING OF THE METHOD TRANSFER

2.1 Responsibilities

Transfers of analytical procedures occur in various contexts with differing complexity: a single test method to a contract laboratory or the whole control strategy of

a product as part of manufacturing transfer, inter-company or between companies, from R&D to industrial QC or between industrial productions sites, between two sites or to multiple sites etc.

Depending on this complexity, the formation of a coordinating Analytical Transfer Team (ATT) may facilitate the process. The ATT should be formed by representatives of SU and RU and should include all affected functions, of course analytics, regulatory, quality assurance, production, etc. In case of production transfer, it may be a sub-team of the technology transfer team. The ATT will manage and coordinate all transfer activities, align schedules, solve issues etc. Here, clear responsibilities must be assigned (Table 1) as recommended in the Guides [2-6]. Typically the responsibility of the SU is emphasized to systematically transfer the knowledge related to the methods in question [2, 6].

Table 1: Responsibilities during an analytical transfer

Sending Unit	**Receiving Unit**
- Provide method-specific training if required - Assist in analysis of quality control testing results - Propose a strategy for all methods to be transferred - Propose experimental design, sampling methods and acceptance criteria - Provide any validation reports and demonstrate robustness of methods - Provide details of the equipment used and any standard reference samples - Provide approved procedures used in testing - Execute the transfer protocol	- Ensure that adequately trained and experienced personnel is in place - Review analytical methods provided by the SU - Formally agree on acceptance criteria before executing the transfer protocol - Ensure that the necessary equipment for quality control is available and qualified - Provide an appropriate documentation system - Execute the transfer protocol - Review and approve transfer reports.

- Compile and approve transfer reports	

In particular, the SU is responsible to provide the expertise and experience to the RU and ideally also for the technical training of the RU staff. However, in some context it may be of advantage to appoint RU responsibility for proposing a strategy and/or the protocol, for example, if larger experience with transfers rests with the RU, or if they have the primary interest in the transfer. Depending on the complexity of the transfer, face-to-face meeting(s) will facilitate a smooth transfer, also because all colleagues concerned can get closer acquainted.

2.2 Success Factors

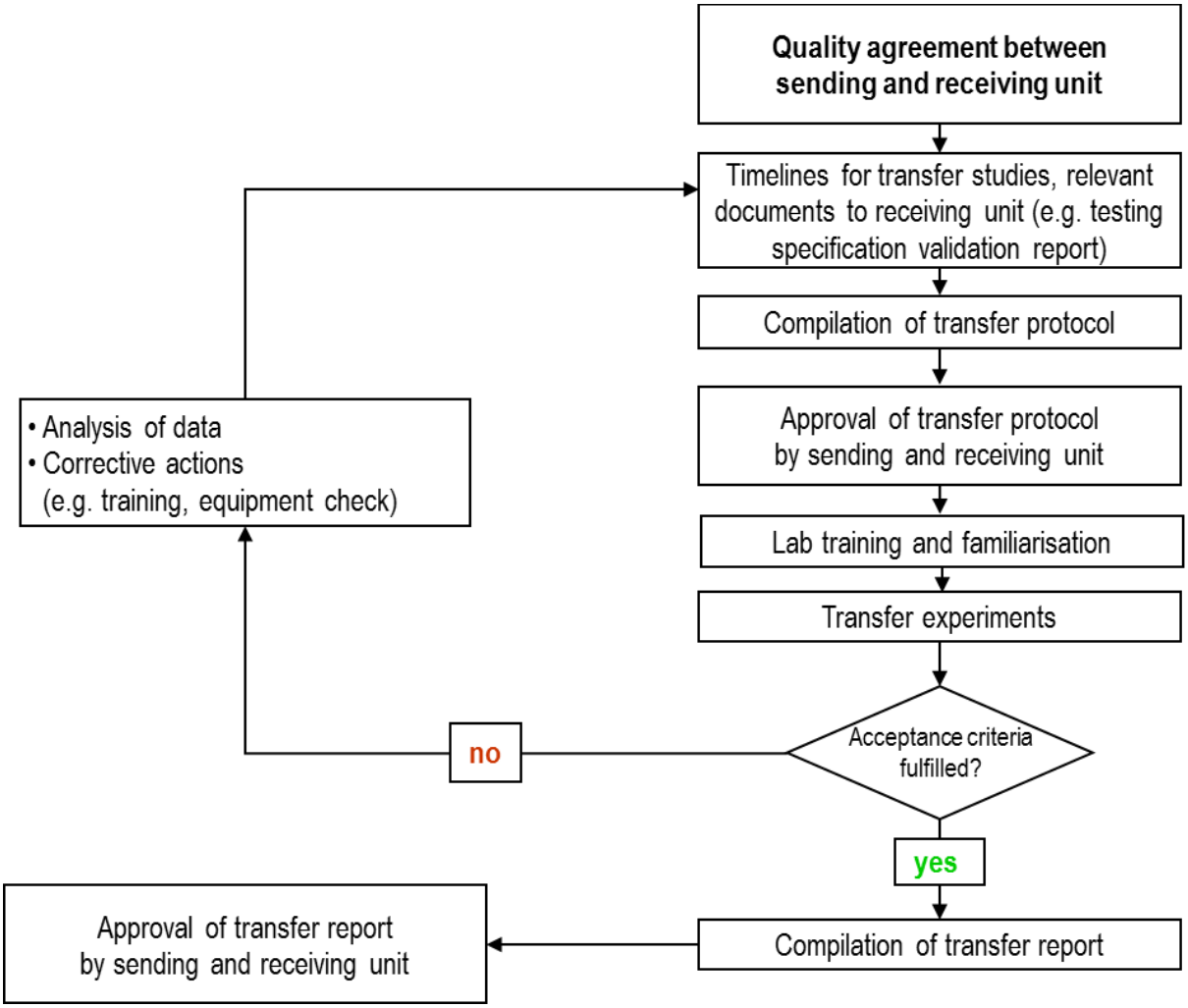
The most important rule for success is to establish an open and reliable communication between both sites. A detailed risk assessment and subsequent consideration of its results concerning definition of transfer scope and strategy as well as training is recommended. The key factors for a successful transfer are summarized in Table 2. In order to avoid any difficulties in the first place, it is a good idea to review potential pitfalls (section 5) right from the start [2].

Table 2: Key factors for success

Key Factors for success:
- Documentation - Information and communication - Risk assessment - Sample handling and storage - Sample preparation - Lab training and experienced staff - Equipment and qualification

- Data evaluation - Procedures for unexpected results or transfer failure

Figure 1: Workflow of method transfer



2.3 Documentation and Knowledge Transfer

In the next step, the SU should provide an up-to-date documentation package including at least the detailed test procedure and its validation, but preferably additional information on routine performance and "behaviour" of the concerned methods. This may include development reports or other knowledge repositories, monitoring of SST-results or other data, control charts, unusual and

OOS-results, information regarding calculation methods (decimal places, average calculation), acceptance criteria and specifications. Stability studies are an excellent source to evaluate the real routine performance of an analytical procedure [2, 12].

2.4 Definition of Transfer Types

After a careful review of the documentation by RU, the Transfer Strategy should be defined by the ATT. The choice of the particular strategy needs justification.

Sometimes, it may be preferable to apply a method by the RU or even to perform some training before deciding about the strategy. The type of transfer is defined for each method based on a risk assessment, taking the complexity and criticality of the analytical procedure and its purpose (e.g. type of material analysed) into account as well as the experience and knowledge of RU. It is essential to address all concerned analytical procedures, in order to ensure a complete documentation. Based on USP<1224>, the types can be differentiated in comparative testing (involving both SU and RU) and "self-qualification" of RU (Table 3) [2].

Table 3: Transfer categories (based on USP<1224 [9])

Category	**Possible design**	**Suitable type of acceptance criteria**	**Examples**
Comparative studies:	Involvement of SU and RU(s)		

- Basic design	1 series with 6 determinations each	Direct comparison)[a]	Less critical methods for API/DP (e.g. water, residual solvents, ions, particle size distribution) Less critical materials: LC for intermediates
- Intermediate design	≥ 2 series each, number of determinations adjusted to number of series	Direct comparison)[a] or equivalence test	Critical or complex methods for API/DP (e.g. LC/GC assay and related substances)
Co-validation:	involvement of RU in method validation usually intermediate design	Dependent on validation characteristics	Critical or complex methods for API/DP
(Re-) Validation	partial or complete method validation by RU according to ICH Q2 for API and drug product methods [16]	= original validation, or tighter	If change is intended or validation status insufficient or no suitable samples available (e.g. cleaning, critical limit tests) Microbiological tests
Verification	demonstration of appropriate performance by RU		
- Comparison with certified result (by SU or reference material)	≥1 analyst, according to test instruction or more determinations	Certified result	Simple methods (e.g. water, loss on drying, …)
- Conformance to SST-criteria or other performance criteria	≥1 analyst, according to SST instruction or more determinations)[b]	SST or defined performance criteria	Compendial methods

Application	by RU, according to control test procedure	acceptance criteria defined in test procedure	Identification tests; compendial standard tests (e.g. sulphated ash, heavy metals, …); limit tests

a) direct comparison of accuracy and precision results with the defined acceptance criteria (point-estimate, see section 4.3)

b) in order to achieve a sufficiently reliable result, e.g. for precision ≥ 6, etc.

2.5 Familiarisation and Training

Before starting any formal transfer exercise, all methods to be transferred by comparative studies should be at least applied at the RU in order to gain experience with the control test as described in the regulatory dossier. This ensures the "RU-readiness", which is essential to maintain regulatory consistency, moreover to understand and address (potential) issues which have to be solved, including equipment, reagents, facilities.

Sample handling and sample preparation are the most critical issues and most common reason for failure of method transfer. It should be verified that the description of the procedures in the testing specifications does reflect all relevant practical aspects of the sample preparation in detail.

Typically the receiving site has less knowledge about the robustness of the procedures. For this reason sample preparation is a hot topic for lab training. For more complex methods, it may be extended to a formal training by SU (as best option) and/or assisted by a video or picture-based documentation prepared by the SU.

Information gained by such "familiarisation" or training may influence the design or even the categorisation of the transfer activities. If the need or wish for changes to the methods is identified, Change Control procedures must be strictly followed and regulatory implications must be evaluated.

2.6 Transfer Samples

Concerning the sample used for transfer, it is important to define the optimum samples for the particular analytical method. Samples for the transfer may include:

- stability- or routine samples
- stressed or spiked samples (purity testing)
- simulated samples (dissolution testing)

As the objective is the successful transfer of the analytical procedure, it is preferable to use one (representative) batch and rather increase the number of determinations than using several batches. Exceptions might be if batch characteristics are known to influence analytical performance, and no “worst-case” batch can be defined, such as an influence of tablet hardness on sample preparation, varying impurity profiles, or particles size distribution.

Data about the relevant properties of samples and standard substances (stability, sensitivity to light / humidity, in particular for biologics) as well as safety precautions (health safety environment) or controlled substances status are very helpful for proper sample handling and provision of the correct transport and storage capabilities [2].

2.7 Transfer Protocol

For transfers of less complex methods, the transfer strategy document can serve as a protocol with established design of the transfer activities and acceptance criteria. Alternatively, separate protocols for each or some analytical procedures can be written, or both approaches can be combined

The transfer (strategy) protocol should be discussed and jointly agreed in the ATT and should include all aspects recommended in the guidelines [2-6].

Design of the experimental studies and acceptance criteria should be defined by a risk assessment, taking the criticality of the concerned material (i.e. API/DP, intermediate, starting material, in-process control) and the criticality and complexity of the test item as well as the experience of the RU into consideration. For higher risks, a formal assessment should be performed [13]. The design can also be influenced or defined based on prior knowledge of the SU, especially during the development phase [14]. The number of determinations should be sufficient to allow a result reliable enough for the given analytical procedure and acceptance criteria (see section 3.4.3.).

2.7.1 Acceptance Criteria

Acceptance criteria should be established to be compatible with the intended use of the method to be transferred. For less complex and less critical methods or materials, a direct comparison of the results with the limits is justified, whereas for more complex and critical applications, statistical equivalence tests are recommended (see section 4.4.). The latter allow a defined decision probability and consequently a direct control of the (consumer's) risk. Statistical significance tests (e.g. t- and F-tests) should be avoided as they do not reflect performance requirements of the intended application (see section 4.2).

The acceptance criteria for direct comparison can be established based on experience (bench-marking) and/or performance requirements derived from the intended use, i.e. specification limits [15]. If more risk control is required, acceptance limits can be established by means of statistical simulations taking the actual performance of the given method into account [2, 14].

2.8 Experimental Investigations

The experimental studies should follow strictly the protocol and any deviation must be documented and evaluated. Failures to meet transfer acceptance criteria must be investigated, properly documented and the root-cause identified. Procedures in case of unexpected results or failure of the method transfer should be defined in advance, e.g. in the transfer protocol or in an SOP.

Repetitions of experimental studies (or of the relevant parts) should only start after appropriate corrective actions have already been implemented. Apart from the transfer type “application” and possibly “comparison with certified result”, the results obtained during transfer studies do not constitute “reportable results” as defined in the control test procedure. Consequently, results outside the release acceptance criteria are formally no OOS results, even if marketed batches were used. Note that during transfer the analytical procedure is not even formally established in the RU. Nevertheless a root cause analysis and proper investigation according to the typical OOS procedure is recommended (see chapter 5) [2].

2.9 Finalization of the method transfer and compilation of transfer report

The transfer report should describe the performed analyses, summarise the results and evaluate the defined parameters with regard to the acceptance criteria. All relevant data will be forwarded to the SU, which compiles the report. Any deviation from the protocol must be described and evaluated. The report must contain a clear conclusion regarding the success of the transfer.

The report should at least contain the following information:

- Unique identifier (title, code, version)
- Indication of the corresponding transfer protocol

- Results of both RU and SU (tabulated)
- Evaluation of the results
- Explicit conclusion
- Root cause analysis in case of failure
    - Description of the proceeding in this case (e.g. additional transfer protocol)
- Attachments: (e.g. raw data, analytical reports, chromatograms, spectra)

After the successful completion of the method transfer, a „post-transfer review“ may take place, where suggestions for increasing the efficiency of the analytical method should be discussed [2].

3 ACCEPTANCE CRITERIA, EVALUATION AND THE USE OF STATISTICAL METHODS

3.1 General considerations

In order to evaluate results from different labs, we need acceptance criteria to distinguish minor and acceptable discrepancies from major ones [2]. Typically the accuracy during the method transfer is monitored by considering the obtained mean values in SU and RU. Further it can be agreed on a comparison of the variability.

There are principally three possibilities to define acceptance criteria (AC) for method transfers:

- absolute limits for differences and variability
- statistical significance tests(e.g. t-test)
- statistical equivalence tests

Other aspects, such as linearity, are usually not part of a method transfer. These aspects are covered by the

successful transfer of the already validated method. An accurate value in the RU implies a valid calibration function, at least at the relevant concentration.

The use of simple statistical significance tests is discouraged [4, 6]. The use of a t-test often causes paradox results, when (favourably) high numbers of experiments or favourably low spread is observed [2]. In addition, only the variance contribution of repeatability (within-series precision) is taken into consideration for the t-test. However, between independent series, as in the case of transfer, usually additional (between-series) variance contributions are present. In these cases the results are more often significant. However, this also happens with very small differences which are not relevant at all [2].

Similar problems can be found during batch-to-batch comparison, within the scope of accuracy testing or recovery rate determination, during the assessment of stability tests and of course in bioequivalence studies. These issues are also described by Hauck et al. in their stimulus paper “Acceptable, Equivalent or Better” [16] and references given therein.

When relevant differences need to be distinguished from irrelevant ones, the t-test is not suitable. The mathematical tool of the equivalence test is a better approach. Thus, equivalence tests are recommended by USP and ISPE to evaluate method transfers.

The implementation of equivalence tests needs a little bit more background. However, as there are tailored spreadsheets for these tests, they are easy to handle and should be used for method transfers, especially if the analytical spread is not (well enough) known a priori (see section 3.4.). If the analytical spread is well known and under control, then the similar but simpler tool of absolute limits is another good alternative, as discussed in section 3.3. A priori-knowledge about the analytical spread is available from experience with techniques and sample pre-treatment scenarios ([17-

18] and references given therein), or can be derived from control charts in the SU.

3.2 Absolute limits

When using absolute limits, then just the mean values of target parameters (e.g. content of a sample investigated during the transfer, precision) are compared to the maximum value acceptable, e.g. 2% difference between the mean contents. This approach is very easy to understand and straightforward. However, it should only be used with a sufficient number of experiments performed in each participating lab. Further, this approach implicitly assumes a certain (i.e. reliably known) analytical variability.

For simple methods, this variability can be obtained from benchmarks, i.e. the typical one for a given analytical (class of) methods [17-21]. However, if a more sophisticated sample pre-treatment is employed, or in case of critical analytical procedures, sufficient validation data become necessary for a proper estimate.

Another approach is to define the target variability from the requirements, i.e. from the specification range available for the analytical variability. Aligned with the definition of method capability or uncertainty, the maximum acceptable standard deviation corresponds to 1/6 or ¼ of this analytical range (1/3 or 1/2 of the one-sided range, i.e. 1/coverage factor), corresponding to 99 or 95% confidence. Based on this estimate of a target (true) standard deviation (TSD), absolute acceptance limits for variability and accuracy can be calculated according to the design of the experimental transfer study [2].

Using the target variability, the (future) distribution of standard deviations can be estimated and an appropriate upper limit can be defined as **precision acceptance limit**. As an approximation, the upper 95% confidence limit can be calculated using the degrees of freedom from the design of the experimental study [15].

$$C_U = \hat{\sigma}_t * \sqrt{\frac{df}{\chi^2(P, df)}} \quad (1)$$

$\hat{\sigma}_t$ = target standard deviation (as an estimate for the true value)

χ^2(P, df) = Chi-square value for the statistical confidence P (usually 95%) and the degrees of freedom df according to the design of the transfer study. Excel®: χ^2 = CHIINV(α, df); α = 1-P

For example, the upper confidence limit for a series of six determinations corresponds to 2.1 times the TSD, for a pooled standard deviation from four series with six determinations each to 1.4 times the TSD [2].

I

The maximum difference between means that may originate from random variability can be calculated (with a simplification for larger number of determinations and 95% probability) according to DIN ISO 5725-3 as 2.8 times the standard deviation of the means. The suitable **accuracy acceptance limit** can then be estimated using the within- and between-series variance contributions obtained in validation studies (or other reliable sources, e.g. from stability studies [12]) and the planned number of repetitive experiments [22].

$$\Delta_{\bar{x}\,\max,95\%} = t(P, df) * \sqrt{2} * s_{\bar{x}} \approx 2.8 * s_{\bar{x}} = 2.8 * \sqrt{\frac{s_b^2}{k} + \frac{s_r^2}{k * n}}$$

with $s_R = \sqrt{s_b^2 + s_r^2}$ (2)

$s_{\bar{x}}; s_r; s_R$ (reliable estimations of) standard deviations of the mean, repeatability, and reproducibility (intermediate precision)

$s_b^2; s_r^2$ between and within series variance

k, n = number of series and determinations per series in the transfer study (assuming the same number in both laboratories)

Sometimes, there may be a lack of reliable precision data Then, the concept of the above described TSD $\hat{\sigma}_t$ can be used, extended by a target ratio f_R between reproducibility and repeatability. Equation 2 can then be rearranged:

$$\Delta_{\bar{x}\,\max,95\%} = 2.8 * \hat{\sigma}_t * \sqrt{\frac{\left(f_R{}^2 - 1\right)}{k} + \frac{1}{k*n}}$$ with

$$f_R = \frac{s_R}{s_r} \qquad (3)$$

The larger the ratio f_R, i.e. the difference between reproducibility and repeatability, the less the random difference between series means can be reduced by increasing the number of repetitions within the series, only by increasing the number of series k.

Using benchmark data for LC assay, the ratio of reproducibility and repeatability standard deviation is found between 1.4 and 3. In LC assay, the reference standard analysis is an essential contributor to the difference between reproducibility and repeatability. The more complex the sample preparation, the smaller will be ratio, reflecting the more dominating effect of the variance contribution of the sample preparation (Table 1.1-5 in [21]). It should be noted, that these acceptance limits will only include the expected random variability, i.e. assuming the absence of any true bias [2].

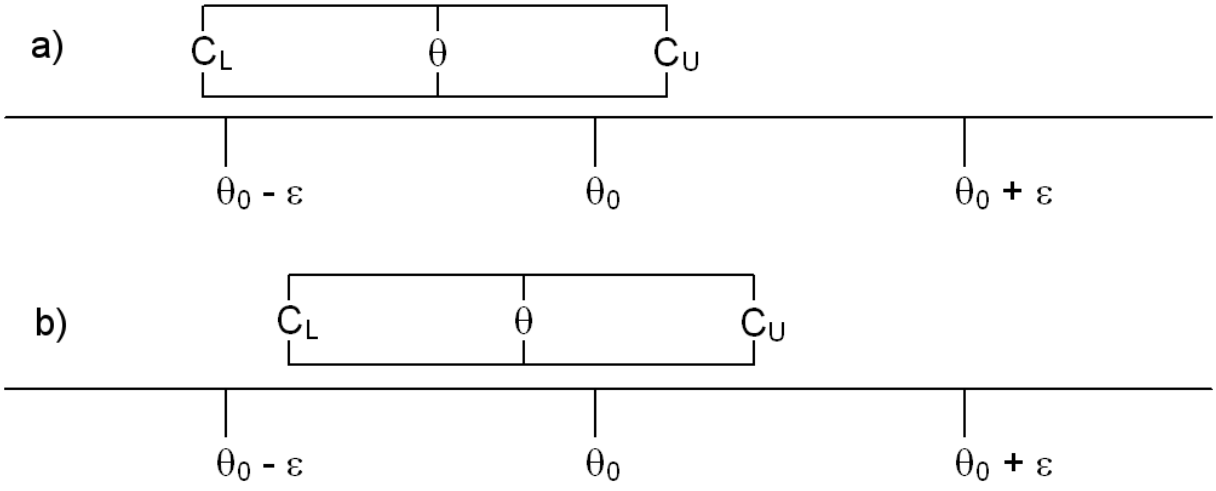
3.3 Equivalence tests

3.3.1 Concepts and calculations

Equivalence tests are applied to decide if an estimate lies within a certain equivalence interval or not [2]. These tests compare the equivalence interval around the nominal or reference value ¸ 0 with the interval around the measured main parameter ¸ . In the case of method transfers, equivalence tests are superior to the classical t-test [2].

Figure 2:

θ is the measured main parameter, θ_0 is the reference value. C_L and C_U are the confidence limits (Eq.s 2 and 3), $\pm\ \varepsilon$ are the acceptance limits (=acceptable deviation). If the confidence interval (C_L d θ d C_U) does not fit completely inside the acceptance interval ($\theta_0 - \varepsilon$ d θ_0 d $\theta_0 + \varepsilon$) non-equivalence is concluded, as the probability to obtain intolerable values smaller (a) than $\theta_0 - \varepsilon$ or larger than $\theta_0 + \varepsilon$ is too high. If the whole confidence interval lies within the acceptance interval (b), equivalence can be concluded and it can be assumed that all measured values can be found inside the acceptance interval $\theta_0 \pm \varepsilon$ with the given error probability ± [32].



This interval concept can now be expanded considering the relevance of a deviation. Essentially, the same confidence interval around ¸ is calculated, but the obtained interval is not compared to the one single value zero but with an interval which is considered as representing acceptable deviations µ, e.g. +- 2% (Fig. 2 a and b). In Fig. 2 a, a part of the confidence interval is outside the interval of relevance. Here there is a considerable probability that the true value is outside the interval of relevance. The possibility of a relevant deviation cannot be neglected. In Fig. 2 b, the confidence interval (CI) lies completely within the interval of relevance (RI). Here the probability of an unacceptable deviation will be very low.

The approach to establish equivalence can be demonstrated most suitably by means of confidence intervals (CIs). For each tested main parameter ¸ ($\mu_1 - \mu_2$, μ_1 / μ_2 or $\hat{\sigma}_2^{\,2} / \hat{\sigma}_1^{\,2}$) a CI is set up. The equivalence hypothesis predicates the equality between ¸ and an appropriate nominal value ¸$_0$. Ideally this nominal value is 0 when testing the difference of mean values ($\mu_1 - \mu_2$). It is ideally 1 when testing the quotient of mean values μ_1 / μ_2 or variances ($\hat{\sigma}_2^{\,2} / \hat{\sigma}_1^{\,2}$).

A symmetrical interval is built for ¸$_0$ with an upper (¸$_0$ + µ) and a lower acceptance limit (AL) (¸$_0$ – µ). This is usually specified by intra-corporate settlements.

A value of 2% has been given as example for an acceptable deviation when comparing mean values during a transfer of a method for quantitation of an API [6]. The following interval is then obtained: [¸$_0$ –2%; ¸$_0$ +2%]

An (1 – 2±) - confidence interval is calculated for ¸ using the test statistics. It is also defined by a lower (C_L) and an upper (C_U) limit. The size of this

interval depends on the measured spread, the available degrees of freedom and the error probability ±.

One can estimate the confidence interval of θ using the classical t-distribution [23, 24].

$$C_{L/U} = 100 \cdot \left(\exp\left(\bar{x}_1 - \bar{x}_2 \mp t_{1-\alpha, n_1+n_2-2} \cdot \hat{\sigma}_\rho \cdot \sqrt{\frac{2}{n}} \right) - 1 \right) \quad (4)$$

The value $t_{\alpha,(2n-2)}$ is chosen as above, the pooled standard deviation $\hat{\sigma}_p$ and the mean values are calculated for logarithmically transformed values and the square root term originates from $\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ for equal n_1 and n_2 (n_1=n_2=n) [2].

Using Eq. 4, all values are first log-transformed, then mean values and SDs are calculated. These values are used to calculate confidence intervals, and finally the limits are retransformed using the exponential function as inverse function to the log-transformation, in order to obtain the confidence limits in the usual scale ([25], sheet Ex1 Series Equiv. Test, cells I23:K32). Methods with exponential functions in their acceptance limits are obviously based on this approach. The log-transformation often leads to normally distributed error probabilities although the original data was not normally distributed; this is the reason why it is often used.

Again, the acceptance limits must be outside this confidence interval CI, or in other words, the whole confidence interval must be embedded into the interval of the acceptance limits (equivalence interval; EI). Both calculation methods provide similar results (given for various scenarios at http://www.pharmchem.tu-bs.de/forschung/waetzig/support/ = [25].

There are different ways to correctly calculate the confidence intervals needed for equivalence tests. These just slightly differ in the assumptions [2]. The obtained results are numerically different, but these differences are not relevant. This is indicated by the close proximity of the obtained values (compare various spreadsheets in [25]). However, in a highly regulated environment such as pharmaceutical QC, one needs unequivocal SOPs and results. Thus one has to decide for one particular way of calculation. We recommend the one given in [23, 26, 27], (Eq. 4), because these are the best referenced and most thoroughly discussed ones in the literature. They are also mainly used in the spreadsheets given at [25].

If the method transfer is performed employing two or more series in each lab, in some cases no (relevant) bias between the series within the same lab will be observed. In this cases all series in lab can be combined to one series for each lab and the procedure above can be directly applied using the spreadsheets for the case of homogeneous variances ([25], “Ex1 to “Ex3 Schuirmann”, 23]).Equivalence of the SDs can also be investigated using equivalence tests. We recommend [28, 29] and USP <1010> [30] for orientation.

However, in many cases a relevant difference between the two series within the same lab is found. The way to calculate the confidence limits is the same, but it could be necessary to treat the data for each series separately [31]. Then the difference between the labs is

$$diff = \frac{\sum_{i=1}^{k} \bar{x}_i}{k} - \frac{\sum_{i=1}^{k} \bar{y}_i}{k} \qquad (5)$$

Where k – number of series in lab, constant for each lab; and $\bar{x}_i$and $\bar{y}_i$-

means from lab 1 and lab 2 respectively.

The standard deviation is calculated as usual, just for more series

$$s = \sqrt{\sum_{i=1}^{k}\left(\frac{s_{x,i}^2}{k^2 \cdot n_i} + \frac{s_{y,i}^2}{k^2 \cdot n_i}\right)} \qquad (6)$$

It is important to properly calculate the degrees of freedom for the t-value in this case. We use here the best known Welch-Satterthwaite procedure

$$\vartheta = \frac{\left(\sum_{i=1}^{k}\left(\frac{s_{x,i}^2}{k^2 \cdot n_i} + \frac{s_{y,i}^2}{k^2 \cdot n_i}\right)\right)^2}{\sum_{i=1}^{k}\left(\frac{s_{x,i}^4}{k^4 \cdot n_i^2 \cdot (n_i - 1)}\right) + \sum_{i=1}^{k}\left(\frac{s_{y,i}^4}{k^4 \cdot n_i^2 \cdot (n_i - 1)}\right)} \qquad (7)$$

An example of this calculation, for the case of two series in each lab, is given in [25] in the sheets “Ex1 to “Ex3 Series Equiv.Test Welch”, (F24, C24, C25) and (K24, N24, K25) for the logarithmically transformed data.

We are aware that there are several approaches to estimate the overall variance and the degrees of freedom, when the variances of the two or more populations, based on independent samples, are not assumed to be equal. These estimations are related to the well-known Behrens-Fisher problem. This has not yet been solved comprehensively for all scenarios. In this work, we follow the best referenced approach obtained from the statistical literature [31].

In equivalence testing the ± error corresponds to the more important risk of accepting an unsuccessful method transfer. The ² error (also known as type II error) stands for the less important risk of rejecting a successful method transfer and repeating it. The acceptance probability 1-² (power) and ² complement one another to 100%. Equivalence tests are designed in a way that it becomes very unlikely that an unacceptable method transfer is wrongly accepted. Thus the effective error probability ± becomes low. However, since ± and ² error always

complement each other, the experimental design must care for low errors of both types. Else equivalence tests will prevent ± at the price of high ² errors. The error to wrongly accept an unacceptable transfer (± error) is more critical, but the unnecessary rejection of an acceptable transfer is also unfavorable since it can cost a lot of money and resources. Thus ISPE proposed a concept which includes an experimental design. This concept was generalized for higher analytical spread and for equivalence intervals (EIs) other than +-2 [28, 32, 33], in order to provide a general strategy to perform equivalence tests. This approach is described in section 3.4.3. In step 3 therein an estimate for the overall experimental error is required.

This overall error during the method transfer can be estimated according to the law of error propagation [32]. Therefore typical error values for HPLC system suitability (0.3% RSD%) and uncomplicated sample preparation (0.6% RSD% repeatability) have been assumed [20].

The variation between independent series typically dominates the overall error. Thus this variation needs to be especially considered, in order to successfully apply this concept. However, this variability is usually unknown. Perhaps it is small for well-described and robust methods, which are performed by experienced personnel. For a case like this, an overall error $\hat{\sigma}_{\bar{x}}$ of 0.37 has been estimated for a typical HPLC setting [32].

However, it is difficult to determine the value reliably for a particular case as this determination itself requires a high number of data. Therefore, usually just estimations will be available for $\hat{\sigma}_{AN}$ and the derived standard error of the mean between laboratories $\hat{\sigma}_{\bar{x}}$ Data from earlier method transfers can be valuable sources for these estimations.

The difference and the advantage of the equivalence test over a classic two-sample t-test is made clear when Figures 4 and 5 of [2] are compared.

Using an acceptance interval to compare with the confidence intervals makes much more sense. Additional supporting material which visualizes the properties of equivalence tests is available at [25]. Guidance to select acceptance criteria and to perform the corresponding equivalence test has also been provided. This approach can readily be customized to one´s own method transfers [2]. In order to do so, first suitable acceptance limits are chosen [2, 6, 32]. Next, the expected variability should be estimated using long-term experience [2, 12, 18]. Then it is considered, if absolute limits are suitable (3.2, Eq. 3; [2]); see EXCEL-File [25]. If yes, this simpler approach is recommended. If no, next a suitable experimental design for an equivalence test is needed [2]. Often a design with two analysts in each lab (e.g. each performing one series of 6 experiments) is sufficient. This design corresponds to UV spectrometric or HPLC-UV methods, or similarly performing ones, with straightforward sample pre-treatment.

After choosing the experimental design, the required experiments can be performed at the SU and RU and then be subsequently evaluated [2, 25]. If the calculated confidence interval lies completely within the acceptance limits, then the method transfer is successfully completed. For more details on this procedure, please refer to [2].

4 AVOIDING POTENTIAL PITFALLS AND MISTAKES

As mentioned in the foregoing chapters the exchange of the relevant method information and subsequent compilation of the transfer protocol avoid a lot of foreseeable problems during method transfers. Nevertheless, potential mistakes should be reviewed [2]. These include

- The calculation of the results
  - Calibration standards and correction factors
  - Rounding
  - integration parameters (e.g. minimum area, threshold, noise, data filtering or “smoothing).
  - reporting limits, summation or averaging procedures
- availability of reagents, samples and standard material
- correct shipment and storage
- equivalent equipment
  - qualification, procedure and acceptance criteria
  - materials, carry-over properties
  - modules equivalency, degree of automation
  - temperature ranges
  - batch-to-batch variability of e.g. column material
  - including equipment for sample pre-treatment and cleaning
    - e.g. properties of ultrasonic baths, centrifuges, filter material etc.

Often marginal discrepancies concerning the sample preparations are the source of systematic discrepancies during a lab-to-lab transfer. Identification of such discrepancies often fails because of a lack of detailed information in the testing procedures. For example different homogenization procedure of tablets (crushing, trituration, milling) could cause variable assay results. Even different cleaning procedures of e.g. glassware can have an effect. Communication between SU and RU is the key to the success of an analytical method transfer. In particular this is true during a failure investigation, and

this is not always easy considering the long distances involved. On the other hand, the use of electronic media facilitates global communication. Not only video conferences can well substitute face-to-face meetings. Video files which demonstrate the procedures in use can easily be shared, large files can be exchanged using dropboxes or even Youtube, as long as the information is not confidential.

As mentioned in chapter 2.8 procedures in case of unexpected results or transfer failure should be defined in advance, e.g. in the transfer protocol or in an SOP [2]. Omitting definition of the procedure often results in a critical delay of the method transfer. A suggestion for a structured approach is described in the Scheme below [2].

Scheme [2]:

Strategy for method transfer failure: formal execution as OOE

(Out of Expectation: result not in accordance with the expectation, e.g. violation of an internal warning limit, statistical parameter or unplausible results)

Acceptance criteria are not met by either SU or RU.

Root cause analysis (investigation strategy according OOS procedure should be followed) will be performed.

A: lab error could be identified or made likely

SU compiles revised transfer protocol, if relevant.

Corrective and preventive actions (CAPA) should take place (e.g. lab training).

The corresponding unit performs the repetition of the transfer investigations.

SU compiles or revises (if relevant) the report covering an overall assessment of the transfer, initial data will be invalidated, approval and signatures of both sites.

B: lab error could not be identified

B1: Acceptance criteria ’ incorrect

SU defines modified acceptance criteria.

Detailed justification of the new acceptance criteria will be compiled in a revised transfer protocol (SU). The transfer protocol contains at least the following information:

- Initial results of the method transfer
- initial and modified acceptance criteria
- reasons for modification

Evaluation of initial data against new acceptance criteria,

initial data will not be invalidated, additional experiments are not necessary.

B2: acceptance criteria ’ correct, experimental design of the transfer (statistical power) likely to be not sufficient

Note: An increase in the number of determinations will always increase the probability to obtain the true parameters.

SU adjusts transfer design and compiles a revised transfer protocol.

Repetition of the transfer at SU and RU. All data will be used for evaluation.

The SU compiles a report covering an overall assessment of the transfer, approval and signatures of both sites.

B3: acceptance criteria ’ correct, expansion of failure investigation

According the responsibilities of the SU and RU the root cause analysis will be extended to:

- storage and transportation of transfer samples

- sample drawing of transfer samples

- manufacturing of transfer samples

After identification of an error SU defines new transfer samples

Transfer will be repeated using new transfer samples

SU compiles or revises if relevant the report covering an overall assessment of the transfer, initial data will be invalidated, approval and signatures of both sites.

5 CLOSING REMARKS

For evaluating analytical method transfers, the equivalence test is the approach of choice. Even though its theoretical framework is challenging, its use and interpretation is made straightforward by a clear routine procedure (section 3.4.3) and elaborated examples of calculations by use of typical lab data and a spreadsheet. The more simple approach of classical t-tests is not suitable to evaluate method transfers due to paradox results which frequently occur. However, the comparison of the difference of the lab mean values to an absolute limit is often an alternative.

As outlined in chapters 3 and 5, a successful transfer requires the exchange and the agreement about a good strategy and very many technical and analytical details. In order to take them all into consideration, several checklists have been provided here. Successful agreements about details also require trusting collaboration and good communication. If you can take care of these, you will be successful in your transfer activities.

GLOSSARY

±	error probability
AI	acceptance interval, see relevance interval
AL	acceptance limit
%AL	scaled standard error of the mean between laboratories
API	active pharmaceutical ingredient
ATT	analytical transfer team
CI	confidence interval
C_L	lower confidence limit
C_U	upper confidence limit
df, Ñ	degrees of freedom
DP	drug product
µ	acceptable deviation
EMA	European Medicines Agency
H_0	null hypothesis
H_1	alternative hypothesis
ISPE	International Society of Pharmaceutical Engineering
µ	true mean value
n_i	number of data of the i-th data set
RI	relevance interval
RSD%	percent relative standard deviation
RU	receiving unit, also called routine unit (site/laboratory)
SD, $\hat{\sigma}$	standard deviation
$\hat{\sigma}_i$	SD of the i-th data set
$\hat{\sigma}_p$	pooled SD

$\hat{\sigma}_{\bar{x}}$	standard error of the mean between laboratories
SOP	standard operating procedure
SU	sending unit, also called developing/reference/originating unit/lab/site
TSD	target standard deviation
T_t	statistic of a t-test
t	tabled values of t-distribution
¸	measured parameter (equivalence test)
¸ 0	reference value (equivalence test)
USP	United States Pharmacopeia
WHO	World Health Organisation
$\bar{x}_i$	mean value of the i-th data set

6 REFERENCES

1 P. Nethercote, J. Ermer: Quality by Design for Analytical Methods: Implications for Method Validation and Transfer. Pharmaceutical Technology 36 (10) (2012) 74-79

2 J. Ermer, M. Limberger, K. Lis, H. Wätzig, The transfer of analytical procedures, Journal of Pharmaceutical and Biomedical Analysis 85 (2013) 262–2

3 WHO Technical Report Series, No. 961, 2011, Annex 7

4 United States Pharmacopeia (USP), General Information Chapter<1224>: Transfer of Analytical Procedures, Rockville

5 EudraLex, Volume 4, EU Guidelines for Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use, Part 1, Chapter 6: Quality Control, Feb. 2013

6 ISPE: International Society for Pharmaceutical Engineering, Technology transfer, ISPE Guide, (2003) 23 – 34, www.ispe.org.

7 U. Schepers, H. Wätzig, Application of the equivalence test according to a concept for analytical method transfers from the International Society for Pharmaceutical Engineering (ISPE), J. Pharm. Biomed. Anal. 39 (2005) 310–314.

8 21CFR §211.165, Code of Federal Regulations, U.S. Food and Drug Administration, http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/cfrsearch.cfm?cfrpart=211

9 21CFR §211.194, Code of Federal Regulations, U.S. Food and Drug Administration

10 EudraLex, EU Guidelines for Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use, Vol. 4, Part II

11 PHARMACEUTICAL INSPECTION CONVENTION(PIC/S): Aide-Memoire: Inspection of pharmaceutical quality control laboratories, PI 023–2 (Sept. 2007)

12 J. Ermer et al.: Precision from drug stability studies. Investigation of reliable repeatability and intermediate precision of HPLC assay procedures. J. Pharm. Biomed. Anal. 38/4 (2005) 653-663

13 ICH Q9: Quality Risk Management (2005)

14 C. Agut, A. Caron, C. Giordano, D. Hoffman, A. Ségalini: Transfer of analytical procedures: A panel of strategies selected for risk management, with emphasis on an integrated equivalence-based comparative testing approach. J.Pharm.Biomed.Anal. 56 (2011) 293– 303

15 M. Broughton, J. Ermer: Chapter 7: Transfer of Analytical Procedures in: J. Ermer, J.H.McB. Miller (Eds.) Method Validation in Pharmaceutical Analysis. A Guide to Best Practice. Wiley VCH, Weinheim 2005, pp. 281-300

16 W. W. Hauck et al., Acceptable, Equivalent or Better: Approaches for Alternatives to Official Compendial Procedures, Pharm. Forum 35(3) (2009) 772 – 778.

17 B. Renger, System performance and variability of chromatographic techniques used in pharmaceutical quality control, J. Chromatogr. B 745 (2000) 167-176.

18 H. Wätzig, J. Ermer, Statistik in der Pharmazeutischen Analytik, Teil 2: Unsicherheit über die Unsicherheit, PZ Prisma, 11(4) (2003) 257-265.

19 A. G. J. Daas, J. H. McB. Miller, Content Limits in the European Pharmacopoeia (Part 2), Pharmeuropa 10(1), (1998) 137-146.

20 U. Schepers, J. Ermer, L. Preu, H. Wätzig, Wide concentration range investigation of recovery, precision and error structure in HPLC, J. Chromatogr. B, 810 (2004) 111-118.

21 J. Ermer: Chapter 2.1 Precision; in: J. Ermer, J. H. McB. Miller (Eds.) Method Validation in Pharmaceutical Analysis. A Guide to Best Practice. Wiley VCH, Weinheim 2005, pp. 21-51.

22 Amtl. Sammlung § 35 LMBG Statistik: Planung und statistische Auswertung von Ringversuchen zur Methodenvalidierung, Mai 2003

23 Schuirmann, D.J. 1987. "A Comparison of the Two One-Sided Tests Procedure and the Power Approach for Assessing the Equivalence of Average Bioavailability." J. Pharm. Biopharm., 15: 657-680.

24 *FDA, "Guidance for Industry: Bioequivalence Guidance" (Nov. 18, 2006).* http://www.fda.gov/AnimalVeterinary/GuidanceComplianceEnforcement/GuidanceforIndustry/ucm123816.htm

25 http://www.pharmchem.tu-bs.de/forschung/waetzig/support/

26 W. W. Hauck and S. Anderson. A new statistical procedure for testing equivalence in two-group comparative bioavailability trials. *J. Pharmacokin. Biopharm.* 12:83-91 (1984).

27 J. .Balthasar. Bioequivalence and Bioequivalency Testing. American Journal of Pharmaceutical Education. Vol. 63: 194-198 (1999)

28 U. Schepers, Dissertation TU Braunschweig 2006. Statistische Beurteilung der Güte von analytischen Ergebnissen, http://www.digibib.tu-bs.de/?docid=00007347.

29 U. Schepers, H. Wätzig, Application of the equivalence test for analytical method transfers: testing precision using the United States Pharmacopoeia concept <10 10>, J. Pharm. Biomed. Anal. 41 (2006) 290–292.

30 United States Pharmacopeia (USP), General Information Chapter<1010>: Interpretation and treatment of analytical data, Rockville

31 B. L. Welch, The generalization of ´Student´s´ problem when several different population variances are involved, Biometrica 34 (1947) 28-35.

32 L. Kaminski, U. Schepers, H. Wätzig, Analytical method transfer using equivalence tests with reasonable acceptance criteria and appropriate effort: extension of the ISPE concept, J. Pharm. Biomed. Anal., 53 (2010) 1124–1129

33 U. Schepers, H. Wätzig, Statistische Signifikanz und Relevanz: Klassische t- und F-tests oder Äquivalenztest, PZ Prisma 15 (2008), pp. 187–196.