Available online at www.sciencedirect.com
ScienceDirect
Analytica Chimica Acta 591 (2007) 239–247
ANALYTICA CHIMICA ACTA
www.elsevier.com/locate/aca

# Improvement of the decision efficiency of the accuracy profile by means of a desirability function for analytical methods validation Application to a diacetyl-monoxime colorimetric assay used for the determination of urea in transdermal iontophoretic extracts

E. Rozet [a,*], V. Wascotte [b], N. Lecouturier [b], V. Préat [b], W. Dewé [c], B. Boulanger [d], Ph. Hubert [a]

[a] *Laboratory of Analytical Chemistry, Bioanalytical Chemistry Research Unit, Drug Research Center, Université de Liège, B 36, B-4000 Liège, Belgium*
[b] *Pharmaceutical Technology Department, School of Pharmacy, Université Catholique de Louvain, B-1200 Brussels, Belgium*
[c] *GSK Bio, 89, Rue de l'Institut, B-1330 Rixensart, Belgium*
[d] *UCB Pharma SA, Chemin du Foriest, B-1420 Braine-L'alleud, Belgium*

Received 21 December 2006; received in revised form 2 April 2007; accepted 3 April 2007
Available online 7 April 2007

**Abstract**

Validation of analytical methods is a widely used and regulated step for each analytical method. However, the classical approaches to demonstrate the ability to quantify of a method do not necessarily fulfill this objective. For this reason an innovative methodology was recently introduced by using the tolerance interval and accuracy profile, which guarantee that a pre-defined proportion of future measurements obtained with the method will be included within the acceptance limits. Accuracy profile is an effective decision tool to assess the validity of analytical methods. The methodology to build such a profile is detailed here. However, as for any visual tool it has a part of subjectivity. It was then necessary to make the decision process objective in order to quantify the degree of adequacy of an accuracy profile and to allow a thorough comparison between such profiles. To achieve this, we developed a global desirability index based on the three most important validation criteria: the trueness, the precision and the range. The global index allows the classification of the different accuracy profiles obtained according to their respective response functions. A diacetyl-monoxime colorimetric assay for the determination of urea in transdermal iontophoretic extracts was used to illustrate these improvements.


*Keywords:* Colorimetric method; Iontophoretic extracts; Validation; Accuracy profile; Tolerance interval; Desirability index

## 1. Introduction

Validation of analytical methods is a well known and accepted concept in all analytical development laboratories of the chemical, pharmaceutical, biopharmaceutical or agro-food industries. Indeed validation is a mandatory step in the life cycle of an analytical method as it is highly regulated: ICH Q2(R1) document for the pharmaceutical industry [1], ISO 5725 and decision 2002/657/EC for chemical and agro-food industries [2,3], FDA guide for bioanalytical studies [4], etc. However, as can be seen from the scientific literature, even if the validation criteria are defined, validation methodology is highly discussed (see for example [5–12]). Recently, an original approach based on the tolerance interval based on the total error and accuracy profile has been introduced [13]. This approach has the advantages that it can be applied to any type of analytical technique, in any kind of industries and is independent of the matrix in which the analyte of interest is analyzed (pharmaceutical formulation, biological fluid, soil/earth, ...). Furthermore by choosing the tolerance interval as the critical parameter to assess the validity of a method, this approach allows to conciliate the validation objectives with the main aim of any analytical method, that is

* Corresponding author at: Laboratory of Analytical Chemistry, Institute of Pharmacy, Université de Liège, CHU, B 36, B-4000 Liège, Belgium.
Tel.: +32 4 3664316; fax: +32 4 3664317.
*E-mail address:* Eric.Rozet@ulg.ac.be (E. Rozet).

0003-2670/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.aca.2007.04.002

quantify accurately during routine use. Indeed during routine use, the samples studied have an unknown concentration of the target analyte that has to be estimated. Therefore, when analyzing unknown samples, only the total error of its measurement is taken into account. It is impossible to estimate separately, on this sample, the systematic and random errors. One advantage of this approach is that it reflects more directly the performance of individual assays and will result in fewer rejected in-study runs than the current procedures that compare point estimates of observed bias and precision. Indeed classical approaches conclude about the validity of an analytical procedure by comparing to a priori fixed acceptance limits on one hand the systematic error and on the other hand the random error. Another advantage is that, this approach allows to control the risk of accepting an unsuitable assay while providing guarantee that the results of the measurements that will be obtained during the future use of the validated method will be included within acceptance limits fixed according to the requirements. This last point is not only essential for the analyst as his conclusions or diagnostics are made on behalf of the results he obtained. It is also an increasing request of new regulatory requirements to manage the risk associated to the use of these methods in routine analysis [14,15].

Accuracy profile is therefore an interpretable decision tool to assess the validity of an analytical method, i.e. its ability to quantify as accurately as possible. The methodology to build such a profile is detailed in the present paper. However, as for any visual tool it has a part of subjectivity. Therefore, there is a need to increase the objectivity of this visual analysis in order to quantify the degree of adequacy of an accuracy profile and therefore of the analytical method, to compare thoroughly different profiles and to allow the classification of the different accuracy profiles build according to their respective response functions. To achieve this a global desirability index was developed based on three major quality criteria, namely the Trueness, Precision and Dosing Range.

The aims of this paper are (1) to present the methodology to compute total error and build accuracy profiles; (2) to develop and apply a desirability function to increase the objectivity of the decision process; and (3) to demonstrate the applicability of this original approach and its improvements to the validation of a newly developed colorimetric method for the quantitation of urea in transdermal iontophoretic extracts.

## 2. Terminology

Before going on with the main aims of this paper it is important to define three crucial terms which are sources of confusion depending on the industrial environment in which one works.

- First, *accuracy* as defined in ISO documents or in documents ICH Q2R1 section terminology is "*the closeness of agreement between the value which is accepted either as a conventional true value or an accepted reference value and the value found*" [1,2,4,13,16]. It therefore refers to total measurement error.
- A second important term is *trueness*, which refers to "*the closeness of agreement between the average value obtained from a large series of test results and an accepted reference value*" [2,13,16]. This concept is therefore related to the systematic error of a measurement process.
- The last term to define is *precision*, which refers to "*the closeness of agreement (degree of scatter) between a series of measurements obtained from multiple sampling of the same homogeneous sample under the prescribed conditions*" [1,2,4,13,16]. This is related to the random error of a measurement process.

As can be understood from these definitions, the main point is that Accuracy is the simultaneous combination of both systematic and random errors, i.e. total error [2,10,13].

## 3. Experimental

### 3.1. Building accuracy profile

The total error of analytical measurement is the difference between a measurement result ($x_i$) and the unknown "true value" ($\mu_T$). Total error is underlying the accuracy concept. One way to estimate this total error is to compute the $\beta$-expectation tolerance interval ($\beta$-TI) introduced by Mee [17], and to compare it to acceptance limits $\lambda$ [13,18]. Indeed this $\beta$-TI is the interval where it is expected that a proportion $\beta$ of future measurements will fall inside. Therefore, as long as this $\beta$-TI is included inside the acceptance limits, conditions of a valid analytical method are fulfilled.

The equation of the $\beta$-expectation tolerance interval is:

$$\left[\hat{\mu}_j - Q_t\left(\nu; \frac{1+\beta}{2}\right)\sqrt{1+\frac{1}{pnB_j^2}}\hat{\sigma}_{IP,j}; \hat{\mu}_j + Q_t\left(\nu; \frac{1+\beta}{2}\right)\sqrt{1+\frac{1}{pnB_j^2}}\hat{\sigma}_{IP,j}\right] \quad (1)$$

where:

- $p$ is the number of series.
- $n$ is the number of independent replicates per series.
- $\hat{\mu}_j$ is the estimate of the mean results of the $j$th concentration level.
- $\hat{\sigma}^2_{IP,j} = \hat{\sigma}^2_{W,j} + \hat{\sigma}^2_{B,j}$ is the estimate of the intermediate precision variance at the $j$th concentration level, which is the sum of the within series variance $\hat{\sigma}^2_{W,j}$ and between series variance $\hat{\sigma}^2_{B,j}$.
- $R_j = \frac{\hat{\sigma}^2_{B,j}}{\hat{\sigma}^2_{W,j}}$;
- $B_j = \sqrt{\frac{R_j+1}{nR_j+1}}$;
- $\nu = \frac{(R+1)^2}{R+(1/n)/(p-1)+1-(1/n)/(pn}$ [19];
- $Q_t\left(\nu; \frac{1+\beta}{2}\right)$ is the $\beta$ quantile of the Student $t$ distribution with $\nu$ degrees of freedom.

As the accuracy profile expresses error values in relative scale (Fig. 1) [13], Eq. (1) of the $\beta$-TI can be rewritten in relative error

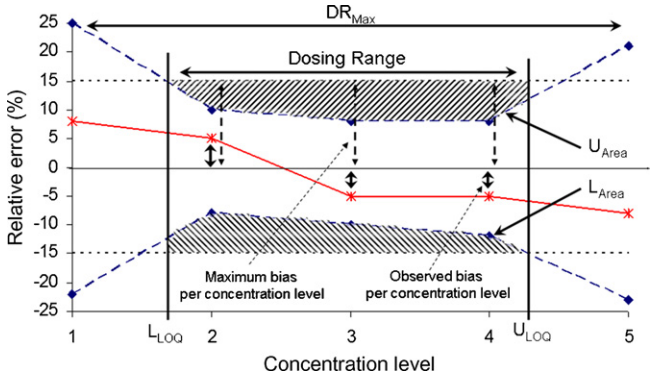


Fig. 1. Schematic representation of an accuracy profile. Illustration of the computation of the indexes: dosing range, trueness and precision. The dotted lines are the upper and lower acceptance limits, the dashed lines are the upper and lower $\beta$-expectation tolerance limits and the continuous line is the relative bias.

scale as follow:

$$\left[\text{bias}(\%)_j - Q_t\left(v; \frac{1+\beta}{2}\right)\sqrt{1+\frac{1}{pnB_j^2}}\text{R.S.D.}_{IP,j}; bias(\%)_j + Q_t\left(v; \frac{1+\beta}{2}\right)\sqrt{1+\frac{1}{pnB_j^2}}\text{R.S.D.}_{IP,j}\right] \quad (2)$$

where:

$$\text{bias}(\%)j = \frac{\hat{\mu}_j - \mu_{Tj}}{\mu_{Tj}} \times 100 \quad (3)$$

$$\text{R.S.D.}IP,\ j = \frac{\hat{\sigma}_{IP,j}}{\hat{\mu}_j} \times 100 \quad (4)$$

One can therefore see that this interval is computed by integrating jointly the estimates of the relative bias and of the intermediate precision R.S.D. obtained for a defined concentration level of the validation standards, together with the level of risk chosen. These estimates are necessary to compute the $\beta$-TI, but they are not compared separately to acceptance limits. They are combined simultaneously into the $\beta$-TI therefore giving a prediction of the future total error of the measurement that is expected. Furthermore, this interval defines a region where it is expected that a defined proportion of future results will lay, being a predictive tool of the behavior of the analytical procedure under investigation.

An ANOVA I model is fitted in order to estimate the bias and the precision through estimates of $\mu_j$, $\sigma^2_{W,j}$ and $\sigma^2_{B,j}$ by the classical least squares formulas [20,21]. Note that these formulas are only correct when the design is balanced, i.e. when there is the same number of replicates per series for each concentration level, which should be the case with validation experiments. If not, maximum likelihood estimators are preferred.

One interesting parameter to observe $R_j$. It shows how important is the series to series (or run to run) variance in comparison to the repeatability variance for one concentration level $j$. High values of $R_j$ will either stress a problem with the variability of the analytical procedure, leading to the redevelopment of the method, or stress a lack of number of series (runs) used during the validation process to obtain a reliable estimate of the between-series variance $\sigma^2_{B,j}$.

The $\beta$-TI is then computed for each concentration level of the validation standards. To obtain an accuracy profile, the relative error is plotted versus the validation standards concentration levels and the upper tolerance limits are joined together on one hand and the lower limits on the other hand. Finally, the prespecified acceptance limits are reported on the graph. If this accuracy profile is included in the acceptance limits, the method can be considered as valid over the whole range studied as shown in Fig. 1. If the profile steps outside these limits then the method cannot be considered valid for these concentration levels.

The accuracy profile can therefore be used as a visual decision tool to assess the validity of an analytical method. However, as for every graphical representation, this visual interpretation contains a part of subjectivity. Therefore, we elaborated a global desirability index [22,23] based on the most important validation criteria, which vary from 0 to 1 to increase the objectivity of the decision made. The first step is to assign an individual desirability or index for each of the following validation criterion: dosing range, trueness and precision.

The first index is the Dosing Range Index $I_{\text{DR}}$ (Fig. 1):

$$I_{\text{DR}} = \frac{U_{\text{LOQ}} - L_{\text{LOQ}}}{\text{DR}_{\text{Max}}} \in [0; 1]$$

where:

- $U_{\text{LOQ}}$ is the upper limit of quantitation.
- $L_{\text{LOQ}}$ is the lower limit of quantitation.
- $\text{DR}_{\text{Max}}$ is the difference between the highest and the lowest concentration levels investigated during the validation.

When this index is equal to one this means that the method is valid on the whole range studied. When inferior to one, the method is valid only on a smaller part of the studied range.

The second index is the Trueness Index $I_{\text{T}}$ (Fig. 1). This index illustrates the trueness of the method, and is computed as follow:

$$I_{\text{T}} = \frac{\text{SSB}_{\text{Max}} - \text{SSB}_{\text{Obs}}}{\text{SSB}_{\text{Max}}} \in [0; 1]$$

where:

- $\text{SSB}_{\text{Max}}$ is the sum of the square of the maximum bias at each concentration level investigated by the validation standards included in the dosing range, i.e. $\sum_{j=1}^{m}\lambda^2 = m\lambda^2$, with $\lambda$ the acceptance limit in % and $m$ the number of validation concentration levels included in the dosing range.
- $\text{SSB}_{\text{Obs}}$ is the sum of the square of the observed bias at each concentration level also included in the dosing range.

A Trueness Index close to 1 implies that the method is almost not biased.

The third one is the Precision Index $I_{\text{P}}$ (Fig. 1)

$$I_{\text{P}} = \frac{U_{\text{Area}} + L_{\text{Area}}}{\text{AREA}_{\text{Max}}} \in [0; 1]$$

where:

- $U_{\text{Area}}$ is the area defined by the upper $\beta$-expectation tolerance limits and the upper acceptance limits $+\lambda$ and included between the $L_{\text{LOQ}}$ and the $U_{\text{LOQ}}$.
- $L_{\text{Area}}$ is the area between the lower $\beta$-expectation tolerance limits and the lower acceptance limits $-\lambda$ and included between the $L_{\text{LOQ}}$ and the $U_{\text{LOQ}}$.
- $\text{AREA}_{\text{Max}}$ is the area defined by the upper and lower acceptance limits $\pm\lambda$ and the $L_{\text{LOQ}}$ and the $U_{\text{LOQ}}$, i.e. $2\lambda(U_{\text{LOQ}} - L_{\text{LOQ}})$.

The closer to one is this index the more precise is the method being validated.

Finally to estimate the overall method quality, a desirability index [22,23], called Accuracy Index ($I_A$), is defined by combining the individual criteria desirabilities as the geometric mean of the three indexes:

$$I_A = \sqrt[3]{I_T I_R I_P} \in [0; 1]$$

Indeed, the interest of accuracy profiles is to estimate total error that is expected over the dosing range, therefore this Accuracy Index is a summary of the three previous ones, showing the overall quality of the method being validated over the range studied.

The choice of the response function and the evaluation of a possible matrix effect can be made using these indexes. It is possible either to choose the accuracy profile with the highest Accuracy Index, or depending on the objective of the method to choose the one with the best Range Index if this is the critical parameter or with the highest Trueness Index. As required by the FDA document [4], choosing the simplest mathematical function of the calibration curve can also be easily justified using these indexes. Indeed, this response function will have to be used during routine analysis.

### 3.2. Chemicals and solvents

Urea (>99.5%), and sulphuric acid (>95%) were supplied by VWR (Leuven, Belgium) and L-Histidine, diacetylmonoxime (>98% pure), thiosemicarbazide, iron (III) chloride hexahydrate (>98%) were purchased from Sigma (Schnelldorf, Germany). All reagents were of analytical grade. Ultrapure water (conductivity <0.065 $\mu S\,cm^{-1}$) was used to prepare the different solutions.

### 3.3. Sample analysis

A diacetyl-monoxime colorimetric method, previously used for determination of very low urea concentration in seawater and biological samples [24], was adapted to dose urea in the samples. Samples and standards (180 μL) were placed in 96-wells plates and 15 μL of reagent A (0.850 g diacetylmonoxime in 25 mL of deionised water together with 1.0 mL thiosemicarbazide solution, 0.95 g in 100 mL deionised water) was added in each well. Forty-eight microlitre of reagent B (30 mL concentrated sulphuric acid diluted to 53.5 mL with distilled water together with 50 μL ferric chloride solution, 0.15 g in 10 mL distilled water) was added immediately after. The plate, wrapped in aluminium foil, was then placed on a stirring plate at 90 shaking/min during 10 min. The samples were then placed at 85 °C during exactly 1 h 15 min. After heating, the samples were cooled with ice during exactly 5 min. After 15 min the absorbance of the 96-wells plate was measured at 520 nm with a Spectra Max 190 apparatus (Molecular Devices, Sunnyvale, CA, USA).

### 3.4. Standard solutions

Each day, stock solutions of urea were prepared at a concentration of $1\,g\,L^{-1}$ in ultrapure water. A stock solution of L-Histidine at 10 mM in ultrapure water was also prepared each day. Both of these solutions were used to prepare newly calibration and validation standards during validation step.

Calibration and validation standards in L-Histidine were daily prepared by diluting the adequate volume of stock solution of urea in order to obtain four concentration levels of urea: 6.0, 12.0, 38.0 and 67.0 μM for both types of standards. Blank samples consisted in L-Histidine samples.

Three series ($p$) of analysis were performed for the validation step. Each calibration and validation samples were analyzed twice ($n = 2$) both at four concentration levels ($m = 4$).

### 3.5. Computations

The validation results of the urea method were obtained by using e.noval version 2.0 software (Arlenda, Liège, Belgium).

## 4. Results and discussion

In order to illustrate this methodology it was applied to the validation of a colorimetric method for the determination of urea in transdermal iontophoretic extracts. Each step of the computations is detailed for the first concentration level. Non-invasive techniques, such as reverse iontophoresis, are under investigation in numerous pathologies to improve comfort of patients for the monitoring of drugs and endogenous substances [25,26]. Indeed, the movements of molecules and ions created by the iontophoretic current allow not only to administer molecules but also to extract substances of clinical interest through the skin [26]. Reverse iontophoresis is based on the existence of a relationship between the quantities of analytes extracted through the skin and the concentration of the blood compartment. For instance, the monitoring of glucose in diabetes using such techniques allows minimizing frequency of hypo- and hyperglycemia episodes [27]. Another possible use of iontophoresis could be to detect renal failure from patient with high risks of developing the disease or to control haemodialysis sessions by determining the concentration of urea in the blood. The actual routine test dedicated to this problem is made by blood sampling, which may lead to infections and increase risks of anemia finally restricting frequency of control [28]. A continuous evaluation of urea level during dialysis could be helpful for the practitioner to control the efficacy of the dialysis session. Transdermal iontophoretic extracts of urea could be used for such evaluation [29].

This paper is part of a study, which investigates the potential use of reverse iontophoresis for monitoring the concentration of urea in patients suffering of renal failure by providing an adequate quantitative analytical method for dosing urea in the extracts. Indeed, the major limitation of reverse iontophoresis consists in the low analysis detection required. Therefore, sensible and reliable methods have to be developed to allow the detection and quantitation of small concentrations of the targeted molecules. To achieve this, analytical methods should be validated in order to demonstrate their ability to quantify. To have guarantees that the analytical method can fulfill its objective, which is to quantify urea in transdermal iontophoretic extracts, the previously described approach was applied.

During the optimization of the operating conditions of the analytical procedure, preliminary studies have demonstrated the absence of matrix effect by comparing results obtained with external calibration curve and internal calibration. The diacetylmonoxime colorimetric method for the determination of urea in transdermal iontophoretic extracts was then validated using accuracy profiles based on $\beta$-expectation tolerance intervals ($\beta = 95\%$) including measurement of total error. The acceptance limits were settled at $\pm 15\%$ following regulatory requirements [4,8,30]. For the validation step, three series ($p$) of two replicates ($n$) with four concentration levels ($m$) ranging from 6.0 to 67.0 μM were used for calibration standards as well as for validation standards.

### 4.1. Response function

The response function of an analytical procedure is, within the range selected, the existing relationship between the response (signal) and the concentration (quantity) of the analyte in the sample [2,4,5,8,13,30,31]. In the present study, eight different response functions were tested and accuracy profiles build for each of them as shown in Fig. 2. The four indexes introduced in the previous section were computed for each accuracy profile and are summarized and sorted according to their Accuracy Index in Table 1. As can be seen by the accuracy profiles, five of the tested models appear to answer the objective of this study: the quadratic regression model, the linear regression after square root transformation, the weighted $1/X$ quadratic model, the weighted $1/X$ linear model and the weighted $1/X^2$ quadratic regression model. However, for the square root model and the weighted $1/X$ linear model there is a higher bias at either the lower and upper region of the range where the $\beta$-TI is reaching the acceptance limits. The three quadratic models examined seemed more adequate. The Indexes in Table 1 allow the classification and comparison of all of these profiles in order to make an objective decision about the selection of the adequate response function. Indeed from this table only the three quadratic models allow to quantify over the whole range studied as their Dosing Range Index is equal to 1. Furthermore when looking at the Accuracy Indexes the quadratic model is the best one and is also simpler than the weighted quadratic regression models. It is therefore the one selected. Indeed this model answers our objective, which is the accurate quantitation of urea over the range of 6.0–67.0 μM. In Table 2, the equations of the quadratic model for each series can be found. Table 3 gives the calculated concentrations of the validation standards with this selected model.

### 4.2. Trueness

Trueness refers to the closeness of agreement between a conventionally accepted value and a mean experimental one [2,5,13,31]. For the first level, the relative bias was computed according to Eq. (3):where:

$$\mu_{T1} = 6.0\,\mu\text{M} \quad \text{and}$$

$$\hat{\mu}_1 = \frac{\sum_{j=1}^{2} \hat{x}_{i1,\text{calc}}}{np} = \frac{37.28}{6} = 6.2133\,\mu\text{M}$$

$$\text{bias}_1(\%) = \frac{6.2133 - 6.0}{6.0} \times 100 = 3.555 \cong 3.6\%$$

All other results for trueness were computed similarly. As can be seen in Table 2, the method trueness is acceptable since the relative bias did not exceed the value of $\pm 15\%$, irrespective to the concentration level.

### 4.3. Precision

The precision of the method was estimated by measuring repeatability and intermediate precision at four concentration levels as described in the protocol (see Section 3.4). The variance of repeatability and time dependent intermediate precision as well as the relative standard deviation (R.S.D.) were calculated from the estimated concentrations [2,4,5,13,16,31]:

Table 1
Indexes of the different regression models tested during the validation phase sorted by accuracy index

| Figure no. | Model | Accuracy index | Dosing range index | Precision index | Trueness index |
|---|---|---|---|---|---|
| 2 (a) | Quadratic regression | 0.7431 | 1.000 | 0.4216 | 0.9734 |
| 2 (b) | Linear regression after SQUARE ROOT transformation | 0.7350 | 0.9995 | 0.4127 | 0.9627 |
| 2 (c) | Linear regression | 0.7284 | 0.9376 | 0.4337 | 0.9505 |
| 2 (d) | Weighted ($1/X$) quadratic regression | 0.7202 | 1.000 | 0.3790 | 0.9856 |
| 2 (e) | Weighted ($1/X$) linear regression | 0.7186 | 0.9980 | 0.3844 | 0.9674 |
| 2 (f) | Weighted ($1/X\hat{}2$) quadratic regression | 0.6457 | 1.000 | 0.2729 | 0.9866 |
| 2 (g) | Weighted ($1/X\hat{}2$) linear regression | 0.5885 | 0.7744 | 0.2750 | 0.9572 |
| 2 (h) | Linear regression after LOGARITHM transformation | 0.5099 | 0.7025 | 0.1973 | 0.9562 |

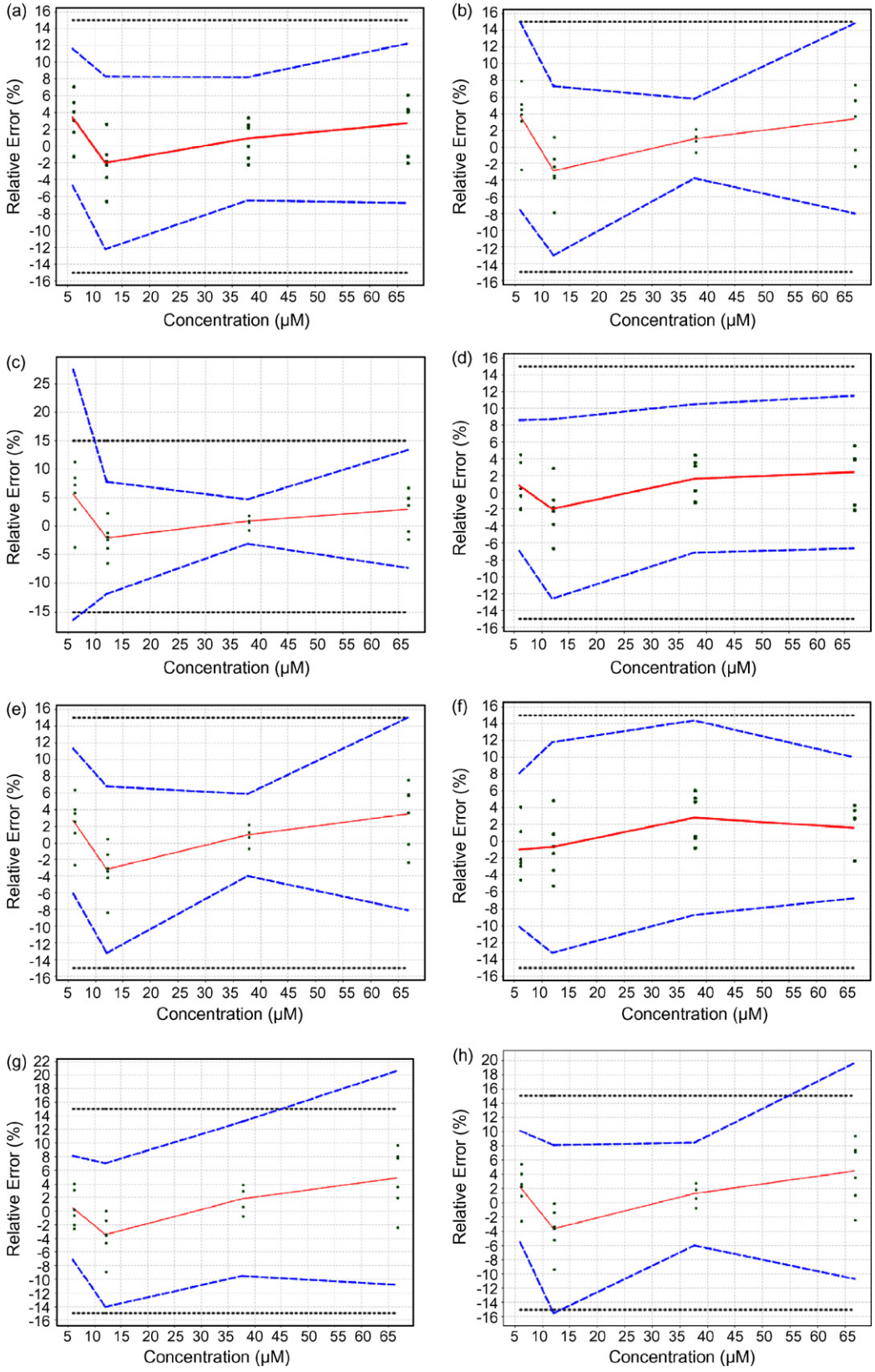


Fig. 2. Accuracy profiles obtained for the determination of urea in transdermal iontophoretic extracts using (a) quadratic regression model; (b) linear regression after square root transformation model; (c) the simple linear model; (d) weighted $1/X$ quadratic regression model; (e) weighted $1/X$ linear model; (f) weighted $1/X^2$ quadratic regression model; (g) weighted $1/X^2$ linear regression model and (h) linear regression after logarithmic transformation model. The continuous lines are the relative bias, the dotted lines are the ±15% acceptance limits and the dashed lines are the upper and lower relative 95%-expectation tolerance limits. The dots represent the relative back-calculated concentrations of the validation standards.

Table 2
Validation results of the method for the determination of urea in transdermal iontophoretic extracts

| Response function ($p=3$; $n=2$) | Quadratic regression model | | |
|---|---|---|---|
| | Calibration range ($m=4$): 6–67 μM | | |
| | Series 1 | Series 2 | Series 3 |
| $x^2$ | 1.10E−05 | 1.05E−05 | −5.75E−06 |
| $x$ | 1.12E−02 | 1.14E−02 | 1.30E−02 |
| Intercept | 3.96E−03 | 7.46E−03 | −3.19E−03 |
| $r^2$ | 0.9995 | 0.9992 | 0.9992 |

| Trueness ($p=3$; $n=2$) | Relative bias (%) |
|---|---|
| 6.0 μM | 3.6 |
| 12.0 μM | −2.0 |
| 38.0 μM | 0.9 |
| 67.0 μM | 2.7 |

| Precision ($p=3$; $n=2$) | Repeatability (R.S.D.%) | Intermediate precision (R.S.D.%) |
|---|---|---|
| 6.0 μM | 2.9 | 2.9 |
| 12.0 μM | 2.4 | 3.2 |
| 38.0 μM | 1.9 | 2.4 |
| 67.0 μM | 3.4 | 3.4 |

| Accuracy ($p=3$; $n=2$) | β-Expectation tolerance limits (%) |
|---|---|
| 6.0 μM | [−4.6; 11.7] |
| 12.0 μM | [−12.2; 8.3] |
| 38.0 μM | [−6.5; 8.2] |
| 67.0 μM | [−6.7; 12.2] |

| Linearity ($p=3$; $n=2$) | β-Expectation tolerance limits (%) |
|---|---|
| Range (μM) | [6.0; 67] |
| Slope | 1.03 |
| Intercept | −0.365 |
| $r^2$ | 0.998 |
| LOD (μM) | 2.5 |
| LOQ (μM) | 6.0 |

$p$: number of series of analysis; $m$: number of concentration levels; $n$: number of replicates per series and per level of concentrations.

$$\begin{aligned}\mathrm{MSM}_1 &= \frac{1}{p-1}\sum_{i=1}^{p} n_{i1}(\bar{x}_{i1.,\mathrm{calc}} - \bar{x}_{.1.,\mathrm{calc}})^2 \\ &= \frac{1}{3-1}\Big[2(6.35-6.2133)^2 \\ &\quad +2(6.16-6.2133)^2 + 2(6.14-6.2133)^2\Big] \\ &= 0.026900\end{aligned}$$

$$\begin{aligned}\mathrm{MSE}_1 &= \frac{1}{\sum_{i=1}^{p} n_{i1} - p}\sum_{i=1}^{p}\sum_{k=1}^{n_{i1}}(x_{i1k,\mathrm{calc}} - \bar{x}_{i1.,\mathrm{calc}})^2 \\ &= \frac{1}{6-3}[(6.26-6.35)^2 + (6.44-6.35)^2 \\ &\quad +(6.20-6.16)^2 + (6.11-6.16)^2 + (6.33-6.14)^2 \\ &\quad +(5.94-6.14)^2] \\ &= \frac{0.0964}{3} = 0.032133\end{aligned}$$

Table 3
Back-calculated concentrations of the validation standards using the quadratic regression model as response function

| Series | Introduced concentration (μM) | Back-calculated concentration (μM) | Relative bias (%) |
|---|---|---|---|
| 1 | 6.0 | 6.26 | 4.2 |
| 1 | 6.0 | 6.44 | 7.2 |
| 2 | 6.0 | 6.20 | 3.2 |
| 2 | 6.0 | 6.11 | 1.8 |
| 3 | 6.0 | 6.33 | 5.3 |
| 3 | 6.0 | 5.94 | −1.1 |
| 1 | 11.98 | 12.31 | 2.8 |
| 1 | 11.98 | 11.88 | −0.9 |
| 2 | 11.98 | 11.73 | −2.1 |
| 2 | 11.98 | 11.21 | −6.4 |
| 3 | 11.98 | 11.78 | −1.7 |
| 3 | 11.98 | 11.55 | −3.6 |
| 1 | 37.68 | 39.00 | 3.5 |
| 1 | 37.68 | 38.67 | 2.6 |
| 2 | 37.68 | 38.54 | 2.3 |
| 2 | 37.68 | 36.88 | −2.1 |
| 3 | 37.68 | 37.73 | 0.1 |
| 3 | 37.68 | 37.21 | −1.3 |
| 1 | 66.66 | 69.64 | 4.5 |
| 1 | 66.66 | 65.94 | −1.1 |
| 2 | 66.66 | 69.62 | 4.4 |
| 2 | 66.66 | 70.79 | 6.2 |
| 3 | 66.66 | 69.49 | 4.2 |
| 3 | 66.66 | 65.38 | −1.9 |

As $\mathrm{MSE}_1 > \mathrm{MSM}_1$ then:

$$\begin{aligned}\hat{\sigma}^2_{W,1} &= \frac{1}{pn_{i1}-1}\sum_{i=1}^{p}\sum_{k=1}^{ni1}(x_{i1k,\mathrm{calc}} - \bar{x}_{.1.,\mathrm{calc}})^2 \\ &= \frac{1}{6-1}[(6.26-6.2133)^2 + (6.44-6.2133)^2 \\ &\quad + +(6.20-6.2133)^2 + (6.11-6.2133)^2 \\ &\quad + +(6.33-6.2133)^2 + (5.94-6.2133)^2] \\ &= \frac{0.152733}{5} = 0.030546\end{aligned}$$

$$\hat{\sigma}^2_{B,1} = 0$$

$$\hat{\sigma}^2_{IP,1} = \hat{\sigma}^2_{W,1} + \hat{\sigma}^2_{B,1} = 0.030546 + 0 = 0.030546$$

$$\begin{aligned}\mathrm{R.S.D.}_{IP,1} &= \frac{\hat{\sigma}_{IP,1}}{\mu_{T1}} \times 100 \\ &= \frac{\sqrt{0.030546}}{6.0} \times 100 = 2.912902 \cong 2.9\%\end{aligned}$$

Intermediate precision relative standard deviations for the other concentration levels were computed using the same methodology. The precision of the bioanalytical method was also very satisfactory; the R.S.D. values for urea were not exceeding 3.4% as illustrated in Table 2.

### 4.4. Accuracy

Accuracy takes into account the total error, i.e. the sum of systematic and random errors, related to the test result

[1,2,5,13,16,31]. For the first concentration level, accuracy is computed as here after:

$$R_1 = 0;$$

$$B_1 = \sqrt{\frac{0+1}{2\times 0+1}} = 1;$$

$$\nu = \frac{(0+1)^2}{(0+(1/2)/3-1)^2+1-(1/2)/3\times 2} = 4.8;$$

$$Q_t\left(4.8; \frac{1+0.95}{2}\right) = 2.60313$$

Therefore the relative $\beta$-expectation tolerance interval for this first concentration level is computed from Eq. (2):

$$\left[3.555 - 2.60313\times\sqrt{1+\frac{1}{3\times 2\times 1^2}}\times 2.912902; 3.555 + 2.60313\times\sqrt{1+\frac{1}{3\times 2\times 1^2}}\times 2.912902\right]$$

$$\Leftrightarrow$$

$$[-4.63521\%; 11.74521\%]$$

$$\Leftrightarrow$$

$$[-4.6\%; 11.7\%]$$

This interval guarantees that with an error of 5% when this analytical procedure is used to measure a sample of a concentration of 6.0 μM in urea, 95 times out of 100 the results obtained will be between 5.7 μM and 6.7 μM of urea.

The relative $\beta$-expectation tolerance interval for the remaining concentration levels of the validation standards were computed using the same procedure and are given in Table 2. As shown in Fig. 2a, the upper and lower $\beta$-expectation tolerance limits (%) did not exceed the acceptance limits settled at ±15% for each concentration level. Consequently, the method can be considered as valid over the concentration range investigated [4,8,30].

This also confirms the choice of the quadratic regression curve for calibration curve, since this model allowed the accurate quantitation of urea in the whole concentration range.

### 4.5. Linearity of the results

The linearity of an analytical method is its ability within a definite range to obtain results directly proportional to the concentrations (quantities) of the analyte in the sample [5,13,31]. Therefore, a linear model was fitted on the calculated concentrations of the validation standards for all series as a function of the introduced concentrations. The regression equation is presented in Table 2. In order to prove method linearity, the absolute $\beta$-expectation tolerance interval was applied [13]. The linearity of the diacetyl-monoxime colorimetric method was also demonstrated since the $\beta$-expectation tolerance limits were included in the absolute acceptance limits for the whole concentration range investigated as shown in Fig. 3.

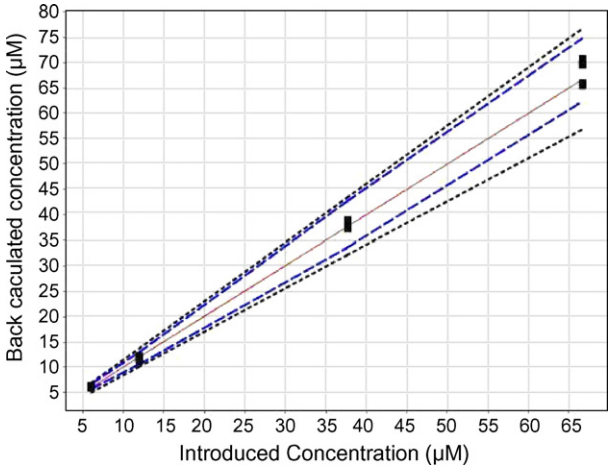


Fig. 3. Linear profile of the analytical method for the determination of urea in transdermal iontophoretic extracts using a quadratic regression model as calibration curve. The dashed limits on this graph correspond to the accuracy profile, i.e. the $\beta$-expectation tolerance limits expressed in absolute values. The dotted curves represent the acceptance limit at ±15% expressed in the concentration unit. The continuous line is the identity line $y = x$.

### 4.6. Detection and quantitation limits

In the present study, the limit of detection (LOD) was estimated using the mean intercept of the calibration model and the residual variance of the regression [20]. The LOD of the colorimetric method was equal to 2.5 μM.

On the other hand, the limit of quantitation (LOQ) of the analytical method is the lowest amount of the targeted substance in the sample, which can be quantitatively evaluated under the experimental conditions used and with a well defined accuracy [2,5,13,31] therefore taking into account simultaneously systematic and random errors. From the accuracy profile in Fig. 2a, the LOQ was fixed to 6.0 μM. Indeed the $\beta$-expectation tolerance interval was perfectly included into the acceptance limits for this lowest concentration level.

## 5. Conclusion

The validation of an analytical method is the ultimate step before its implementation in routine use. In order to guarantee the ability of the method to provide accurate results, the tolerance interval has to be used to assess its validity. The accuracy profile uses this approach by providing a visual decision tool. The way to compute accuracy profiles has been detailed in this paper. Moreover, to make objective the decision process a global desirability index based on the most important validation criteria has been developed allowing a methodical comparison and classification of the profiles.

This approach and its improvements have been successfully applied to the validation of a diacetyl-monoxime colorimetric method for the determination of urea in transdermal iontophoretic extracts. Indeed, the accuracy profile gave the guarantee that at least 95% of the future results obtained with the validated method will be within the ±15% acceptance limits over the whole concentration range studied.

## Acknowledgement

Research grant from the Walloon Region and the European Social Fund to one of the author (E. Rozet) is also gratefully acknowledged (First Europe Objective 3 project no. 215269).

## References

[1] International Conference on Harmonization (ICH) of Technical Requirements for registration of Pharmaceuticals for Human Use, Topic Q2 (R1): Validation of Analytical Procedures: Text and Methodology, Geneva, 2005.
[2] 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.
[3] Commission Decision 2002/657/EC, Off. J. Eur. Commun., L 221 (2002) 8.
[4] Guidance for industry: Bioanalytical Method Validation, US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), Rockville, May 2001.
[5] Ph. Hubert, P. Chiap, J. Crommen, B. Boulanger, E. Chapuzet, N. Mercier, S. Bervoas-Martin, P. Chevalier, D. Grandjean, Ph. Lagorce, M.C. Laparra, M. Laurentie, J.C. Nivet, Anal. Chim. Acta 391 (1999) 135–148.
[6] H.T. Karnes, G. Shiu, V.P. Shah, Pharm. Res. 8 (1991) 421–426.
[7] C. Hartmann, J. Smeyers-Verbeke, D.L. Massart, R.D. McDowall, J. Pharm. Biomed. Anal. 17 (1998) 193–218.
[8] V.P. Shah, K.K. Midha, J.W.A. Findlay, H.M. Hill, J.D. Hulse, I.J. McGilveray, G. McLay, K.J. Miller, R.N. Patnaik, M.L. Powell, A. Tonelli, C.T. Viswanathan, A. Yacobi, Pharm. Res. 17 (2000) 1551–1557.
[9] B. Boulanger, W. Dewe, P. Chiap, J. Crommen, Ph. Hubert, J. Pharm. Biomed. Anal. 32 (2003) 753–765.
[10] M. Feinberg, B. Boulanger, W. Dewe, Ph. Hubert, Anal. Bioanal. Chem. 380 (2004) 502–514.
[11] S.P. Boudreau, J.S. McElvain, L.D. Martin, T. Dowling, S.M. Fields, Pharm. Technol. (2004) 54–66.
[12] J.W. Lee, V. Devanarayan, Y.C. Barrett, R. Weiner, J. Allinson, S. Fountain, S. Keller, I. Weinryb, M. Green, L. Duan, J.A. Rogers, R. Millham, P.J. O'Brien, J. Sailstad, M. Khan, C. Ray, J.A. Wagner, Pharm. Res. 23 (2006) 312–328.
[13] 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) 579–586.
[14] Process Analytical Technology (PAT) Initiative, US Food and Drug Administration, Rockville, MD, 2004; http://www.fda.gov/cder/OPS/PAT.htm.
[15] Food and Drug Administration, International Conference on Harmonization (ICH): Quality Risk Management (Q9), Fed. Regist., vol. 71 (2006) 32105–32106.
[16] The Fitness for Purpose of Analytical Methods, A Laboratory Guide to Method Validation and Related Topics, Eurachem, Teddington, 1998.
[17] R. Mee, Technometrics 26 (1984) 251–254.
[18] G. Gonzalez, M.A. Herrador, Talanta 70 (2006) 896–901.
[19] F. Satterthwaite, Psychometrika 6 (1941) 309–316.
[20] J.C. Miller, J.N. Miller, Statistics for Analytical Chemistry, fifth ed., Ellis Horwood, New York, 2000, p. 96.
[21] P. Dagnelie, Théorie et méthodes statistiques vol. 1 et 2, Les presses agronomiques de Gembloux, Gembloux, 1994.
[22] G.C. Derringer, R. Suich, J. Qual. Technol. 12 (1980) 214–219.
[23] G.C. Derringer, Qual. Prog. June (1994) 51–58.
[24] P.F. Mulvenna, G. Savidge, Estuarine, Coastal Shelf Sci. 34 (1992) 429–438.
[25] A. Sieg, R.H. Guy, M.B. Delgado-Charro, Clin. Chem. 50 (2004) 1383–1390.
[26] B. Leboulanger, R.H. Guy, M.B. Delgado-Charro, Physiol. Meas. 25 (2004) R35–R50.
[27] M.J. Tierney, J.A. Tamada, R.O. Potts, R.C. Eastman, K. Pitzer, N.R. Ackerman, S.J. Fermi, Ann. Med. 32 (2000) 632–641.
[28] A. Farkas, R. Vamos, T. Bajor, N. Mullner, A. Lazar, A. Hraba, Exp. Eye Res. 76 (2003) 183–192.
[29] I.T. Degim, S. Ilbasmis, R. Dundaroz, Y. Oguz, Pediatr. Nephrol. 18 (2003) 1032–1037.
[30] V.P. Shah, K.K. Midha, S. Dighe, I.J. McGilveray, J.P. Skelly, A. Yacobi, T. Layloff, C.T. Viswanathan, C.E. Cook, R.D. McDowall, K.A. Pittman, S. Spector, J. Pharm. Sci. 81 (1992) 309–312.
[31] 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.