Available online at www.sciencedirect.com
SCIENCE DIRECT
Talanta 68 (2006) 1166–1175
Talanta
www.elsevier.com/locate/talanta

LC method for the determination of R-timolol in S-timolol maleate: Validation of its ability to quantify and uncertainty assessment

R.D. Marini [a,1], P. Chiap [b,1], B. Boulanger [c], S. Rudaz [d], E. Rozet [a], J. Crommen [b], Ph. Hubert [a,*]

[a] *Laboratory of Analytical Chemistry, Institute of Pharmacy, University of Liège, CHU, B36, B-4000 Liège, Belgium*
[b] *Department of Analytical Pharmaceutical Chemistry, Institute of Pharmacy, University of Liège, CHU, B36, B-4000 Liège, Belgium*
[c] *Lilly Development Centre, Statistical & Mathematical Sciences, rue Grandbompré, 11, B-1348 Mont-Saint-Guibert, Belgium*
[d] *Laboratory of Pharmaceutical Analytical Chemistry, School of Pharmaceutical Sciences, EPGL, University of Geneva, 20, Bd d'Yvoy 1211 Geneva 4, Switzerland*

Received 27 May 2005; received in revised form 6 July 2005; accepted 14 July 2005
Available online 29 August 2005

Abstract

This article presents the validation results of a chiral liquid chromatographic (LC) method previously developed for the quantitative determination of R-timolol in S-timolol maleate samples. A novel validation strategy based on the accuracy profiles was used to select the most appropriate regression model, to assess the method accuracy within well defined acceptance limits and to determine the limits of quantitation as well as the concentration range.

The validation phase was completed by the investigation of the risk profiles of various acceptable regression models in order to ensure the risk of obtaining the future measurements outside the acceptance limits fixed a priori.

On the other hand, the present paper also shows how data used in this validation approach can be used to estimate the measurement uncertainty. The uncertainty derived from β-expectation tolerance interval (σ^2_{Tol}), which is equal to the uncertainty of measurements as well as the expanded uncertainty (U_x) using a coverage factor $k = 2$ was estimated. The uncertainty estimates obtained from validation data were finally compared with those obtained from interlaboratory and robustness studies.

Keywords: Timolol maleate; R-timolol impurity; Liquid chromatography; Validation; Accuracy profile; Risk; Uncertainty assessment

1. Introduction

During the development of analytical methods, it becomes more and more obvious and essential that after the optimisation step the analysts have to demonstrate that the obtained results are reliable for the intended use of the method. In this way, many procedures are available, such as those established by ICH and SFSTP commissions [1–4]. However, in a statistical point of view and by considering the decision making according to the defined acceptance limits and the risk related to the future use of the method, some drawbacks were noticed. Recently, a novel validation strategy based on the use of accuracy profiles has been introduced [5,6]. The notion of including the use of accuracy profiles is in accordance with the objective of an analytical method that can be summarized as its ability to quantify as accurately as possible each of the unknown quantities that the laboratory will have to determine. In fact, what is expected from all analysts when using an analytical method is that the difference observed between the measured result (x_i) and the "true value" (μ_{T}) of the sample (which will always remain unknown) is inferior to an acceptance limit (λ), as can be expressed in the following Eq. (1):

$$|x_i - \mu_{\mathrm{T}}| < \lambda \quad (1)$$

* Corresponding author. Tel.: +32 4 3664316; fax: +32 4 3664317.
E-mail address: ph.Hubert@ulg.ac.be (Ph. Hubert).
[1] These authors contributed equally to this work.

0039-9140/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.talanta.2005.07.026

The acceptance limits are different and depend on the requirement of the analyst or the objective of the analytical method [5,6]. At each concentration level, the accuracy profile is obtained by computing the β-expectation tolerance interval that allows to predict where β% of the future measurements are expected to lie. Therefore, this new strategy clearly shows an advantage over the commonly used procedures by allowing the control of the risks associated to the use of the method. In fact, this notion of risk is linked to the notion of guarantee concerning the future analysis of unknown samples using the validated method [5,6]. As suggested in our previous paper [7], a procedure can be qualified as acceptable if it is very likely, i.e. with a "guarantee", that the difference between every measurement (x_i) of a sample and its "true value" (μ_T) is inside the acceptance limits predefined by the analyst. From this, one can refer to the risk expressing the proportion of measurements that are expected to fall out of acceptance limits ($\pm\lambda$) during the routine analysis. That risk can be evaluated by means of a profile by level of investigated concentration and can be translated by the following Eq. (2):

$$Pr[|x_i - \mu_T| > \lambda] \leq \beta \quad (2)$$

where Pr is the probability that a measurement will fall outside the acceptance limits and β the maximum risk that the analyst is able to take during routine use [6–8].

On the other hand, by considering this new validation strategy, Feinberg et al. [9] introduced the possibility to estimate the uncertainty using the validation data. The definition of uncertainty can be found in the Eurachem guide [10]. From an analytical perspective, this can be considered as straight forward for many analysts. Indeed, even though few approaches have been described for the estimation of uncertainty from validation process [11–16], there is still a need to clarify the relationship between validation and uncertainty for many analysts and particularly to show how the validation data can be practically used to estimate the uncertainty measurement. A recent draft of guide ISO/DTS [17] suggests that experimental data obtained from repeatability, reproducibility and trueness studies could be used to determine uncertainty measurement [9]. Other approaches, such as those proposed by the International Organization for Standardization (ISO) [18] and the Analytical Methods Committee [19] can be applied to estimate the uncertainty.

The first objective of this paper is to fully validate the liquid chromatographic (LC) method for the determination of R-timolol in S-timolol samples, applying this novel validation strategy based on the accuracy profiles. Indeed, the method was previously developed for the simultaneous determination of R-timolol and other related substances in S-timolol maleate bulk material but was not validated [20].

The second objective is to estimate the measurement uncertainty from validation data for the determination of R-timolol content. For this purpose, the approach described in [9] is applied.

Finally, the third objective is the comparison of different studies to evaluate the uncertainty, namely the present validation, the interlaboratory [21] and the robustness [22] studies. In these three studies, the present LC method was concerned to analyze S-timolol maleate samples containing R-timolol impurity at similar concentrations.

2. Experimental

2.1. Chemical and reagents

Samples of S-timolol maleate, R-timolol maleate, isotimolol, dimer maleate and dimorpholinothiadiazole were obtained from the European Pharmacopoeia Secretariat (Strasbourg, France).

N-Hexane of LC grade was purchased from Hipersolv (Poole, England), 2-propanol for analysis from Merck (Darmstadt, Germany) and diethylamine (DEA) for analysis from Sigma (St. Louis, MO, USA).

2.2. Apparatus

The chromatographic system from Shimadzu (Shimadzu Corporation, Kyoto, Japan) was composed of a model LC-10 AT pump, a model SIL-10 AVL automatic injector, a model CTO-10 AC oven and a model SDP-M10 AVP diode array detector. To control the LC system, a Class LC-10 software from Shimadzu was loaded on a Pentium 166 MHz computer. A model CBM-10 Shimadzu interface was used to send the signals from the detector to the computer.

A Chiralcel OD-H column (250 mm × 4.6 mm, i.d.) packed with cellulose tris(3,5-dimethylphenylcarbamate) coated on silica particules (5 μm) from Daicel Limited Industries (Tokyo, Japan) was used. A guard column (4 mm × 4 mm, i.d.) packed with LiChrospher 100 Diol (5 μm) (Merck) and maintained with a holder was used.

The accuracy profiles as well as the statistical calculations including the validation results and the different uncertainty estimates were obtained using the e-noval® software (Arlenda, Belgium). JMP® software Version 5.1 for Windows (SAS Institute, Cary, NC, USA) was also used for further statistical calculations.

2.3. Analytical conditions

The chromatographic separation was carried out using a mobile phase consisting of a mixture of hexane, 2-propanol and DEA pumped at a constant flow rate of 1.0 mL min^{-1}. UV detection was set at 297 nm. Prior to use, the mobile phase was degassed for 15 min in an ultrasonic bath. The injection volume was 10 μL.

2.4. Preparation of standard solutions

The dissolution of analytes and dilution of sample solutions were realized in 2-propanol containing 1% (v/v) of

Table 1
Preparation of standard solutions related to R-timolol

Concentration level (% relative to 1.5 mg mL^{-1} of S-timolol maleate)	Concentration of R-timolol (μg mL^{-1})	
	Calibration standards	Validation standards
0.1	1.5	1.5
0.2	3.0	3.0
0.4	–	6.0
0.8	12.0	12.0
1.6	24.0	24.0
Total	12 samples/day	15 samples/day

DEA. Prior to use, the prepared final solutions were filtered through a 0.5 μm PTFE filter type FH (Millipore Corporation, Bedford, MA, USA).

2.4.1. Solutions used for calibration

A stock solution of R-timolol was prepared by dissolving in a 25-mL volumetric flask an accurately weighted amount of approximately 11.25 mg of R-timolol maleate chemical reference substance (CRS). To allow a complete dissolution in the mixture of 2-propanol and DEA, the stock solution was sonicated in an ultrasonic bath for at least 15 min. Then, subsequent dilutions were performed in order to obtain several solutions at the concentration levels as mentioned in Table 1. These solutions are used as calibration standards (CS). Each solution was injected three times. According to the protocol in [6], the number of concentration levels is sufficient to generate different regression models.

2.4.2. Solutions used for validation

Independent stock solutions of R-timolol were prepared in the same way as mentioned in point Section 2.4.1. Subsequent dilutions were carried out in order to obtain intermediate solutions. Then, the final solutions at the concentrations mentioned in Table 1 were made in 10-mL volumetric flasks previously containing 15 mg of S-timolol maleate accurately weighed. Ultrasonic bath was used for at least 15 min to allow a complete dissolution of samples. Three replicates ($n = 3$) were prepared per concentration level ($m = 5$). The overall preparation step was repeated for 3 days ($p = 3$). These validation standards (VS) were independently prepared in the matrix simulating as much as possible the future routine analysis of impurities in S-timolol maleate samples. In this purpose, S-timolol maleate sample CRS (batch number 8060) was used as matrix since it was one of the most pure sample [20] and was also quantitatively available.

For the evaluation of method selectivity, a blank solution of the solvents mixture was prepared as well as a solution containing a mixture of S-timolol, R-timolol, isotimolol, dimer maleate and dimorpholinothiadiazole at the concentration of about 10 μg mL^{-1}.

3. Results and discussion

3.1. Analytical method

The liquid chromatography (LC) separation of timolol enantiomers and other related substances (isotimolol, dimer maleate and dimorpholinothiadiazole) was carried out on a cellulose tris(3,5-dimethylphenylcarbamate) chiral stationary phase packed in a Chiralcel OD-H column thermostated at 22 °C and using a mobile phase composed of *n*-hexane, 2-propanol and diethylamine (965/35/1, v/v/v). These LC conditions were previously optimized from a Box–Benhken three level design with fifteen experimental points [20,23,24].

3.2. Validation

The strategy that remains usually used for the validation of an analytical method is based on acceptance criteria considering only estimates of observed bias and variance [6,25]. However, in the present study a novel approach using accuracy profiles is applied. It is based on β-expectation tolerance intervals for the total error measurement that includes trueness (bias) and intermediate precision (standard deviation). The advantage of this strategy is the possibility to control the risk of accepting an unsuitable assay by using accuracy profiles while providing the guarantee that the results of the measurement that will be obtained during the future use of the validated method will be included within acceptance limits fixed according to the requirements. Consequently, the accuracy profile can be considered as a useful decision tool to accept or reject a method according to its intended use. Moreover, the accuracy profiles can also be used to select the most appropriate regression model for calibration, to determine the quantitation limits (upper and lower) and subsequently to select a concentration range for the assay.

In a practical point of view, the validation approach based on the accuracy profiles can be carried out applying six important basic steps:

a. selection of the acceptance limits taking into account the intended use of the method;
b. fitting of a regression model from the calibration standards;
c. calculation of the concentrations of all validation standards according to the selected model;
d. determination of the mean bias at each concentration level;
e. calculation of two-sided β-expectation tolerance limits of the mean bias at each concentration level considering the standard deviation for intermediate precision;
f. plotting of the accuracy profile, representing as a function of concentration, the mean bias, the β-expectation tolerance intervals as well as the acceptance limits.

3.2.1. Selectivity

The method selectivity was checked by comparison of typical chromatograms obtained by injecting a blank solution of 2-propanol containing 1% of diethylamine (DEA)

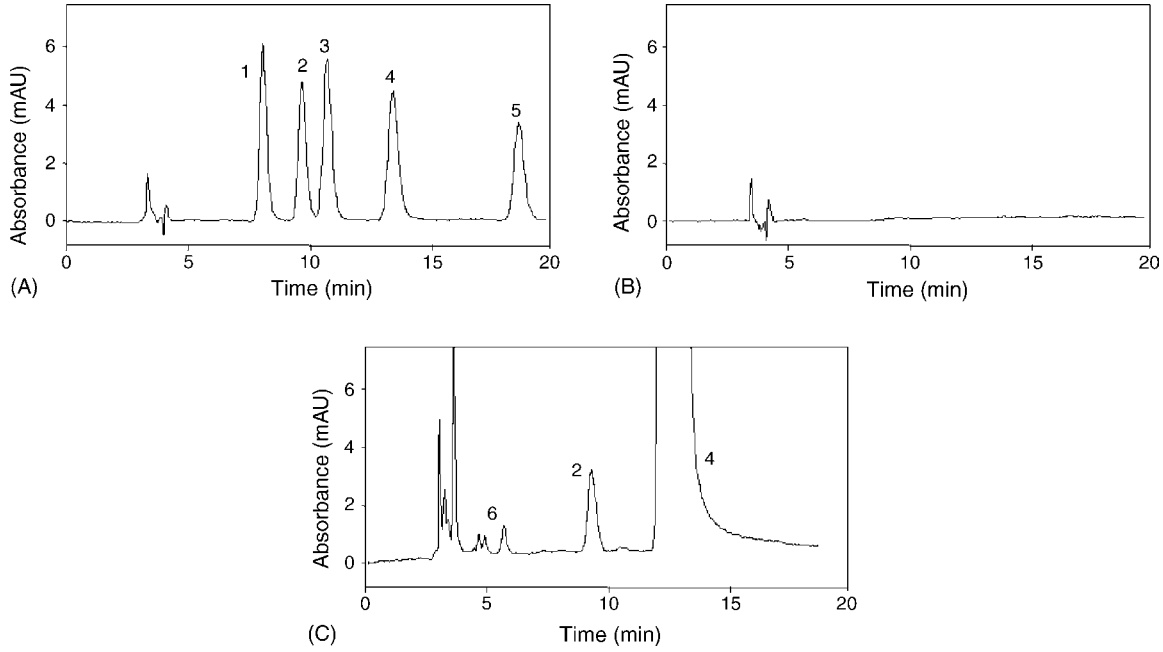

Fig. 1. Typical chromatograms of the mixture solution (A), of the dissolution solvent (B) and of a validation standard at 0.2% (C). *Peaks*: dimer maleate (1), R-timolol (2), isotimolol (3), S-timolol (4), dimorpholinothiadiazole (5) and solvent front (6). *Concentration of analytes*: between 5 and 10 μg mL^{-1} in (A) and about 3 μg/mL of R-timolol calculated against S-timolol maleate in (C).

and a diluted solution of S-timolol, R-timolol, isotimolol, dimer maleate and dimorpholinothiadiazole. As illustrated in Fig. 1A and B, no peak or no interference was observed at the retention times corresponding to those analytes. In addition, as can be seen in Fig. 1C, the peak corresponding to R-timolol at 0.2% concentration level can be easily integrated, and consequently quantified, in presence of the peak corresponding to S-timolol at 1.5 mg mL^{-1}.

3.2.2. Analysis of the response function and selection of the most appropriate regression model

This step constitutes one of the most important steps since the reliability of the validation results that will be obtained are depending on the selected regression model. The response function was evaluated from three calibration curves constructed from the calibration standards using four concentration levels (Table 1) ranging from 0.1 to 1.6% (1.5–24.0 μg mL^{-1}). Then, several regression models were fitted in order to analyze the relationship between concentration (μg mL^{-1}) and analytical response (peak area).

From each regression curve obtained, the concentrations of the validation standard were calculated, which allowed obtaining at each concentration level the mean relative bias, the upper and the lower β-expectation tolerance limits by considering the standard deviation for intermediate precision. Then, different accuracy profiles were plotted from these data as can be seen in Figs. 2 and 3.

The acceptance limits were settled to ±10% since we are dealing with impurities. By considering a risk of 5% and the simple linear regression model (Fig. 2A), it can be noticed that the β-expectation tolerance intervals are out of the acceptance limits at the lowest concentration level. Even with a risk of 10% (Fig. 2B) the lower β-expectation tolerance limit still remains outside the acceptance limit while the upper one becomes borderline the limit at the lowest concentration level.

Other regression models were also tested. The weighted linear regression model with a weighting factor of $1/X^2$ coupled to a risk of 5% was found appropriate as illustrated in Fig. 2C. Indeed, the β-expectation tolerance intervals are comprised within the acceptance limits over all concentration range.

Even if this regression model seems to be the most appropriate to describe adequately the relationship between concentration and analytical response, its application in routine analysis can be time-consuming because it needs the preparation of a calibration curve for the impurity assay. In addition, as it is currently practiced in the Pharmacopoeia's monographs, only one concentration level is used for calibration in order to determine the impurities. In this context, other accuracy profiles considering 24 or 12 μg mL^{-1} as calibration standards levels were investigated. As can be seen in Fig. 3A, by keeping a risk of 5% and using 24 μg mL^{-1} as calibration level, the accuracy profile exceeded the acceptance limits with an increase of the relative error at the lowest concentrations while by using 12 μg mL^{-1} as calibration level (Fig. 3B), the profile was inside the acceptance limits. Even if an increase of relative error was observed at the lowest concentration levels, the objective of the method remained fulfilled, i.e. the ability to quantify R-timolol impurity in S-timolol maleate samples knowing that the maximum content of this chiral impurity tolerated by the monograph of S-timolol maleate is 1.0% [26]. Consequently, this simple

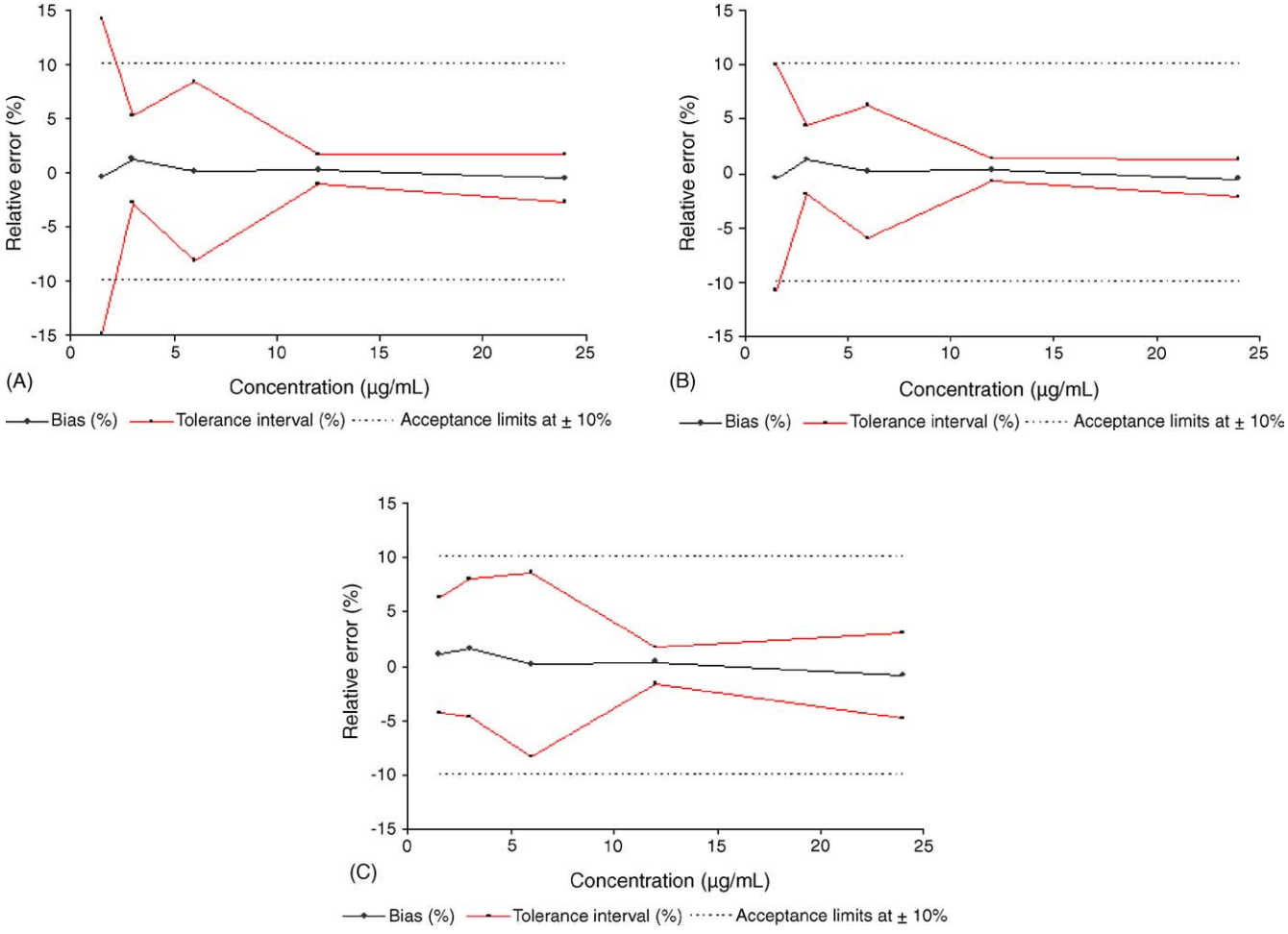

Fig. 2. Accuracy profiles obtained by considering the linear regression model with a risk of 5% (A) and a risk of 10% (B). Accuracy profiles obtained by considering the weighted linear regression model with a risk of 5% (C).

regression model fulfilling its objective can be used in routine analysis and was then used to evaluate the different validation criteria. Moreover, 95 times out of 100, the future measurements will be included within the acceptance limits.

3.2.3. Trueness

The results of trueness [6,25] were expressed in terms of absolute bias (in $\mu g\,mL^{-1}$) or relative bias (%) and were assessed by means of validation standards in the matrix at

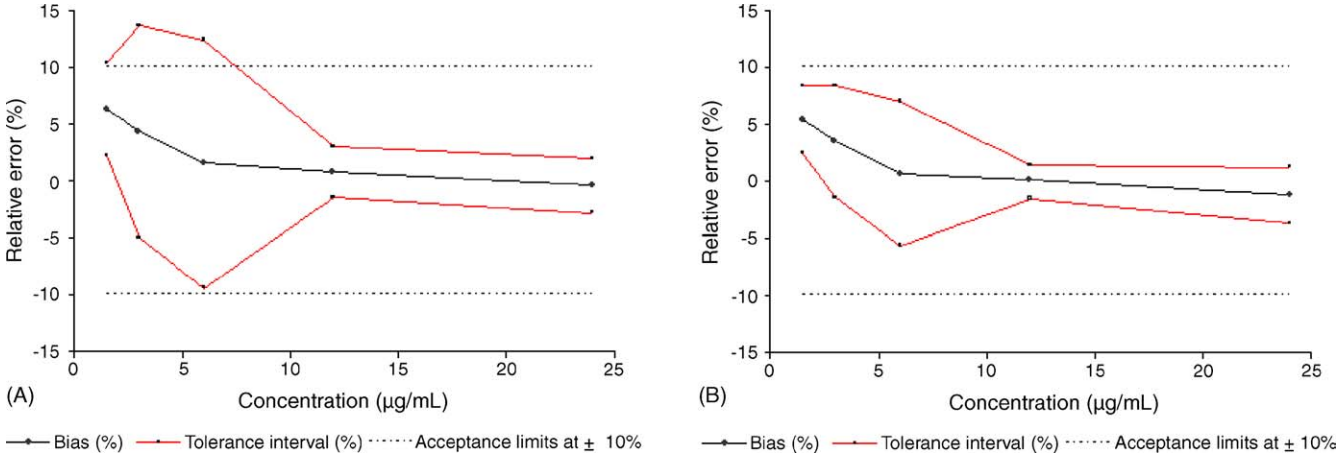

Fig. 3. Accuracy profiles obtained by considering the linear regression model through 0 and fitted using 24.0 $\mu g\,mL^{-1}$ (A) and 12.0 $\mu g\,mL^{-1}$ (B), as concentration levels, with a risk of 5%.

Table 2
Validation results referred to R-timolol using the linear regression model through 0 fitted with 12 μg mL^{-1} as concentration level

Response function (0–12 μg mL^{-1}) (n = 3, m = 1, p = 3)	Slope
Day 1	5016
Day 2	5083
Day 3	5073

Trueness (n = 3; p = 3)	Absolute bias: μg mL^{-1} (relative bias: %)	Recovery (%)
1.5 μg mL^{-1}	0.082 (5.4)	105.4
3.0 μg mL^{-1}	0.106 (3.5)	103.5
6.0 μg mL^{-1}	0.042 (0.7)	100.7
12.0 μg mL^{-1}	−0.002 (0.0)	100.0
24.0 μg mL^{-1}	−0.303 (−1.3)	98.7

Precision (n = 3; p = 3)	Repeatability (R.S.D., %)	Intermediate precision (R.S.D., %)
1.5 μg mL^{-1}	1.2	1.2
3.0 μg mL^{-1}	1.2	1.7
6.0 μg mL^{-1}	1.2	1.9
12.0 μg mL^{-1}	0.6	0.6
24.0 μg mL^{-1}	0.9	1.0

Accuracy (n = 3; p = 3)	β-Expectation tolerance limits in μg mL^{-1} (in %)
1.5 μg mL^{-1}	1.55/1.64 (2.5/8.4)
3.0 μg mL^{-1}	2.98/3.28 (−1.4/8.5)
6.0 μg mL^{-1}	5.71/6.48 (−5.6/7.0)
12.0 μg mL^{-1}	11.9/12.3 (−1.5/1.5)
24.0 μg mL^{-1}	23.3/24.5 (−3.8/1.3)

Linearity (n = 3; m = 5; p = 3), N = 45	
Range (μg mL^{-1})	1.5–24
Slope	0.9825
Intercept	0.1492
r^2	0.9998
LOD (μg mL^{-1}) (% relative to 1.5 mg mL^{-1})	0.27 (0.02%)
LOQ (μg mL^{-1}) (% relative to 1.5 mg mL^{-1})	1.51 (0.1%)

n = replicates; m = concentration levels; p = days.

five concentration levels ranging from 1.5 to 24 μg mL^{-1}. As can be seen in Table 2, the relative biases of the developed method were found acceptable since they are relatively close to 0, except at the lowest concentration levels for which they are around 4 or 5% as discussed above.

3.2.4. Precision

The precision of the developed method was estimated by calculating repeatability and intermediate precision at each concentration level used in validation [4,25]. As shown in Table 2, the relative standard deviations values never exceeded 1.3 and 2.0% for repeatability and intermediate precision, respectively, illustrating the good precision of the developed method.

3.2.5. Accuracy

Accuracy refers to the closeness of agreement between *the test result* and *the accepted reference value*, namely the conventionally true value. The accuracy takes into account the total error, i.e. systematic and random errors, related to the test result [6,25]. It is represented from the accuracy profile illustrated in Fig. 3B. As can be seen from the results in Table 2, the proposed method was accurate over the concentration range investigated, since the different limits of tolerance of bias did not exceed the acceptance limits of ±10% at each concentration level. The accuracy was particularly good around 1.0% (15 μg mL^{-1}) that corresponds to the maximum content of R-timolol tolerated by the monograph of S-timolol maleate [26].

3.2.6. Linearity

The linearity of an analytical method is its ability within a given range to obtain results (not signals) directly proportional to the concentrations (quantities) of the analyte in the sample [4,6,27]. For all series, a regression line was fitted on the calculated concentrations versus the introduced concentrations by applying the linear regression model for which the determination coefficient (r^2) the slope and the intercept are presented in Table 2.

In order to demonstrate the method linearity, the approach based on the absolute β-expectation tolerance limits was applied. As illustrated in Fig. 4 the absolute β-expectation tolerance limits were within the absolute acceptance limits demonstrating the linearity of the present LC method.

3.2.7. Detection and quantitation limits

The limit of detection (LOD) is the smallest quantity of the targeted substance that can be detected, but not accurately quantified in the sample [4,6]. The LOD was estimated using the mean intercept of the calibration model and the residual variance of the regression (Table 2). The lower limit of quantitation (LOQ) is the smallest quantity of the targeted substance in the sample that can be assayed under experimental con-

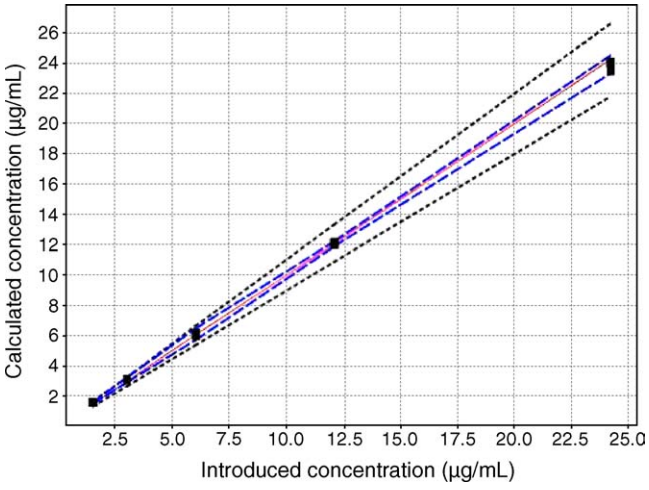

Fig. 4. Linearity graph related to R-timolol LC method. *Legends*: The plain line is the identity line ($y = x$) on which the points should be located in case of validated method; the dashed lines correspond to the accuracy profile, i.e. the β-expectation tolerance limits expressed in absolute values and the dotted lines represent the acceptance limits at ±10%.

ditions with a well-defined accuracy [4,6]. The LOQ was obtained by calculating the smallest concentration beyond which the accuracy limits or β-expectation tolerance limits go outside the acceptance limits. As the accuracy profile was included inside the acceptance limits (Fig. 3B), the first concentration level (1.51 μg mL^{-1}) was considered as the lower LOQ taking into account the selected regression model.

3.3. Profile of risk

The risk profile expresses, by level of concentration that has been investigated, the expected probability that a measurement will fall outside the acceptance limits during routine use, according or conditionally to the estimated bias and precision parameters obtained during the validation phase. The risk or probability to have measurements falling outside the acceptance limits is computed according to Mee [8]. When many determinations are envisaged in routine, this risk expresses the proportion of measurements that are expected to fall outside the acceptance limits during the routine analysis, if everything else remains equivalent, i.e. if no change occurs in the analytical method. Profile of risks can be considered as a new tool among available ones that enable scientific, risk-managed pharmaceutical development and quality assurance as in the present case, since it can provide effective and efficient means for acquiring information to facilitate process understanding and achieve continuous improvement of the analytical method [28]. In this study, the risk was investigated at each concentration level of the validation standards by taking into account the most appropriate regression models previously determined by use of the accuracy profile as decision tool. Therefore, with respect to this strategy, two profiles of risk were plotted on the basis of the weighted linear regression model and the linear regression model through 0 with 12 μg mL^{-1} as concentration level for calibration. As can be observed in Fig. 5, the risk of having future measurements out of the acceptance limits of 10% was practically null at concentration levels above 12 μg mL^{-1} while this risk increased and culminated at 6 μg mL^{-1} (risk equals to 4.1%) and 3 μg mL^{-1} (risk equals to 2.3%) by considering the weighted linear regression model and the linear regression model through 0, respectively. In all cases, the risks did not exceed 5% and were even lower by considering this last regression model justifying its selection for calibration. When analyzing S-timolol samples containing R-timolol at concentrations equivalent to about 3.0 μg mL^{-1} (0.2%), 97.5 times out of 100 the future measurements given by the validated LC chiral method will be included in the acceptance limits of 10%. This proportion still increases when the concentration of the chiral impurity is higher than 0.4%.

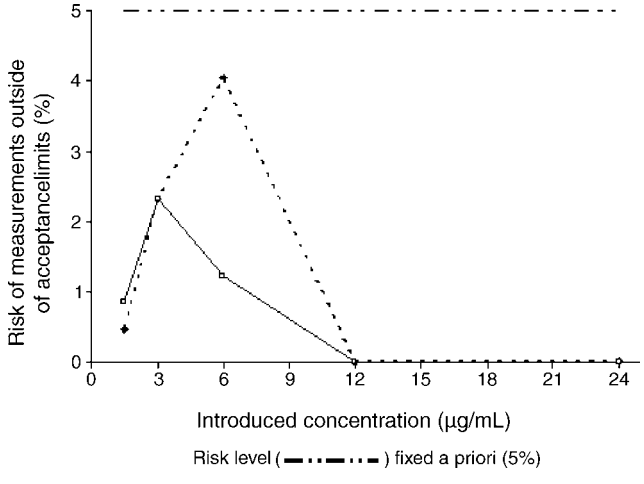

Fig. 5. Risk profiles obtained by considering the weighted linear regression model (dotted line) and the linear regression model through 0 and fitted using 12.0 μg mL^{-1} (continuous line) as concentration level.

3.4. Uncertainty

3.4.1. Assessment from validation data

The data used to carry out the validation using the accuracy profile were also used for the estimation of uncertainty of measurements. In this way, the uncertainty is derived from the variance used to construct the β-expectation tolerance limits and is equal to:

$$\sigma^2_{\mathrm{Tol}} = \left[\sqrt{1 + \frac{1}{pnB^2}}\,\hat{\sigma}_{\mathrm{M}}\right]^2 \tag{3}$$

with $B = \sqrt{\frac{A+1}{nA+1}}$, $A = \hat{\sigma}^2_{\mathrm{B}}/\hat{\sigma}^2_{\mathrm{W}}$ and $\hat{\sigma}^2_{\mathrm{M}}$ is the estimated total variance or intermediate precision, σ^2_{B} and σ^2_{W} the between- and the within-series variances, respectively, n and p the number of replicates and series, respectively. In the present case, only the day effect was investigated. Feinberg et al. [9] demonstrated that σ^2_{Tol} is equal to the uncertainty of the measurements. On this basis, several uncertainty results were generated and are presented in Table 3. They were obtained by using the last regression model selected during validation. As shown, the values are comprised between 0.006 and 0.094 μg mL^{-1}and 0.019 and 0.253 μg mL^{-1} for the bias uncertainty and the measurement uncertainty, respectively. It was remarked that the uncertainty seems to increase exponentially with the concentration. The expanded uncertainty (U_X), which equals to the standard uncertainty multiplied by a coverage factor (k) [10,12,29], defines an interval around the mean value in which the unknown "true value" is retrieved with a defined probability. The choice of this factor is based on the confidence level desired. By considering $k = 2$ [10,19,29] which means that U_X is approximately equivalent to 95% level of confidence, our expectation was fulfilled since all the measurements observed for each concentration level obtained using the selected regression model were comprised within the acceptance limits of 10%. In addition, the relative expanded uncertainties (%) obtained by dividing the corresponding expanded uncertainties with the corresponding introduced concentrations (Table 3) are not exceeding 5%.

Table 3
Point estimates of the different uncertainties related to R-timolol content at each concentration level of the accuracy profile using the selected regression model

Concentration level ($\mu g\,mL^{-1}$)	Uncertainty of the bias ($\mu g\,mL^{-1}$)	Uncertainty ($\mu g\,mL^{-1}$)	Expanded uncertainty ($\mu g\,mL^{-1}$)	Relative expanded uncertainty (%)
1.5	0.61×10^{-2}	1.93×10^{-2}	3.86×10^{-2}	2.55
3.0	2.34×10^{-2}	5.55×10^{-2}	11.11×10^{-2}	3.67
6.0	5.78×10^{-2}	12.97×10^{-2}	25.94×10^{-2}	4.28
12.0	2.66×10^{-2}	7.75×10^{-2}	15.50×10^{-2}	1.28
24.0	9.32×10^{-2}	25.25×10^{-2}	50.51×10^{-2}	2.09

3.4.2. Comparison of different studies to evaluate the uncertainty

The studies concerned here, namely the validation, the interlaboratory [21] and the robustness [22] can be compared since the same S-timolol maleate samples containing R-timolol impurity at similar concentrations were analyzed in each study with the same analytical method. To allow this comparison, the results of R-timolol content expressed in percentage in the robustness and the interlaboratory studies have to be transformed in $\mu g\,mL^{-1}$ as in validation study.

The main variance components in the three sets of studies were estimated as described in the previous paper [21,22], namely the intermediate precision obtained from validation ($IP_{Validation}$), from robustness ($IP_{Robustness}$) and from inter-laboratory study ($IP_{Inter\text{-}laboratory}$) as well as the reproducibility issued from inter-laboratory study ($Repro_{Inter\text{-}laboratory}$). The variance components were modelled as a function of the concentration and as a function of the study (4).

$$\text{Variance} = f(\text{concentration} + \text{study}) + \varepsilon \quad (4)$$

As the variance exponentially increases with the concentration, the naperian logarithm (ln) was used to linearize the relation between variance and concentration. The model fitted was (5):

$$\ln \text{variance} = \alpha + \beta \ln \text{concentration} + \gamma\,\text{study} + \varepsilon \quad (5)$$

where β is the estimate of the slope and γ the vector of estimates of the effect of the kind of study, α and ε the intercept and the residual error of the regression model, respectively. The regression coefficient (r^2) of this model equals to 0.9465 suggesting that variances are adequately modelled by that way with corresponding intercept of −5.547 and slope of 1.624, suggesting that for this method the variances of measurements increase nearly as a function of the square of the concentration. This model allows to statistically compare the various variance component estimates obtained in the various study sets. Using the residual error of the model, then various hypotheses can be tested using contrast *t*-test. The estimated contrasts presented in Table 4 show clearly that $IP_{Robustness}$, $IP_{Inter\text{-}laboratory}$ and $Repro_{Inter\text{-}laboratory}$ variance component estimates are significantly greater than $IP_{Validation}$ components. This result is obvious when looking at the graphic of Fig. 6A. It is also reasonably expected since validation studies are usually performed in very well controlled and strict conditions that are not comparable to conditions used in inter-laboratory studies, and, by definition, not with robustness studies since within this case analytical conditions are deliberately modified [2].

Related to this debate, another interesting question is to evaluate if $IP_{Robustness}$ is predictive of the variances levels that could be obtained in inter-laboratory studies, as one could expect. Contrasts performed by comparing $IP_{Inter\text{-}laboratory}$ and $Repro_{Inter\text{-}laboratory}$ to $IP_{Robustness}$ as presented in Table 4 show that $IP_{Robustness}$ is not significantly different from $IP_{Inter\text{-}laboratory}$ but smaller than $Repro_{Inter\text{-}laboratory}$. The levels of intermediate precision variance (s^2_{IP}), i.e. Within- plus between-day variances are comparable when estimated from robustness and inter-laboratory studies (see Eq. (6)). This result is consistent with expectations.

$$s^2_{IP} = s^2_{\text{Within-day}} + s^2_{\text{Between-day}} \quad (6)$$

A closer look at the data issued from collaborative study as already indicated in a previous paper [21] shows that one laboratory presents outlying values that make the between-laboratory variance ($s^2_{\text{Between-laboratory}}$) large and so the inter-laboratory reproducibility ($s^2_{\text{Reproducibility}}$) large as well, since the later corresponds to:

$$s^2_{\text{Reproducibility}} = s^2_{IP} + s^2_{\text{Between-laboratory}}. \quad (7)$$

or

$$s^2_{\text{Reproducibility}} = s^2_{\text{Between-replicates}} + s^2_{\text{Between-days}} + s^2_{\text{Between-laboratory}}. \quad (8)$$

Table 4
Comparison of variance components estimates of the three studies involving the LC method related to R-timolol determination and including all the laboratories in the interlaboratory study

	Contrast by *t*-test				
$IP_{Validation}$					
$IP_{Robustness}$					
$IP_{Inter\text{-}Laboratory}$					
$Repro_{Inter\text{-}Laboratory}$					
Estimated difference between curves	2.7152	4.3633	3.5043	1.648	0.7891
S.E.	0.5697	0.5697	0.5697	0.6806	0.6806
t-Ratio	4.7659	7.6585	6.151	2.4214	1.1594
p-Value[a] (Student's test)	**0.0031**	**0.0003**	**0.0008**	0.0518	0.2903

[a] Significant values at 5% level are printed in bold.

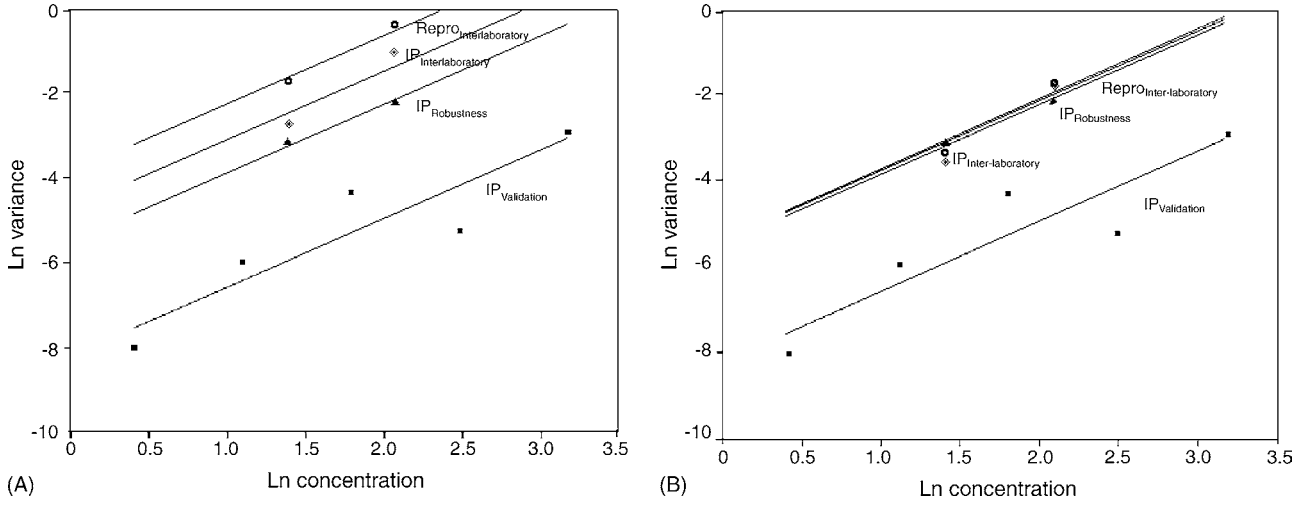

Fig. 6. Regression plots of variance components as function of concentration in validation, robustness and interlaboratory studies including all the laboratories (A) and without laboratories 7 (B).

The scenario of excluding this laboratory was envisaged for the purpose of the present paper in order to examine what are the best conditions to determine the variability and subsequently the uncertainty of analytical results. This practice of excluding data is not recommended when dealing with a specific interlaboratory study. As previously mentioned in [21], results from laboratory 7 suggested that this laboratory has not correctly applied the recommendation for using the validated LC method. By removing this laboratory, one can examine more appropriately what belong to the method and what belong to the laboratory practices for the purpose of this paper only, i.e. comparing the uncertainty from different kind of studies. Once more, it can be noticed that the variance is adequately modelled using the same model previously applied with r^2 equaling to 0.9245. As illustrated in Fig. 6B and as indicated on the contrasts in Table 5, $IP_{Inter\text{-}laboratory}$, $Repro_{Inter\text{-}laboratory}$ and $IP_{Robustness}$ are significantly greater than the $IP_{Validation}$. However, the intermediate precision obtained with robustness study ($IP_{Robustness}$) predicts very well the uncertainty obtained with inter-laboratory study. Indeed, $IP_{Inter\text{-}laboratory}$ and $Repro_{Inter\text{-}laboratory}$ are not different from $IP_{Robustness}$. This was expected since the analytical method has been applied in slightly various experimental conditions from laboratory to laboratory that have been "simulated" by the robustness study. This later was precisely intended to demonstrate that the analytical method can be transferred by deliberately and slightly changing the experimental conditions. With that respect, the present robustness study fulfills thus its objective. The difference observed between the validation and the inter-laboratory or robustness studies clearly demonstrate that uncertainty heavily depends on the conditions on which it has been obtained, i.e. the experimental design. Consequently, it can be stressed that if few variation sources are introduced in the experimental design, the estimate of uncertainty obtained in very similar analytical conditions that may likely not be powerful enough (from a statistical point of view) to mimic the conditions of routine analysis making a poor and possibly not relevant estimate of uncertainty.

Table 5
Comparison of variance components estimates of the three studies involving the LC method related to R-timolol determination and excluding laboratory 7 in the interlaboratory study

	Contrast by *t*-test				
$IP_{Validation}$					
$IP_{Robustness}$					
$IP_{Inter\text{-}Laboratory}$					
$Repro_{Inter\text{-}Laboratory}$					
Estimated difference between curves	2.7157	2.8427	2.7968	0.1271	0.0811
S.E.	0.5715	0.5715	0.5715	0.6827	0.6827
t-Ratio	4.7522	4.9745	4.8942	0.1861	0.1189
p-Value[a] (Student's test)	**0.0032**	**0.0025**	**0.0027**	0.8585	0.9093

[a] Significant values at 5% level are printed in bold.

Therefore, in case R-timolol impurities are to be quantified in S-timolol maleate samples applying the developed LC method by the same operator in the same laboratory, the uncertainty estimates obtained from validation remain appropriate estimates of uncertainty of the results provided by that laboratory. However, if one intends to transfer the method for a large scale use especially to several laboratories, then the estimates obtained from robustness study are the most appropriate ones to predict the estimates of the reproducibility study.

4. Conclusion

In the present paper, a novel validation strategy based on the accuracy profiles was successfully applied to demonstrate the ability of the LC method to quantify R-timolol in

S-timolol maleate. By considering the risks of 5 and 10% and the adequate regression model, several accuracy profiles were constructed from which a suitable regression model was selected namely the linear regression model through 0 and using only one calibration standard (12 $\mu g\,mL^{-1}$). The different validation criteria were evaluated applying that selected regression model. As a risk-management tool in quality assurance, the profiles of risks were also investigated in order to evaluate the probability to obtain the future measurements outside the defined acceptance limits. Moreover, data used in this validation approach was also used to estimate the uncertainty of bias as well as the expanded uncertainty at each concentration level. Both seemed to be concentration dependent. The comparison of different studies to evaluate the uncertainty showed that the uncertainty estimates obtained from validation are significantly different to those obtained in robustness and in inter-laboratory studies. Nevertheless, this approach based on the use of validation data is appropriate to calculate the uncertainty estimates if the validation experimental design is adequate and the analytical method remains in the laboratory, which validated it.

Acknowledgements

Thanks are due to the Belgian Government (The Prime Minister Services – Federal Office for Scientific, Technical and Cultural Affairs, Standardisation Programme) for the financial support of this project (No. NM/12/23). Research grant from the Walloon Region and the European Social Fund to one of the author (E. Rozet) is gratefully acknowledged (First Europe Objective 3 project no. 215269). The authors are also grateful to the European Pharmacopoeia Secretariat for providing them with samples.

References

[1] Food and Drug Administration: International Conference on Harmonization, Definitions and Terminology, Fed. Regist. 60 (1995) 11260.
[2] Food and Drug Administration: International Conference on Harmonization, Methodology, Fed. Regist. 62 (1997) 27463.
[3] J. Caporal-Gauthier, J.M. Nivet, P. Algranti, M. Guilloteau, M. Histe, M. Lallier, J.J. N'guyen-Huu, R. Russoto, S.T.P. Pharma Pratiques 2 (4) (1992) 205.
[4] E. Chapuzet, N. Mercier, S. Bervoas-Martin, B. Boulanger, P. Chevalier, P. Chiap, D. Granjean, Ph. Hubert, P. Lagorce, M. Lallier, M.C. Laparra, M. Laurentie, J.C. Nivet, S.T.P. Pharma Pratiques 7 (3) (1997) 169.
[5] B. Boulanger, P. Chiap, W. Dewé, J. Crommen, Ph. Hubert, J. Pharm. Biomed. Anal. 32 (2003) 753.
[6] Ph. Hubert, J.J. Nguyen-Huu, B. Boulanger, E. Chapuzet, P. Chiap, N. Cohen, P.A. Compagnon, W. Dewé, M. Feinberg, M. Lallier, M. Laurentie, N. Mercier, G. Muzard, C. Nivet, L. Valat, S.T.P. Pharma Pratiques 13 (3) (2003) 101.
[7] Ph. Hubert, J.-J. Nguyen-Huu, B. Boulanger, E. Chapuzet, P. Chiap, N. Cohen, P.-A. Compagnon, W. Dewé, M. Feinberg, M. Lallier, M. Laurentie, N. Mercier, G. Muzard, C. Nivet, L. Valat, J. Pharm. Biomed. Anal. 36 (2004) 579.
[8] R. Mee, Commun. Stat. – Theory Meth. 17 (5) (1988) 1465.
[9] M. Feinberg, B. Boulanger, W. Dewé, Ph. Hubert, Anal. Bioanal. Chem. 380 (2004) 502.
[10] Eurachem/Citac Guide, Quantifying the Uncertainty in Analytical Measurement, second ed., 2000.
[11] E. Hund, D.L. Massart, J. Smeyers-Verbeke, Anal. Chim. Acta 480 (2003) 39.
[12] A. Díaz, L. Vàzquez, F. Ventura, M.T. Galceran, Anal. Chim. Acta 506 (2004) 71.
[13] J.O. De Beer, P. Baten, C. Nsengyumva, J. Smeyers-Verbeke, J. Pharm. Biomed. Anal. 32 (2003) 767.
[14] A. Maroto, J. Riu, R. Boqué, F.X. Rius, Anal. Chim. Acta 391 (1999) 173.
[15] P. Dehouck, Y. Vander Heyden, J. Smeyers-Verbeke, D.L. Massart, J. Crommen, Ph. Hubert, R.D. Marini, O.S.N.M. Smeets, G. Decristoforo, W. Van de Wauw, J. De Beer, M.G. Quaglia, C. Stella, J.-L. Veuthey, O. Estevenon, A. Van Schepdael, E. Roets, J. Hoogmartens, Anal. Chim. Acta 481 (2003) 261.
[16] A.G. González, M.A. Herrador, A.G. Asuero, Talanta 65 (2005) 1022.
[17] ISO/DTS 21748, Guide to the Use of Repeatability, Reproducibility and Trueness Estimates in Measurement Uncertainty Estimation, ISO, Geneva, 2003.
[18] International Organization for Standardization, Guide for the Expression of Uncertainty in Measurements, International Organization for Standardization, Geneva, 1993.
[19] Analytical Methods Committee, Analyst 120 (1995) 2303–2308.
[20] R.D. Marini, P. Chiap, B. Boulanger, W. Dewé, Ph. Hubert, J. Crommen, J. Sep. Sci. 26 (2003) 809.
[21] R.D. Marini, N. Matthijs, Y. Vander Heyden, J. Smeyers-Verbeke, P. Dehouck, J. Hoogmartens, P. Silvestre, A. Ceccato, Ph. Goedert, J. Saevels, C. Herbots, G. Caliaro, R. Herráez-Hernández, J. Verdú-Andrès, P. Campíns-Falcó, W. Van de Wauw, J. De Beer, B. Boulanger, P. Chiap, J. Crommen, Ph. Hubert, Anal. Chim. Acta 546 (2005) 182.
[22] R.D. Marini, B. Boulanger, Y. Vander Heyden, P. Chiap, J. Crommen, Ph. Hubert, Anal. Chim. Acta 531 (2005) 131.
[23] G.E.P. Box, D.W. Behnken, Technometrics 2 (1960) 455.
[24] W. Dewe, R.D. Marini, P. Chiap, Ph. Hubert, J. Crommen, B. Boulanger, Chemom. Intell. Lab. Syst. 74 (2004) 263.
[25] ISO 5725-1, Application of the Statistics—Accuracy (Trueness and Precision) of the Results and Methods of Measurement. Part 1: General Principles and Definitions. International Organization for Standardization (ISO), Geneva, 1994.
[26] European Pharmacopoeia, fifth ed., Council of Europe, Strasbourg, France, 2004, p. 2796.
[27] Ph. Hubert, P. Chiap, J. Crommen, B. Boulanger, E. Chapuzet, N. Mercier, S. Bervoas-Martin, P. Chevalier, D. Grandjean, P. Lagorce, M. Lallier, M.C. Laparra, M. Laurentie, J.C. Nivet, Anal. Chim. Acta 391 (1999) 135.
[28] Food and Drug Administration, Process Analytical Technology (PAT) Initiative, 2004, http://www.fda.gov/cder/guidance/index.htm.
[29] EA-4/16, EA Guidelines on the Expression of Uncertainty in Quantitative Testing, 2004, http://www.european-accreditation.org.