Forensic Science International 251 (2015) 32–39

Contents lists available at ScienceDirect

Forensic Science International

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

What a validation strategy means for the quantitation of cocaine and heroin?

Laurence Dujourdy [a,*], Céline Charvoz [b], Marion Dalmasso [b], Anne-Béatrice Dufour [c]

[a] National Forensic Institute, Central Service of Laboratories, 31 Avenue Franklin Roosevelt, 69134 Ecully Cedex, France
[b] National Forensic Institute, Forensic Laboratory of Lyon, 31 Avenue Franklin Roosevelt, 69134 Ecully Cedex, France
[c] University of Lyon, F-69000, Lyon, University Lyon 1, CNRS UMR CNRS 5558, Laboratory of Biometry and Evolutive Biology, 69622 Villeurbanne France

ARTICLE INFO

Article history:
Received 19 May 2014
Received in revised form 8 March 2015
Accepted 17 March 2015
Available online 24 March 2015

Keywords:
Statistical validation
Accuracy profile
Weighted linear regression
Illicit drugs
Chromatography
Uncertainty

ABSTRACT

A method of separation by gas chromatography with a flame ionisation detector was developed for quantifying cocaine and heroin in powders seized by law enforcement. The method was validated by studying parameters of calibration, trueness, precision based on trueness error (or systematic bias) and random error. Total error, which is the combination of these errors, verified its adequacy with the objectives fixed by the analyst. Accuracy profile proved to be an efficient decision tool for that purpose.

Results obtained with weighted regression model were analysed and allowed to conclude that the method enables quantitation of heroin and cocaine in powders on 2–100% concentration (w/w) range with acceptance limits fixed at 10% and a risk at 5%.

The possible sources of uncertainty were evaluated and measurement of their contribution was integrated. The combined standard uncertainty and expanded uncertainty were determined.

1. Introduction

In this study, a method for quantifying heroin and cocaine and their main adulterants in a same run with a relatively short time (e.g. less than 6 min) was sought. Adulterants which are frequently encountered with these illicit drugs are also quantified to feed the National Substance Identification System (SINTES): a monitoring of substances seized by law-enforcement or collected among drug users managed by the French Monitoring Centre for Drugs and Drug Addiction (OFDT) [1].

Reliable analytical methods are needed to be in compliance with national and international regulations in all areas. The same demand should be applicable to forensic laboratories and appropriate measures must be taken to ensure data with required quality level is provided to the customer. Method validation is one of them. Validation is also the ultimate phase before the routine use of the method, great care must therefore be taken to check it. The purpose of this validation was to show that the developed analytical method was suitable for its intended use. A global approach based on measurement of total error and accuracy profile as a decision tool was selected [2]. This approach was initially produced to help professionals from the pharmaceutical industry to validate their analytical procedures and it is now widely used in various fields such as food safety [3], toxicology [4] and building trade [5] for instances.

A method of analysis is a dynamic process that goes through several successive steps called lifecycle of the method [6]. This paper presents the selected and optimised analytical method and some key steps of the validation procedure, which is besides well described in literature [2,6–10]. Key steps can be summarised as follows:

1. Definition of the desired criteria i.e. validation domain of the analytical method in terms of concentration levels and its objective in terms of acceptance limits.
2. Definition of the experimental design for calibration and validation steps.
3. Analysis of the calibration model.
4. Drawing of the accuracy profile i.e. validation of the criteria defined in point 1.

At the end of the study the enlarged uncertainty of the analytical procedure was determined.

* Corresponding author. Tel.: +33 4 72 86 84 74; fax: +33 4 72 86 86 90.
E-mail address: laurence.dujourdy@interieur.gouv.fr (L. Dujourdy).

http://dx.doi.org/10.1016/j.forsciint.2015.03.009
0379-0738/

2. Material and methods

2.1. Analytical method and reagents

2.1.1. Preparation of internal standard solution and analytical samples

A solution at 0.7 mg/ml of tetracosane in ethanol/chloroform (50/50, v/v) was prepared. It constituted the internal standard solution. Samples were directly weighed in a GC vial, between 2.0 and 2.2 mg for cocaine and 3.0 and 3.2 mg for heroin. Balances used were Mettler Toledo® MT5 analytical microbalances, capacity – 5.1 mg and readability – 0.001 mg.

One ml of internal standard solution was added using certified 1000 μl fixed-volume pipettes (Eppendorf®). Certified laboratory glassware (10 ml volumetric flasks) was used to prepare solutions.

Microbalances and micropipettes are certified annually.

Vials were sealed with a PTFE/rubber cap and vortexed.

This validation approach consists in using two kinds of samples called standards: the calibration standards used to set up the calibration model while the validation standards are used to estimate the precision, trueness and accuracy of the method.

Calibration and validation standards are samples of known concentrations prepared with reference materials. They are either pure heroin/cocaine dissolved in the internal standard solution, or completed with a simulated matrix (cutting agents and impurities).

Matrix for cocaine was made of phenacetin, lidocaine, levamisole, diltiazem, hydroxyzine, caffeine, procaine and mannitol (see Table 1).

Matrix for heroin was composed of caffeine, paracetamol and noscapine. Noscapine was reported in one paper as an adulterant in illicit heroin samples and it is also frequently encountered in high proportion among impurities [11].

The selection of cutting agents was based on statistics extracted from the national database fed by police forensic laboratories [12].

2.1.2. Chromatographic conditions

The choice of a method has to be made in accordance with the specific issue being addressed. Because gas chromatography coupled with flame ionisation detector (GC-FID) has the advantages to be easily implemented, it was selected for quantitative analyses in routine. In the analytical procedure of our laboratory, identification of the target compounds is always performed by GC–MS prior to quantitation with GC-FID. Any possible interference between constituents would therefore be spotted at this stage.

Gas chromatography measurements were performed on a ThermoFisher® Trace GC Ultra GC-FID using Chrom-Card v2.4.1 software equipped with a Triplus AS autosampler and a split/splitless injector. A DB-1 capillary column, 100% apolar (10 m × 0.18 mm × 0.18 μm) was used. The injection liner is a Thermo splitless focusliner 5 mm i.d. × 8 mm o.d. × 105 mm length for 70 mm needle containing quartz wool.

The carrier gas was helium at a constant flow of 0.6 ml/min and a split ratio of 100:1 with an injection volume of 1.0 μL. The injector temperature was set to 280 °C, with an initial oven temperature of 200 °C held for 1.5 min, then ramped at 26 °C/min to 300 °C and held there for 1.06 min. Total chromatographic run was 6.5 min long. The detector temperature was set at 300 °C, hydrogen and air flows were set at respectively 35 ml/min and 350 ml/min.

Peak areas of cocaine, heroin and tetracosane were monitored. With these chromatographic conditions and the use of a less alcoholic solvent mixture, no degradation of cocaine or heroin was observed. Especially, no products of heroin transacetylation in presence of paracetamol (monoacetylmorphine formation and acetylparacetamol) were detected [13]. No interaction between illicit drug, internal standard and adulterants was observed. Selectivity is adequate.

2.1.3. Definition of the desired criteria

The objective of the method was to quantify the mass percent fraction of molecules of interest in a powder specimen with a matrix. It also had to fulfil the following requirements: the covered concentration range had to vary from 2 to 100% (w/w) and no matrix effect on calibration model shall be observed.

The risk for the customer, police or justice services in our case, i.e. false negative result, is defined as the risk to wrongly accept inaccurate results. By contrast, the risk for the supplier, here the forensic laboratory, i.e. false positive result, is defined as the risk to wrongly reject correct results [14].

Analytical results must be provided to the law enforcement and justice with guaranties that every expected results obtained in routine analysis will be as close as possible to the true unknown value of the analyte in the specimen. In statistical terms, this is reflected in a proportion of obtained results Π higher than a minimum proportion included in the interval $\mu \pm \lambda$; λ being the expectation limit, i.e. the limit of the deviation between the true value (μ_T) (unknown) and obtained results (μ being the mean of the results).

It is important to note that these limits could be imposed by a regulation or by the customer, and can sometimes vary depending on the concentration level.

Based on our experience in illicit drugs analysis, for this quantitative method, we estimated dispersion as 10% (acceptance limits $\lambda = \pm 10\%$) with a 5% risk (95% of results will be expected to be within these limits λ). The model is selected to meet objectives of the method in routine analysis.

2.2. Experimental design

When editing the validation protocol, one of the first questions to rise is: "how many experiments are required?". Due to time and other constraints such as reference materials' cost, this often

Table 1
List of the reference materials used during the validation of cocaine and heroin quantitation by GC-FID.

Name	Description	Purity	Supplier
Cocaine base	Illicit drug	>98.5%	Lipomed®
Heroin base	Illicit drug	>98.5%	Lipomed®
Noscapine	Impurity of heroin	>98.5%	Lipomed®
Phenacetin	Analgesic (removed in France from 1980s)	≥98%	Sigma Aldrich VWR®
Lidocaine	Local anaesthetic	≥99%	Sigma Aldrich®
Paracetamol	Analgesic	NA	Lipomed®
Caffeine	Stimulant	≥99%	Sigma Aldrich VWR®
Diltiazem	Calcic inhibitor, bradycardic	≥99%	Sigma Aldrich®
Levamisole	Pest	>99%	Sigma Aldrich®
Procaine	Local anaesthetic	>98.5%	Lipomed®
Hydroxyzine	Antihistamine	≥98%	Sigma Aldrich®
Mannitol	Polyol	NA	VWR

translates as: "what is the absolute minimum required?" When referring to calibration experiments, this relates to the number of standards that need to be analysed and the amount of replicates at each concentration level.

An experimental design was selected in compliance with the French Society of Pharmaceutical Science and Technology (SFSTP) commission guidelines on validation of quantitative analytical procedures [2,7,8]. The protocol took into account quantification procedures constraints and some peculiarities (matrix effect, number of calibration standards and their concentrations, etc.).

In this case, 3 series of calibration standards with and without matrix, at 3 concentration levels were prepared (see Table 2). A concentration level corresponds here to the ratio between the concentration of targeted substance and the internal standard concentration.

For cocaine: level 1 is 0.057 (0.04 mg/ml cocaine/0.7 mg/ml tetracosane) – level 2 is 1.430 (1.00 mg/ml cocaine/0.7 mg/ml tetracosane) – level 3 is 3.14 (2.20 mg/ml cocaine/0.7 mg/ml tetracosane).

For heroin: level 1 is 0.071 (0.05 mg/ml heroin/0.7 mg/ml tetracosane) – level 2 is 2.140 (1.50 mg/ml heroin/0.7 mg/ml tetracosane) – level 3 is 4.57 (3.20 mg/ml heroin/0.7 mg/ml tetracosane).

Each pair of series (with and without matrix) being weighed by a different operator.

Validation standards are samples reconstituted in the matrix or in any other reference material with known concentration (true values) set by consensus and used to validate the analytical procedure.

Each validation point was prepared twice by the operator. Values for the replicates resulted of the averaged measurements of two samples, because in routine work, results are calculated in that manner.

In total 63 concentrations were measured:

- 2×18 calibration standards: without matrix (SE) and with matrix (SEM),
- 27 validation standards (considering the average value obtained from 2 determinations, carried out on separate samples of a given powder).

Before starting the validation process, calibration itself should be examined, as it can affect the strategy to adopt in order to achieve an optimal design. For this reason, it is important to know if it is plausible that calibration curve (a) could be linear, (b) could pass through the origin and (c) is not affected by matrix (a selectivity study ensures that there is no interference between peaks of interest: cocaine, heroin and tetracosane). That's why matrices with cutting agents usually encountered in cocaine and heroin have been simulated.

Matrix components (even if their peaks are separated from the analyte of interest's) can influence the recovery of the analyte. "Inadequate recovery" is demonstrated as a systematic error (method bias) [15].

Table 2
Experimental design for calibration validation standards.

Calibration standards						
Series	3 Concentrations/2 replicates					
	With matrix			Without matrix		
1	••	••	••	••	••	••
2	••	••	••	••	••	••
3	••	••	••	••	••	••
Validation standards						
Series	3 Concentrations 3 Replicates (average) With matrix					
1	$\bullet^{(A)}\bullet^{(A)}\bullet^{(A)}$			$\bullet^{(A)}\bullet^{(A)}\bullet^{(A)}$		
2	$\bullet^{(A)}\bullet^{(A)}\bullet^{(A)}$			$\bullet^{(A)}\bullet^{(A)}\bullet^{(A)}$		
3	$\bullet^{(A)}\bullet^{(A)}\bullet^{(A)}$			$\bullet^{(A)}\bullet^{(A)}\bullet^{(A)}$		

3. Results and discussion

3.1. Regression model selection and its evaluation for calibration

The relationship between the response (i.e. the chromatographic signal) and the concentration (amount) of the analyte in the sample is a very important parameter that must be considered in the validation of an analytical method.

When building a model, it is necessary to express the y – instrumental response as a function of the x concentration using a mathematical relationship. The simplest relationship is described by the equation of the linear regression $y = a + bx + e$, where b is the slope, a the intercept with the y-axis and e the residuals.

When variations from environmental sources are introduced, this model becomes more complex. Let i be an assay linked to a concentration x_i and a response y_i.

For each i assay, a weighting w_i is introduced [16] such that residuals can be written:

$$e_i = \sqrt{(w_i)}(y_i - (bx_i + a)) \quad (1)$$

with $w_i = \frac{1}{x^2}$

In classical regression models, the y-instrumental response is continuous, normally distributed and the error e is supposed random. These models are meaningful only when some hypotheses are verified: constant variance, no outlier and a symmetric distribution [17,18].

The statistical assessment of normality and regressions will be examined later on.

Note that response functions of an analytical instrument depend on its nature (separation instruments, spectrophotometry, absorption, etc.) and the model is selected only after carrying validation testing. With our GC-FID, available response functions were linear regressions without transformation [19].

Calculations were performed using open source software R [20] and a lab-made web application called VALMETH using CGIwithR, an R package that allows a server to run R scripts and to build HTML pages. Collected data from validation experiments are inserted in the tool, parameters are fixed by the user and a validation report is generated. Functionalities of VALMETH were inspired by various validation softwares especially e-Noval® [21], statistics were found in [2,6–10,14]. VALMETH is a decision support tool but it does not replace statistical process to meet the validation objectives. To make a statistical decision, the level of significance $\alpha = 0.05$ was classically retained.

In supplementary files, data used for calculations are provided. R commands are available on demand.

The response functions were evaluated using 3 calibration curves built with calibration standards with and without matrix at 3 concentration levels (see Section 2.2).

Several regression models have been tested using VALMETH tool in order to (i) study the relationship between cocaine or heroin concentration (mg/ml) and the analytical response (ratio between peak areas of cocaine/heroin and IS (tetracosane)) and (ii) select the most appropriate one.

A selection of various models is available in VALMETH (simple linear regression or linear weighted regression ($1/x^2$ and $1/x$),

linear regression after logarithmic or square root transformation, simple quadratic regression or weighted quadratic regression) based on 4 indexes:

- determination interval (DI),
- trueness index (TI),
- accuracy index (AI),
- precision index (PI),

Criteria are ranked in descending order of interest DI, TI, AI and PI (values from 0 to 1), and the out-coming index is used to select the most appropriate model.

The selection of the regression model which will be used for routine calibration had to fit the following chosen criteria:

- a DI equal to 1, meaning that high and low quantification limits correspond to low and high calibration levels (0.057 and 3.14 for cocaine and 0.071 and 4.57 for heroin), to be able to quantify these two illicit drugs from 2 to 100% (w/w),
- with respect to the previous criteria, with and without matrix, meaning that there is no matrix effect, hence future routine calibrations could be prepared with only analyte reference material and internal standard solution, without addition of a simulated matrix, which saves time,
- same model for heroin and cocaine, in order to quantify them in a same run (within the GC software restrictions).

With these conditions, the most relevant model was found to be weighted linear regression, with and without matrix.

3.1.1. Weighted linear regression

In linear models, response is supposed to be normally distributed (Gaussian). Despite the reduced number of individuals ($n = 6$), a Shapiro–Wilks test [17,18] for each series was performed to check this hypothesis.

For each series p-values are significant enough (>0.05) compared to the risk, fixed at the beginning of the analysis, not to reject the hypothesis of a Gaussian response.

Research of regression parameters (a and b) consists of minimising residual sum of squares, i.e. minimisation of Eq. (2):

$$\sum_{i=1}^{6} e_i^2 = \sum_{i=1}^{6} (\sqrt{w_i} y_i - \sqrt{w_i}(bx_i + a))^2 \quad (2)$$

with e_i residual associated to an assay i (see Eq. (1)).

Equation of weighted regression for series n° 3 is: $\hat{y}_i = 0.6764x_i - 0.0029$ which is the model value for response i.

A variance decomposition analysis (one-factor ANOVA [17]) was applied and results are displayed in Table 3.

For the sake of clarity, only variances linked to the model and residual variances within series were reported here. These two variances enable Fisher's test leading to the validation of the model (relevant or not). Critical probabilities (p-values) being very low, it means that "neutral" model or simple linear regression (no X-variable effect on Y-response) is rejected and weighted model is accepted. Difference between simple and weighted linear regressions seemed to be negligible but sufficient to adjust data.

Finally, when the model is selected, its validation is possible only if residuals are normally distributed and have the same variance within series (see Table 4).

Results show that the hypotheses of normality of residuals and equality of variances cannot be rejected because p-values are sufficiently greater compared to the chosen risk (5%) i.e. higher than the risk chosen at the beginning of the analysis. Conditions are therefore verified.

3.2. Validation

When the choice of the model from statistical theory point of view is secured, coherence between experimental objectives, especially in terms of determination interval and accuracy, is required. Concentrations of validation standards (VS) were calculated from the model, for each concentration level. From these predicted inverse concentrations, the mean relative bias and the high and low tolerance interval limits considering intermediate precision standard deviation can be determined. Tolerance interval is the domain where future proportion of observations will fall within the predefined acceptance $[-\lambda,+\lambda]$ at β level (fixed here at 5%). Afterwards, an accuracy profile can be built with these data (Figs. 1 and 2) [14].

The tolerance limits, obtained via the weighted $1/x^2$ linear regression model used as calibration curve, fitted our requirements (acceptance limits ± 10%) for all the concentration levels and for both analytes (cocaine and heroin).

The mean relative bias does not exceed 3% in absolute value for cocaine and heroin.

The recovery rate, for cocaine, varies from 96.79% to 100.2% between the levels (0.057 and 3.140). For heroin, it varies from 97.27% to 101.2% between the levels (0.071 and 4.57).

The GC-FID analytical method is thus valid.

The quantitation limits are the extreme values that can be quantified with a given accuracy.

The lower limit of quantitation (LLOQ) is obtained by calculating the lowest concentration below which accuracy or tolerance limits are out of acceptance limits. In our case, accuracy profile is within the acceptance limits, so the lowest level of calibration (corresponding to 0.04 mg/ml for cocaine and 0.05 mg/ml for heroin) is the LLOQ.

Moreover, with the weighted $1/x^2$ linear regression model, expected measurements will be included in the acceptance limits with a guaranty of 95%.

Linearity domain covers concentration ranges from 0.057 to 3.14 (i.e. 0.04 to 2.20 mg/ml) and from 0.071 to 4.57 (i.e.0.05 to 3.20 mg/ml) for cocaine and heroin respectively; corresponding to w/w content ranges 1.8% to 100% of 2.2 mg sample and 1.6% to 100% of 3.2 mg sample.

3.3. Calculation of uncertainty measurement

As mentioned above, measurement error can be defined as the difference between the measured value x and the true value X

Table 3
Results of the decomposition of response variance.

Series	Variance of estimated model (ddl = 1)	Estimated residual variance (ddl = $n-2$)	Fisher statistics	p-Value
1	1.7709	0.0004	4829.4	2.6E−07
2	1.8055	0.0003	5767.7	1.8E−07
3	1.7229	0.0009	1975.7	1.5E−06

For cocaine experiments ($n = 6$ points by series – 3 concentrations and 2 replicates).

Table 4
Final tests on the normality of residuals (Shapiro-Wilks) and the identity of variances (Bartlett [17]).

Series	Normality of residuals	Equality of variances
1	0.7997 (0.0584)	0.7810 (0.6767)
2	0.8429 (0.1377)	0.5203 (0.7709)
3	0.8374 (0.1241)	0.0254 (0.9874)

The first value gives the statistics of the test; the value in brackets gives the p-value.

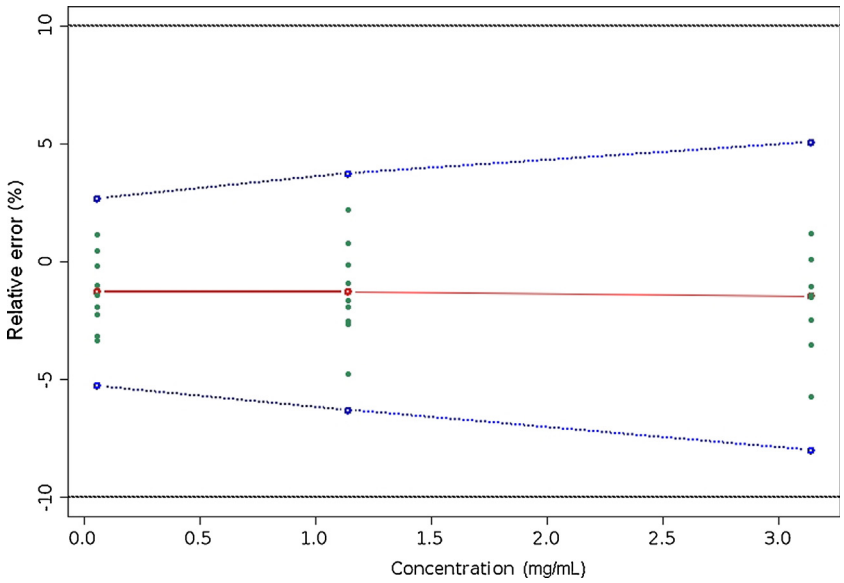

Fig. 1. Accuracy profile of cocaine quantitation using GC-FID analytical method obtained from VALMETH with weighted regression model at 3 concentration levels: 0.057; 1.14; 3.14 (cocaine/tetracosane). Black dotted lines bound the interval of acceptance; blue dotted lines bound the interval of tolerance calculated from the standard deviations of intermediate standard deviation for every level. The red line represents the average relative error. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

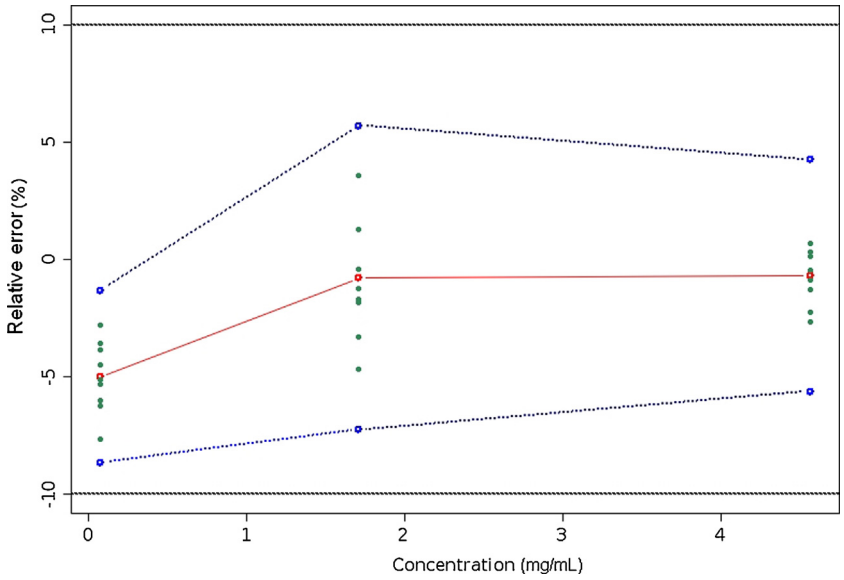

Fig. 2. Accuracy profile of heroin quantitation using GC-FID analytical method obtained from VALMETH with weighted regression model at three concentration levels: 0.071; 1.71; 4.57 (heroin/tetracosane). Black dotted lines bound the interval of acceptance; blue dotted lines bound the interval of tolerance calculated from the standard deviations of intermediate standard deviation for every level. The red line represents the average relative error. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

which is, in general, unknown. There are two types of errors: random and systematic. Uncertainty reflects scientific attempts to estimate the magnitude of random error. In the absence of systematic error, uncertainty defines an interval around the measured value which includes the true value with given probability.

The uncertainty on the result may arise from many possible sources (sampling, matrix effects and interferences, environmental conditions, uncertainties of masses and volumetric measuring equipment, purity of reference material, etc.) [22]. It can be reduced to two main source categories: sample heterogeneity uncertainty and analytical uncertainty [23–25].

The overall combined standard uncertainty was calculated using the following formula:

$$u_c^2 = u_h^2 + u_a^2 \tag{3}$$

with u_c the combined standard uncertainty, u_h the uncertainty coming from heterogeneity of the material and u_a the analytical uncertainty.

Results of heroin or cocaine quantitation are given in weight percentage $\%_{(ww)}$ using the formula:

$$\%_{(ww)} = \frac{100 \times (C \times V)}{W} \tag{4}$$

Table 5
Calculation of the different uncertainties coming from analytical process and heterogeneity of cocaine.

Concentration level	Weight-%	Uncertainty from analytical process			Uncertainty from heterogeneity of material
		$u^2_{\%_{(w/w)}}$	u^2_{round}	u^2_{Ref}	u^2_h
1	1.80	0.0009	0.0208	2.70×10^{-5}	0.6936
2	36.40	0.4650		1.10×10^{-2}	0.0017
3	100.00	1.3585		8.33×10^{-2}	0.0002

Table 6
Calculation of the different uncertainties coming from analytical process and heterogeneity of heroin.

Concentration level	Weight-%	Uncertainty from analytical process			Uncertainty from heterogeneity of material
		$u^2_{\%_{(w/w)}}$	u^2_{round}	u^2_{Ref}	u^2_h
1	1.60	0.0006	0.0208	2.13×10^{-5}	0.8778
2	37.50	1.2263		1.17×10^{-2}	0.0016
3	100.00	2.5581		8.33×10^{-2}	0.0002

with C concentration of target compound (mg/ml) obtained with the calibration curve (response function), W sample weight in mg and V volume of internal standard solution used to dilute the sample (V = 1 ml).

The weight percentage is rounded up to the nearest whole number which introduces another uncertainty called rounding uncertainty (u_{round}).

The purity of the reference standard is potentially an additional uncertainty source called standard uncertainty (u_{Ref}).

Overall analytical uncertainty is given by Eq. (5):

$$u_a^2 = u^2_{\%_{(ww)}} + u^2_{round} + u^2_{Ref} \quad (5)$$

$$\text{with } u^2_{\%_{(ww)}} = \left(\frac{100 \times C \times V}{W}\right)^2 \left[\left(\frac{u(C)}{C}\right)^2 + \left(\frac{u(V)}{V}\right)^2 + \left(\frac{u(W)}{W}\right)^2\right]$$

When calculating measurement uncertainty two types, correctly known as type A and type B must be assessed and combined [26]. $u(C)$ is quantified by calculation of the estimated standard deviation from the set of repeated measurements (type A evaluation). Its value is obtained from the validation study using absolute bias at each concentration level.$u(V)$ is estimated from data provided in calibration certificates (type B evaluation). The calibration certificate for the automatic pipette with a fixed volume (1 ml) used states the measurement uncertainty over the range at ±0.6 μL at a 95% confidence level:

$$u(V) = \frac{I_{V\text{calib}}}{2} = \frac{0.6}{2} = 0.3\ \mu\text{L} \quad \text{so} u(V) = 0.0003\ \text{mL} \quad \text{for} V = 1\ \text{mL}.$$

$u(W)$ is estimated from data provided in calibration certificates (type B evaluation). The calibration certificate for Mettler Toledo® microbalances states the measurement uncertainty over the range at ±0.0295 mg at a 95% confidence level:

$$u(W) = \frac{I_{W\text{calib}}}{2} = \frac{0.02965}{2} = 0.01475\ \text{mg}$$

$u^2_{round}() = \left(\frac{\omega/2}{\sqrt{3}}\right)^2$ (rectangular distribution) with ω = measuring range [22]. ω = 0.5 because it is limited to a maximum of 0.5%: $u^2_{round} = 0.02083$.

Purities of cocaine and heroin standard must be greater than or equal to 99% therefore there is 1% relative deviation on each level concentration. Then (rectangular distribution) with $\omega_{Ref} = \%_{(ww)} \times 0.01$. For each level, the weight percentage is given in Table 5 for cocaine and Table 6 for heroin.

When p distinct replicates are taken from different parts of the specimen, and analysed separately, the composition of the

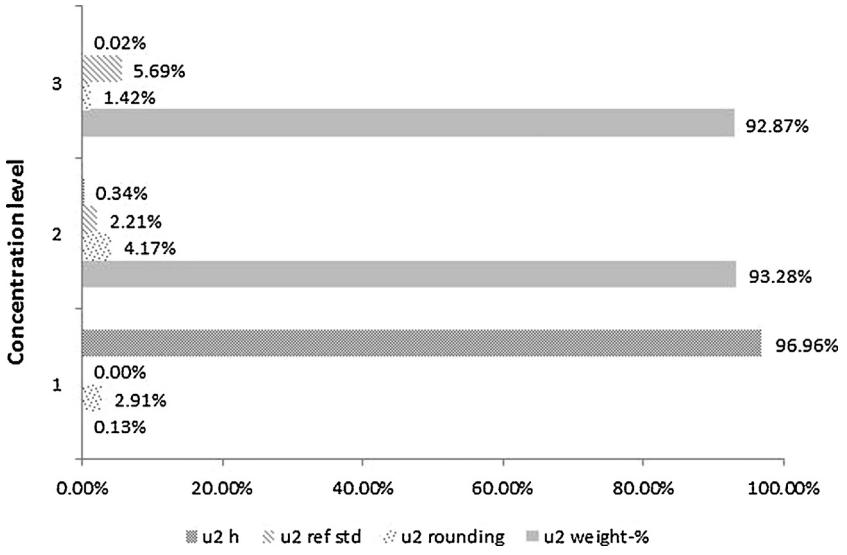

Fig. 3. Contribution of various types of uncertainties at the three levels of concentrations for cocaine quantitation.

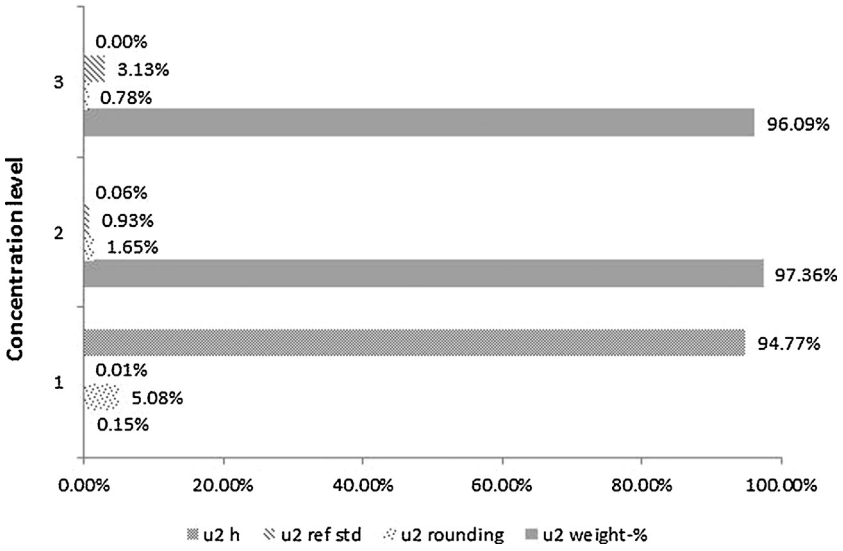

Fig. 4. Contribution of various types of uncertainties at the three levels of concentrations for heroin quantitation.

specimen can be determined by taking the mean value of these p separate measurements, and the relative standard uncertainty on the mean value can be used to determine the best estimate of the uncertainty due to target heterogeneity [22,27]:

$$u_h^2 = \left(\frac{S_p}{\sqrt{p}}\right)^2 \quad (6)$$

where s_p is the relative standard deviation determined by the analysis of p samples taken from the same material. It is understood that material has been homogenised (using mortar and pestle) and s_p corresponds to the relative standard deviation of the sampling after comminution.

The different contributions are shown in Figs. 3 and 4 for cocaine and heroin quantitation.

For low concentration level (level 1), the uncertainty is mainly influenced by heterogeneity of the material. For medium and high concentration levels, analytical uncertainty with calculation of concentration is the main contributor to the overall uncertainty. Uncertainties from rounding and reference materials are negligible.

The expanded uncertainty U is obtained by multiplying u_c by a coverage factor k. The choice of the factor k is based on the level of confidence desired. For 95% confidence level, k is usually set to 2 [22].

4. Conclusion

Validation of quantitative method is essential in analytical chemistry and acutely pertinent in forensic science. It calls upon statistical methods based on strong hypotheses such as normality, heteroscedasticity (homogeneity of variances). Interpretation of results only makes sense if analytical and statistical techniques are properly handled.

This combination has been illustrated with quantification of cocaine and heroin, in real conditions (small dataset). A specific linear model was exploited here: weighted regression based on accuracy, trueness and precision criteria. GC-FID method was validated over the 1.8–100% range for cocaine determination, and 1.6–100% for heroin.

In a practical perspective, this validation strategy can be summarised in the 4 following steps:

(1) selection of acceptance limits in compliance with expected method use,
(2) definition of the experimental design and choice of a regression model for the calibration,
(3) analysis of results with this calibration model,
(4) for each concentration level, calculation of tolerance limits and set up of the accuracy profile. A decision on the validity of the analytical method is then taken.

We carried out the evaluation of uncertainties and their contribution. The result showed that for low levels, the contribution of heterogeneity is maximal and for other levels, analytical uncertainty is the main contributor to the overall error.

The analytical method validated using this approach has been accredited by the French national accreditation body (accreditation n° 1-2322, available on http://www.cofrac.fr. And all the computations are available in VALMETH using the R framework inside.

Other information, such as chromatograms and detailed results of statistical analysis, is provided in supplementary data, datasets used for this study as well.

Acknowledgements

Thanks to Ronan Maron and Frédéric Flores for their contributions to the development, the improvements and updates, of VALMETH statistical tool used for the validation of results in the Forensic framework. Thanks to Virginie Ladroue for the proofreading. Thanks to the reviewers for their relevant comments improving this paper in a substantial way.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.forsciint.2015.03.009.

References

[1] French Monitoring Centre for Drugs and Drug Addiction (OFDT), SINTES dispositive. Available from: URL: http://www.ofdt.fr/ofdtdev/live/reserve/sintes.html (cited 03.02.15).
[2] Ph. Hubert, J.-J. Nguyen-Huu, B. Boulanger, E. Chapuzet, P. Chiap, N. Cohen, P.-A. Compagnon, W. Dewe, M. Feinberg, M. Lallier, M. Laurentie, N. Mercier, G. Muzard,

C. Nivet, L. Valat, Harmonization of strategies for the validation of quantitative analytical procedures. A SFSTP proposal – part I, J. Pharm. Biomed. Anal. 36 (3) (2004) 579–586.
[3] M. Laurentie, J.M. Delmas, The Accuracy Profile: A Tool for Choosing an Analytical Method and Evaluating Its Validity, EuroReference No. 3, 2010, pp. 9–11 Available from: URL: http://www.afssa.fr/euroreference/numero3/PN50I0.htm.
[4] F. Smith, M. Augsburger, V. Varlet, Accuracy profile validation of a new analytical method for propane measurement using headspace-gas chromatography–mass spectrometry, J. Anal. Toxicol. 38 (2) (2014) 73–79.
[5] C. Marlet, G. Lognay, Development and validation by accuracy profile of a method for the analysis of monoterpenes in indoor air by active sampling and thermal desorption-gas chromatography–mass spectrometry, Talanta 82 (3) (2010) 1230–1239.
[6] M. Feinberg, Validation of analytical methods based on accuracy profiles, J. Chromatogr. A 1158 (2007) 174–183.
[7] P. 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, E. Rozet, Harmonization of strategies for the validation of quantitative analytical procedures. A SFSTP proposal – part II, J. Pharm. Biomed. Anal. 45 (1) (2007) 70–81.
[8] P. Hubert, J.J. Nguyen-Huu, B. Boulanger, E. Chapuzet, N. Cohen, P.A. Compagnon, W. Dewé, M. Feinberg, M. Lallier, M. Laurentie, N. Mercier, G. Muzard, L. Valat, E. Rozet, Harmonization of strategies for the validation of quantitative analytical procedures. A SFSTP proposal – part III, J. Pharm. Biomed. Anal. 45 (1) (2007) 82–96.
[9] P. Hubert, J.J. Nguyen-Huu, B. Boulanger, E. Chapuzet, N. Cohen, P.A. Compagnon, W. Dewé, M. Feinberg, M. Laurentie, N. Mercier, G. Muzard, L. Valat, E. Rozet, Harmonization of strategies for the validation of quantitative analytical procedures: a SFSTP proposal part IV. Examples of application, J. Pharm. Biomed. Anal. 48 (3) (2008) 760–771.
[10] 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, The SFSTP guide on the validation of chromatographic methods for drug bioanalysis: from the Washington Conference to the laboratory, Anal. Chim. Acta 391 (1999) 135–148.
[11] S. Klemenc, Noscapine as an adulterant in illicit heroin samples, Forensic Sci. Int. 108 (1) (2000) 45–49.
[12] Office central pour la Répression du Trafic Illicite des Stupéfiants, STUPS OSIRIS, Internal Report, OCRTIS, Nanterre, 2011.
[13] H. Huizer, A.J. Poortman, Some Aspects of the Gas Chromatographic (GC) Analysis of Heroin, United Nations, Scientific and Technical Notes, SCITEC/5, 1989.
[14] M. Feinberg, Labo-Stat. Guide de validation de méthodes d'analyse, Lavoisier TEC&TOC, 2009.
[15] D.B. Hibbert, Systematic errors in analytical measurement results, J. Chromatogr. A 1158 (2007) 25–32.
[16] R.J. Carroll, D. Ruppert, Transformation and Weighting in Regression. Monographs on Statistics and Applied Probability 30, Chapman & Hall, New York, 1988.
[17] J.N. Miller, J.C. Miller, Statistics and Chemometrics for Analytical Chemistry, 6th ed., Prentice Hall, Edinburg, 2010.
[18] NIST/SEMATECH, e-Handbook of Statistical Methods. Available from: URL: http://www.itl.nist.gov/div898/handbook/ (cited 03.02.15).
[19] Thermo Scientific Gas Chromatography Data System. CHROM-CARD S/W for TRACE/FOCUS GCs. Operating Manual PN 317 084 00, 2008 April Available from: http://www.samsi.no/Products/GAS/tmanuals/ChromCard_15_4-2008book.pdf (cited 02.08.13).
[20] R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Available from: URL http://www.R-project.org/ (cited 03.02.15)
[21] Arlenda e-Noval®. Available from: URL http://www.arlenda.com/ (cited 03.02.15).
[22] EURACHEM/CITAC Guide CG 4, Quantifying Uncertainty in Analytical Measurement, 3rd ed., 2012.
[23] L. Zamengo, G. Frison, M. Gregio, G. Orru, R. Sciarrone, Determination of illicit drugs in seized materials: role of sampling and analysis in estimation of measurement uncertainty, Forensic Sci. Int. 208 (2011) 108–123.
[24] L. Zamengo, C. Bettin, G. Frison, M. Gregio, R. Sciarrone, DrugsWorkBook (DWB): a tool for the analysis of illicit drugs in seized materials, Sci. Justice 53 (4) (2013) 415–424.
[25] L. Dujourdy, C. Charvoz, M. Dalmasso, A.-B. Dufour, À la recherche du juste dosage: statistique appliquée à l'analyse de la cocaïne, Ann. Pharm. Fr. 71 (3) (2013) 193–200.
[26] K. Birch, Estimating Uncertainties in Testing, Measurement Good Practice Guide No. 36, British Measurement and Testing Association, Teddington, 2003.
[27] T. Cseszregi, M. Bovens, L. Dujourdy, A. Franc, J. Nagy, Sampling of illicit drugs for quantitative analysis – part III: sampling plans and sample preparations, Forensic Sci. Int. 241 (2014) 212–219.