Journal of Pharmaceutical and Biomedical Analysis 53 (2010) 1124–1129

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

journal homepage: www.elsevier.com/locate/jpba

Review

Analytical method transfer using equivalence tests with reasonable acceptance criteria and appropriate effort: Extension of the ISPE concept

L. Kaminski [1], U. Schepers [1], H. Wätzig*

Technical University of Braunschweig, Institute of Pharmaceutical Chemistry, Beethovenstraße 55, 38106 Braunschweig, Germany

ARTICLE INFO

Article history:
Received 17 February 2010
Received in revised form 23 April 2010
Accepted 26 April 2010
Available online 4 May 2010

Keywords:
Equivalence test
Method transfer
Acceptance criteria
ISPE

ABSTRACT

A method development process is commonly finalized by a method transfer from the developing to the routine laboratory. Statistical tests are performed in order to survey if a transfer succeeded or failed. However, using the classic two-sample *t*-test can lead to misjudgments and unsatisfying transfer results due to its test characteristics. Therefore the International Society of Pharmaceutical Engineering (ISPE) employed a fixed method transfer design using equivalence tests in their Guide for Technology Transfer. Although it was well received by analytical laboratories worldwide this fixed design can easily bring about high β-errors (rejection of successful transfers) or high workload (many analysts employed during transfer) if $\hat{\sigma}_{AN}$ (error due to different analysts) exceeds 0.6%. Hence this work introduces an extended concept which will help to circumvent this disadvantage by providing guidance to select a personalized and more appropriate experimental design. First of all it demonstrates that former *t*-test related acceptance criteria can be scaled by a factor of 1.15, which allows for a broader tolerance without a loss of decision certainty. Furthermore a decision guidance to choose the proper number of analysts or series at given percentage acceptance limits (%AL) is presented.

Contents

Abbreviations: Abbreviations: AAPS, American Association of Pharmaceutical Scientists; AL, acceptance limit(s) (acceptable bias); %AL, percentage acceptance limit(s); API, active pharmaceutical ingredient; C_L, lower limit of a confidence interval; C_U, upper limit of a confidence interval; ε, limit of an acceptance interval; FDA, Food and Drug Administration; H_0, null hypothesis; H_1, alternative hypothesis; ISPE, International Society of Pharmaceutical Engineering; $\hat{\sigma}_{\bar{x}}$, total error; $\hat{\sigma}_{AN}$, error due to different analysts; RSD%, relative standard deviation; TAP, transfer of analytical procedures (method transfer); USP, United States Pharmacopeia.

* Corresponding author. Tel.: +49 5313912764.
E-mail address: h.waetzig@tu-bs.de (H. Wätzig).

[1] These authors equally contributed to this article.

0731-7085/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2010.04.034

1. Introduction

1.1. Analytical method transfer

The transfer of an analytical method from a laboratory, where it was originally developed and validated, to another laboratory, which is close to an additional production site, as well as the transfer for outsourcing purposes became an important issue during the life cycle of a product [1].

The appropriate organization of analytical method transfer is an essential part of the quality assurance system, when pharmaceuticals are produced and analyzed at different sites. As the cGMP requirement 21CFR §211.165 states: the accuracy, sensitivity, specificity, and reproducibility of test methods employed by the company shall be established and documented [2]. 21CFR §211.194 requires, that "The suitability of all testing methods shall be verified under actual conditions of use" [3]. This can only be provided by on-site validated methods or by successful method transfers. Of course the importance of this topic had already been recognized and also taken on by the forum of the United States Pharmacopeia (USP). In a recently published stimulus paper [18] different alternative types of the transfer of analytical procedures (TAP) were distinguished.

Although there is no particular guideline how to perform a method transfer, several regulatory authorities including FDA, MCA, HPB and AAPS collaborated with the International Society of Pharmaceutical Engineering (ISPE) in order to publish a guide on technology transfer with regard to analytical methods [4]. This ISPE guide suggests acceptance criteria and experimental designs. In addition, it provides very useful checklists and templates for transfer protocols. Thus the ISPE guide offers a sound basis for own method transfer activities, which should be further developed to own detailed concepts. Further valuable sources of information are included in [1,5–7].

1.2. Why to use equivalence tests

Let us start with a familiar scenario: a method for quantitative analysis shall verify that the concentration of an API (active pharmaceutical ingredient) lies within the specified limits of 95–105 percent of the nominal value. The relative standard deviation (RSD%) is 0.2% in the developing laboratory. Once the method is developed and established in routine, 0.3% RSD can be achieved. Hence, one can say that both laboratories yield data of very high quality.

Subsequently the method is being transferred from the developing (=sending) to the routine (=receiving) laboratory. The same sample is measured in both laboratories. The following concentration levels are obtained: 100.1% and 99.6%.

According to common sense, one would assess this method as excellent and the method transfer as great success. However, the application of a t-test based on $n = 18$ measurements for each lab shows significant differences of the concentration levels. Thus the method transfer would be considered as a failure! Nevertheless, refusing this method transfer would be a misjudgment because the results provided by both labs deviate only slightly from 100%. In fact, both laboratories are perfectly suitable to control the range between 95 and 105 percent. Obviously the classic two-sided t-test is not perfectly suitable to assess a method transfer. This issue is also described by Hauck et al. in their stimulus paper "Acceptable, Equivalent or Better" [19] and references given therein. 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.

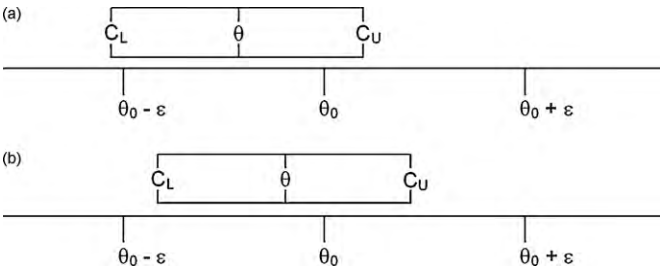

Fig. 1. θ is the measured main parameter, θ_0 is the reference value. C_L and C_U are the confidence limits, $\pm\varepsilon$ are the acceptance limits (acceptable bias). If the confidence interval ($C_L \le \theta \le C_U$) does not fit completely inside the acceptance interval ($\theta_0 - \varepsilon \le \theta_0 \le \theta_0 + \varepsilon$) non-equivalence is concluded, as the probability to obtain intolerable values smaller than $\theta_0 - \varepsilon$ (a) 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 α.

1.2.1. The concept of equivalence tests

Equivalence tests, with a given error probability α, are applied to decide if an estimate lies within a certain equivalence interval or not. Primarily they were developed by Westlake [8] and were used for bioequivalence studies [9]. They compare the equivalence interval around the nominal or reference value θ_0 with the interval around the measured main parameter θ. The classic t-test compares only the nominal value θ_0 with the interval around θ [Fig. 1(a) and (b)].

The approach to establish equivalence can be demonstrated most suitably by means of confidence intervals:

For each tested main parameter θ ($\mu_1 - \mu_2$, μ_1/μ_2 or $\hat{\sigma}_2^2/\hat{\sigma}_1^2$) a confidence interval is set up. The equivalence hypothesis predicates the equality between θ and an appropriate nominal value θ_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 variances ($\hat{\sigma}_2^2/\hat{\sigma}_1^2$).

A symmetrical interval is build for θ_0 with an upper ($\theta_0 + \varepsilon$) and a lower acceptance limit (AL) ($\theta_0 - \varepsilon$). Either regulatory authorities or intra-corporate settlements specify this acceptance tolerance (bias).

2% is a general accepted bias for the comparison of mean values when a method for quantitative analysis, regarding the quantitation of an API, is transferred [4]. The following interval is then obtained: [$\theta_0 - 2\%$; $\theta_0 + 2\%$].

An $(1 - 2\alpha)$ – confidence interval is calculated for θ 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 α.

The limits of the confidence interval are compared to the acceptance limits to verify the hypotheses. The null hypothesis (H_0) is accepted either if $C_L < \theta_0 - \varepsilon$ or if $C_U > \theta_0 + \varepsilon$ is true. In this case, the confidence interval lies partially or completely beyond the acceptance interval [Fig. 1(a)].

Null and alternative hypotheses are:

$$H_0 : \theta \le \theta_0 - \varepsilon \text{ or } \theta \ge \theta_0 + \varepsilon$$

$$H_1 : \theta_0 - \varepsilon < \theta < \theta_0 + \varepsilon$$

The null hypothesis is nullified if the whole confidence interval is included within the acceptance interval. $C_L > \theta_0 - \varepsilon$ and $C_U < \theta_0 + \varepsilon$ must be true [10]. In this case the alternative hypothesis (H_1) is accepted and hence equivalence can be concluded [Fig. 1(b)].

The confidence interval of the equivalence test is consisting of a lower limit C_L and an upper limit C_U. Each limit is related to an

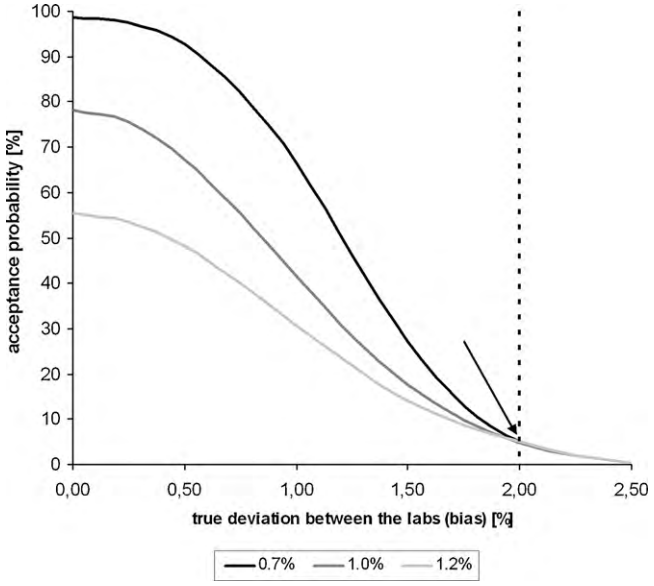

Fig. 2. Acceptance probability of the equivalence test with an acceptance limit of 2% (dashed line) subjected to the true deviation (true bias) between the laboratories (0–2%). The three different curves exemplify transfer situations, which differ only in precision (RSD%) of the executive laboratories (0.7%, 1.0% und 1.2%). An equivalence test can reliably control the risk of accepting an intolerable method transfer. Even if the inter laboratorial deviation is as high as 2% the risk is being controlled with a probability of 5% (arrow).

one-sided test and it is calculated according to Eq. (1) [11,12]:

$$C_U = 100\left[\left(\frac{\bar{x}_1}{\bar{x}_2}\right)e^{+(t_{\alpha,(2n-2)}\hat{\sigma})} - 1\right] \quad C_L = 100\left[\left(\frac{\bar{x}_1}{\bar{x}_2}\right)e^{-(t_{\alpha,(2n-2)}\hat{\sigma})} - 1\right] \tag{1}$$

with:

$$\hat{\sigma} = \sqrt{\frac{1}{2n}(\hat{\sigma}_1^2 + \hat{\sigma}_2^2)\left(\frac{1}{\bar{x}_1^2} + \frac{1}{\bar{x}_2^2}\right)} \tag{2}$$

whereas $\bar{x}_1$ and $\bar{x}_2$ represent the laboratory means $\hat{\sigma}_1$ and $\hat{\sigma}_2$ the corresponding standard deviations. Further, the value $t_{\alpha,(2n-2)}$ is derived from the *t*-distribution with n degrees of freedom and the one-sided error probability α.

1.2.2. Advantages of equivalence tests over a two-sample t-test

Is the difference between two mean values larger than a negligible ε? This question seems to be appropriate in the context of analytical method transfer. An equivalence test can respond to this question if the right hypotheses (H_0: $\mu_1/\mu_2 \leq \theta_0 - \varepsilon \vee \mu_1/\mu_2 \geq \theta_0 + \varepsilon$ and H_1: $\theta_0 - \varepsilon < \mu_1/\mu_2 < \theta_0 + \varepsilon$) are set up.

The difference and the advantage of the equivalence test over a classic two-sample *t*-test become clear again when Figs. 2 and 3 are compared. The two-sided *t*-test controls the risk of rejecting a successful method transfer (α error = type I error) with 5% at a bias of zero (arrow in Fig. 3). The equivalence test also controls the α error with 5% but at the acceptance limit of 2% (bias = 2%) (arrow in Fig. 2). It also becomes obvious that the equivalence test controls the α error independently of particular method precision. Furthermore, due to the changed hypotheses (see above), in equivalence testing the α error represents the more important risk of accepting an unsuccessful method transfer. This means, that an insufficient method transfer is accepted by mistake if the null hypothesis is wrongly rejected. On the other hand, the β 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 $1-\beta$ (power) and β complement one another to 100% within the acceptance limits.

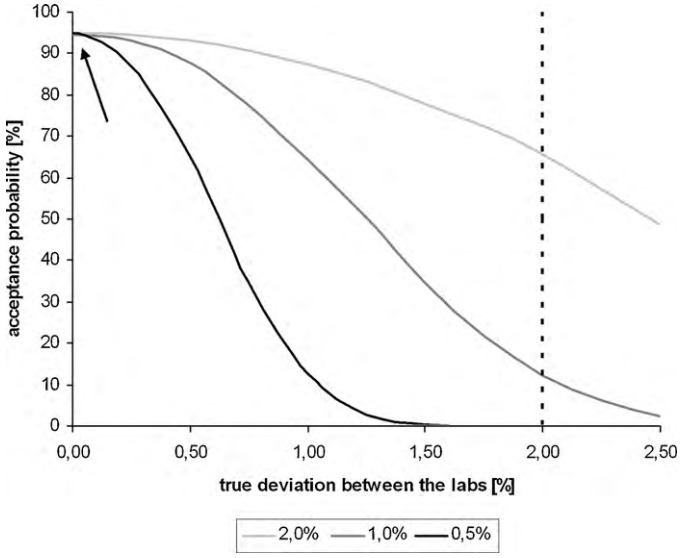

Fig. 3. Acceptance probability of two-sample *t*-tests subjected to the true deviation between the laboratories. The RSD% of these three populations was respectively 0.5%, 1.0% and 2.0%. Each sample size was $n = 6$. The dashed line marks the acceptance limit. It becomes clear, that the *t*-test rewards imprecise working.

The intra-laboratory precision influences the power. When precision increases in equivalence tests, the power increases as well. Let us assume a true deviation between the laboratories of 0.5%. In this case, the power is 68% for an intra-laboratorial precision of 1.0%. In the very same case, the power increases to 93% when the precision is 0.7%. Therefore, equivalence tests reward precise working [16].

Additional supporting material which visualizes the properties of equivalence tests is available at www.pharmchem.tu-bs.de/forschung/waetzig/support/. Furthermore an excel sheet with some of the calculations given here is also provided. This sheet can be easily customized to own scenarios. Equivalence tests to compare precision between labs have been discussed as well [17].

1.3. The ISPE concept: vision and limitation

The ISPE Guide for Technology Transfer is the best generally accepted fundamental text about analytical method transfer. As shown in [12] and [13], this requires either high measurement precision or very extensive and time-consuming experimental designs. This can be exemplified by means of the following scenarios.

In Scenario 1 the ISPE method transfer design [4] where "at least two analysts at each laboratory independently analyze three sample lots in triplicate; resulting in three distinct executions of the method" is applied. Let us assume that in these fictional laboratories a total RSD% of 0.37% is not exceeded. In this case the ISPE design works perfect as the power ($1-\beta$) is 89% (see Fig. 3 in [12] or (identical) Fig. 4 2-2 in [14]).

Imagine a second scenario which is equal to the previous one except for the total RSD%. Here we have 0.62%, which of course still is more than acceptable. However, for the classic ISPE design (2*3*3) a power of only 50% is obtained. Increasing the number of analysts is the only possibility to increase the power. With three analysts the power will raise to 88% in this scenario.

As a kind of rule of thumb Fig. 3 in [12] and 4.2-2 in [14] reveal that an increase of the total RSD% by 0.2% will require at least one additional analyst to maintain an acceptable power level.

The total error $\hat{\sigma}_{\bar{x}}$ is governed by the error due to different analysts $\hat{\sigma}_{AN}$ doing the same analyses. Unfortunately, this error component is often not exactly known. Thus it often remains

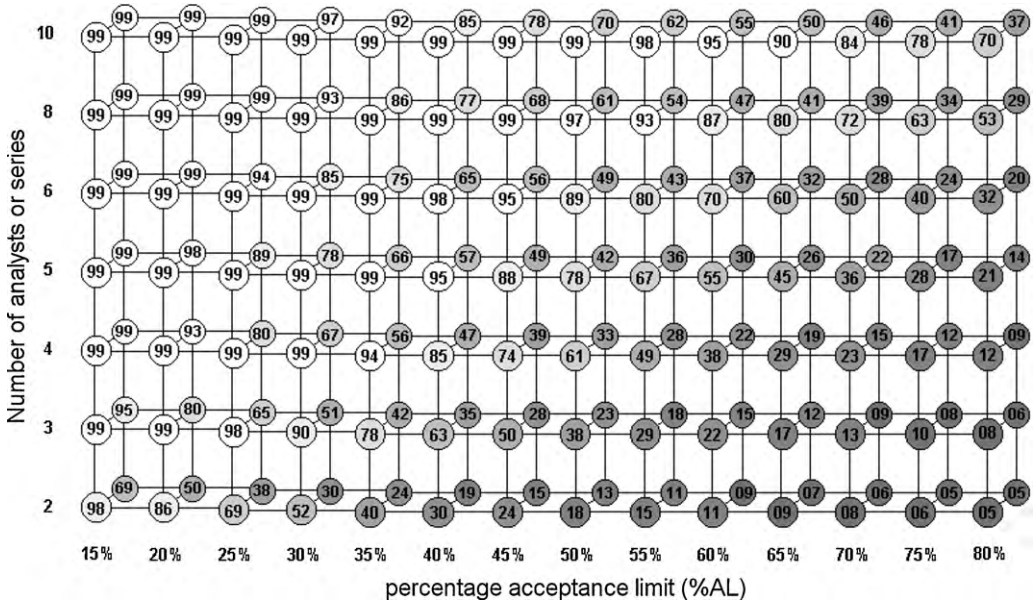

Fig. 4. The encircled values represent the power or $1-\beta$ (probability of correctly accepting a successful method transfer). The bias (true deviation between the labs) in the front layer is set to zero, while the bias in the back layer is half of the acceptance limit. The percentage acceptance limits (%AL) according to Eq. (3) are displayed on the abscissa and the number of analysts/series can be found on the ordinate. Each power value was calculated with a 5% error probability.

unclear as well, if this error is small enough. Perhaps it is, especially for robust methods. However, if it is not, the rejection of in fact successful transfer results will become likely, or an enormous workload will be necessary. These drawbacks can be avoided by using an extended concept, which will be presented in the following.

2. Experimental

The simulation experiments and underlying calculations have been described in detail in [12,14]. Briefly, a basic population was generated, providing 2000 mean values for each laboratory. A specified bias of 0% or 1% was integrated between two simulated laboratories.

The error probability α (one-sided) was chosen to be 0.05. Then the equivalence test was performed (see Section 1.2) using two mean values per site resulting in 2000 tests. The relation of the accepted tests to the number of all performed tests was the probability to correctly accept the transfer. These probability values should be understood with an uncertainty of ±0.5.

All the calculations were performed by various versions (97, 2003, 2007, 2010beta) of Microsoft® Excel (Munich, Germany). The $N(\mu, \sigma^2)$ normal distributed single values x_i were generated by the function: norminv(random(), μ, σ). The mean value μ was always 1.0 and the variance σ^2 (here in terms of standard deviation) according to the simulated conditions (e.g. $\sigma = 0.003 = 0.3/100 = 0.3\%$, see Fig. 4).

3. Results: acceptance criteria according to an extended ISPE concept

3.1. Extended concept

In order to successfully apply the original ISPE concept, the error due to different analysts $\hat{\sigma}_{AN}$ needs to be small [12]. However, this variability is usually unknown. Furthermore, it will strongly depend on the particular method. Perhaps $\hat{\sigma}_{AN}$ is small for robust methods, but this is difficult to determine reliably as this determination itself requires a high number of data. Therefore, usually just estimations will be available for $\hat{\sigma}_{AN}$.

Applying the ISPE concept, rather high errors have to be taken into account when $\hat{\sigma}_{AN}$ is 0.6% or larger. This can be circumvented using a higher number of analysts and series, but this certainly means great effort. Undoubtedly great effort has to be provided if necessary, but it should be avoided whenever possible. It can be avoided by an extended design. The number of necessary data to obtain a certain power depends on the analytical variability, but also on the acceptance criteria.

Acceptance limits (AL) can be reconsidered, too. The ISPE suggests acceptance tolerances of 2%, based on the consideration, that 2% is a frequently used acceptance tolerance for drug substances. If one lab shows 100%, another should not show less the 98% for the same substance. This means that wider acceptance intervals of relevance are suitable when methods control wider acceptance criteria. The absolute difference of the mean values between labs should be equal or less than the method test acceptance limit.

However, the acceptance limits for the calculation of the equivalence test may regularly be chosen wider, because the equivalence test principle intrinsically contains additional safety measures. Thus, even if a method controls 2%, an acceptance tolerance of 2.25% for the method transfer seems reasonable. This is based on the following: consider 1% RSD% measurement spread and a ±2% acceptance limit. If this limit is related to the traditional *t*-test, then 12% error at the acceptance limit must be expected (Fig. 3). The error of refusing a truly acceptable transfer only falls below the limit of 5% error probability (according to the ISPE concept) at a ±2.3% limit. Thus 5% error probability considering a reasonable spread immanently corresponds to a 2.3% limit for the two-sided *t*-test.

Putting it differently, a 12% error probability would correspond to approximately 1.65% deviation for an equivalence test, if 5% error probability corresponds to a 2% limit (Fig. 2). Now it seems reasonable to allow the same error probability for *t*-test and equivalence test at the limit of 2%. This means, the 12% error value can be moved to the 2% acceptance limit. This further means, that the whole error distribution can be rescaled by a factor of 2/1.65. Doing this, the 2% AL would be scaled into a 2.42% limit.

Both considerations show, that using the same error probabilities, wider limit ranges are acceptable using the superior calculation method of the equivalence test. If the measurement

spread is lower, then each test method will provide reasonable results. If the measurement spread is higher, e.g. 2% RSD%, then the traditional *t*-test has more than 60% probability at the acceptance limits. On the contrary, the equivalence test always provides controlled errors close to the acceptance limits.

These considerations above apply to all acceptance limits. Thus all test acceptance criteria can be scaled by a factor of 1.15 related to the traditional *t*-test-related value. If e.g. 1% deviation was the accepted criterion, 1.15% could be used as acceptance limit for the equivalence test (Figs. 2 and 3).

Now having a valid and not unnecessary strict acceptance limit, one can calculate the necessary data number from the given standard error of the mean between laboratories $\hat{\sigma}_{\bar{x}}$ and acceptance limits (Eq. (3), [14]):

$$\%\mathrm{AL} = \frac{\hat{\sigma}_{\bar{x}} 100\%}{\mathrm{AL}} \tag{3}$$

The obtained (percentage) relative acceptance limit, %AL, is used in Fig. 4 to estimate the power of a given experimental design. Fig. 4 corresponds to Fig. 3 in [12]. However, this earlier figure was just valid for the fixed acceptance criteria ±2%, whereas the newly obtained Fig. 4 describes all scenarios with individually chosen acceptance limits.

Let us consider a typical example: $\hat{\sigma}_{\bar{x}} = 0.62$, also equaling RSD%, if the mean value is set to 100% for the sake of simplicity. If now just two analysts are involved in each lab, the error to reject a successful transfer is 50%, even if there is no true bias! If three analysts are involved each, this error becomes acceptable (12%) if there is no bias at all, but again increases dramatically to 51%, if the true bias is 1%, a value which is in fact very much acceptable with acceptance criteria ± 2% (Fig. 3 in [12] and Fig. 4 2-2 in [14]).

Let us now assume the acceptance criteria could be set to ±3% due to specifications given in the same range. Scaling this value to $3 \times 1.15 = 3.45$ and inserting this value for AL into Eq. (3) results in 17.97 for %AL. Looking at the corresponding percentage values in Fig. 4 (between 15% and 20%) demonstrates, that acceptable power is already achieved with just two analysts when no bias is present. If bias is present (in this example half of the acceptance limit), a third analyst may still be advisable to reduce the error to reject a successful transfer, but it will be sure that a third person will always be sufficient. In any case the workload will be substantially reduced using the presented extended concept.

3.2. Guidance to select acceptance criteria and to perform the corresponding equivalence test

The described approach can now readily be customized to ones' own method transfers using the following steps:

1. Set up acceptance limits (AL). AL = acceptance limit of the method (acceptable bias) multiplied by the factor 1.15 (see Section 3.1)
2. Consider expected lab spread. Long-term experience gained by using control charts or by estimations from repeatability (see [15] and references given therein) are most suitable for these considerations.
3. Calculate %AL according to Eq. (3) introduced above.
4. Use the obtained %AL to select a design which guarantees sufficient power.
5. Calculate the lower and the upper limit (C_L, C_U) of the confidence interval for θ according to Eq. (1).
6. Draw the confidence interval [$C_L < \theta < C_U$] into the interval of relevance (interval of acceptance) [$-\mathrm{AL} < \theta_0 < +\mathrm{AL}$].

If it lies completely within: The method transfer was successfully completed. Write a report for documenting purposes.

Else: Reconsider assumptions, especially the expected lab spread. Continue at point 3 of this checklist.

4. Conclusions

In previous works it was already shown that during method transfers, testing for equivalence is more appropriate than the classic two-sample *t*-test. The difference and the advantage become clear again when Figs. 2 and 3 are compared (see also supplementary material on the internet). A fundamental and generally accepted text about analytical method transfer is provided in the Guide for Technology Transfer by the ISPE as it avoids the often described paradoxes by using the appropriately chosen equivalence test instead of a traditional *t*-test. It has proven to be a very sustainable concept which introduces a fixed acceptance tolerance ε of 2%. These 2% were generally accepted for the comparison of mean values for a method transfer of a quantitative analysis. However, there is a shortcoming of this strict setting. As demonstrated in this work and in the mentioned literature as well, it is imminent that $\hat{\sigma}_{\mathrm{AN}}$ (the error due to different analysts) needs to be kept very small. If it exceeds an RSD% of 0.6%, a value which is not uncommon for $\hat{\sigma}_{\mathrm{AN}}$, the ISPE concept leads to unsatisfying results with further costs and time consuming extra work.

Therefore an improvement for the ISPE concept was conceived and introduced in this manuscript in order to avoid these drawbacks. Interestingly, it is the nature of an equivalence test itself, which basically allows for this new extension. The equivalence test intrinsically contains additional safety margins and it always controls the error close to the acceptance limit (Figs. 2 and 3 and supplementary material). These facts led to the assumption that all acceptance limits can be scaled by a factor of 1.15 related to the traditional *t*-test-related value. Nevertheless, the absolute difference of the mean values between labs should be equal or less than the method test acceptance limit.

Furthermore, the extended ISPE concept provides a generalized approach to calculate the necessary data number from the given standard error of the mean between laboratories $\hat{\sigma}_{\bar{x}}$ and acceptance limits (Eq. (3)). It allows for setting up percentage relative acceptance limits (%AL) in order to obtain individually selectable experimental designs. As demonstrated by an example, the workload can be substantially reduced using this extended concept. For the sake of convenience, an easy understandable master standard operation procedure (SOP) was presented in this work which guides the user in only six steps to the experimental design which fits best for his purposes. Further, a template spreadsheet is provided.

Acknowledgement

We are very grateful to Claudia Cianciulli for her skillful support during the preparation of the manuscript.

References

[1] E. Rozet, W. Dewé, E. Ziemons, A. Bouklouze, B. Boulanger, Ph. Hubert, Methodologies for the transfer of analytical methods: a review, J. Chromatogr. 23 (2009) 2214–2223.
[2] CFR, Code of Federal Regulations Chapter 21 §211 (2009) 21CFR211.165.
[3] CFR, Code of Federal Regulations Chapter 21 §211 (2009) 21CFR211.194.
[4] ISPE: International Society for Pharmaceutical Engineering, Technology Transfer, ISPE Guide, 2003, pp. 23–34.
[5] E. Rozet, W. Dewé, R. Morello, P. Chiap, F. Lecomte, E. Ziemons, K.S. Boos, B. Boulanger, J. Crommen, Ph. Hubert, Risk-based approach for the transfer of quantitative methods: bioanalytical applications, J. Chromatogr. A 1189 (2008) 32–41.
[6] M. Broughton, J. Ermer, in: J. Ermer, J.H.McB. Miller (Eds.), Method Validation in Pharmaceutical Analysis, Wiley-VCH Verlag, Weinheim, Germany, 2005, pp. 281–300.
[7] M. Limberger, Transfer of Analytical Methods, http://www.phast.de/Veroeffentlichungen.

[8] W.J. Westlake, Use of confidence intervals in analysis of comparative bioavailability trials, J. Pharm. Sci. 61 (1972) 1340–1341.
[9] D.J. Schuirmann, On hypothesis testing to determine if the mean of a normal distribution is contained in a known interval, Biometrics 37 (1981) 617.
[10] S. Wellek, Statistische Methoden zum Nachweis von Äquivalenz, Fischer, Stuttgart, 1994.
[11] NOVIA GmbH, MVA – Method Validation in Analytics, NOVIA GmbH, Frankfurt, Germany, 2001, available at www.novia.de.
[12] 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.
[13] G. de Fontenay, Analytical method transfer: new descriptive approach for acceptance criteria definition, J. Pharm. Biomed. Anal. 46 (2008) 104–112.
[14] U. Schepers, Dissertation. Statistische Beurteilung der Güte von analytischen Ergebnissen, http://www.digibib.tu-bs.de/?docid=00007347.
[15] H. Wätzig, J. Ermer, Statistik in der Pharmazeutischen Analytik Teil 2: Unsicherheit über Unsicherheit, PZ Prisma 11 (2003) 257–265.
[16] U. Schepers, H. Wätzig, Statistische Signifikanz und Relevanz: Klassische *t*- und *F*-tests oder Äquivalenztest, PZ Prisma 15 (2008), pp. 187–196.
[17] U. Schepers, H. Wätzig, Application of the equivalence test for analytical method transfers: testing precision using the United States Pharmacopoeia concept <1 0 1 0>, J. Pharm. Biomed. Anal. 41 (2006) 290–292.
[18] Quattrochi, et al., Transfer of analytical procedures: a proposal for a new general information chapter, Pharm. Forum 35 (2009) 1380–1382.
[19] Hauck, et al., Acceptable, equivalent or better: approaches for alternatives to official compendial procedures, Pharm. Forum 35 (2009) 772–778.