METHOD VALIDATION BY MEANS OF THE ACCURACY PROFILE: A REALISTIC APPROACH FOR ROUTINE CHEMICAL LABORATORIES

Petru Jitaru[1], *Rachida Chekri*[1], *Paola Fisicaro*[2] *and Thierry Guérin*[1]

[1] Université Paris-Est, Anses, Laboratory for food safety, F-94700 Maisons-Alfort, France. E-mail: petru.jitaru@anses.fr

[2] Laboratoire National de Métrologie et d'Essais (LNE), Department of Biomedical and Inorganic chemistry, 1 Gaston Boissier, 75015 Paris, France. E-mail: paola.fisicaro@lne.fr

Abstract

This paper describes the advantages and applications of a method validation approach based on the construction of an *accuracy profile*. This is currently employed by the laboratory of *Trace Metals and Minerals* of the French Agency for Food, Environmental and Occupational Health & Safety (ANSES). The main analytical performance characteristics obtained for a selection of analytes by means of an in-house developed and validated multi-element ICP-MS method are presented. A brief comparison of the uncertainty calculated by using the GUM recommendations and by ANSES approach is also provided.

Keywords: food chemistry; ICP-MS; accuracy profile; intermediate precision, expanded uncertainty.

1. INTRODUCTION

Method validation is the fundamental step that allows the application of a developed method to routine chemical analysis [1]. It relies generally on the assessment of several analytical performance characteristics in order to verify method's *fit-for-purpose* [2]. Among them, the most important are the precision and the accuracy, which in most cases are asses independently. This means that the method is developed in order to achieve, for instance, repeatability <10% and accuracy within 10% compared to a reference value. Taking into account that accuracy (in terms of relative bias) could be either negative or positive, a global error of 30% could be encountered in this case (see Fig. 1).

Whereas national metrological laboratories utilize primary methods such as isotope dilution-mass spectrometry to provide reference values [3], the approach currently employed by the National Reference Laboratory (NRL) of the French Agency for Food, Environmental and Occupational Health & Safety (ANSES) for the validation of methods for determination of trace metals in foodstuff is based on the accuracy profile [4].

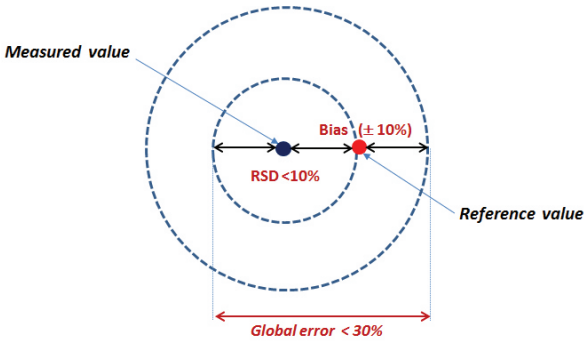

Fig. 1. Individual interpretation of the precision and accuracy of a chemical measurement method (*adapted from [5]*).

The accuracy profile approach relies on a criterion, namely the *acceptability interval* taking into account simultaneously the accuracy and precision of the method. In this respect, a β-*expectation tolerance* (generally 80 to 90%) is *a priori* set for constructing the accuracy profile, which means that the risk of expected results falling outside these limits is below 10-20%. In addition, *method acceptance limits* (λ) are set according to the criteria expected of repeatability and intermediate precision [6]. For multi-element analysis methods, λ is generally set between 20-30 %.

According to the accuracy profile approach, a method is validated when the β-expectation tolerance interval for a given result is comprised within the *a priori* set acceptability limit (λ). A graphical representation of a validated result when using β and λ criteria as defined above is shown in Fig. 2.

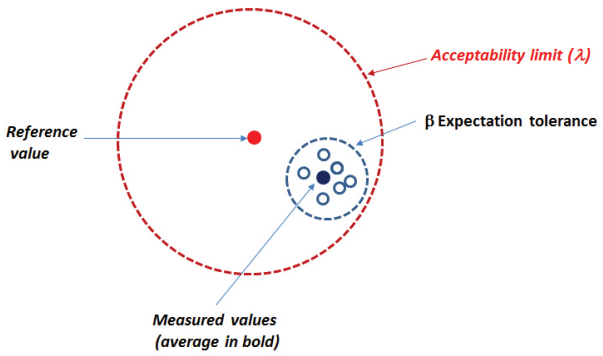

Fig. 2. Example of a valid result when employing the accuracy profile approach (*adapted from* [5]).

In this study, an example of method validation by means of the accuracy profile obtained in real-life laboratory conditions for multi-element analysis is provided. It is worth noting that ANSES' NRL is accredited by COFRAC (*French comity for standardization*) for two multi-element analysis methods, one for four analytes, namely As, Pb, Cd, and Hg (*M1*) that is used for truly routine analyses and one for simultaneous determination of 31 major, minor and trace metals (*M2*) that is employed both for routine and research projects. Both methods are based on (closed) microwave samples digestion (acidic) and quantification by inductively coupled plasma mass spectrometry (ICP-MS) [4].

The main analytical characteristics such as limit of quantification (LOQ), repeatability (S_r), intermediate precision (S_R) and expanded uncertainty (U) obtained by ICP-MS multi-element analysis (M_2) by using the accuracy profile approach are reported. Additionally, a comparison in terms of expanded uncertainty assessed by means of the accuracy profile approach as well as the guide to the expression of uncertainty in measurement (GUM)[7] is provided. For simplicity, although the ANSES method (M_2) is validated for the simultaneous determination of 31 elements, data for As, Pb, Cd and Hg solely are reported. These are the analytes for which maximum levels in foodstuffs are regulated by the European commission [8-11].

2. METHOD VALIDATION BY MEANS OF THE ACCURACY PROFILE

Method validation by means of the accuracy profile relies on the graphical representation of the accuracy generally measured in terms of recovery factor over a range of analyte levels comprised between the limit of quantification (LOQ) and an in-house defined highest concentration (generally lower than the maximum of the calibration curve).

The analysis of certified reference materials (CRM) or spiked samples generally at five concentration levels is carried out in duplicate in different days during a time span of at least three months (different operators are also employed). For each recovery factor (Fig. 3), the tolerance interval is calculated either in positive (Recovery factor+β, %) or negative (Recovery factor -β, %) as described in detail in section 2.1. A typical example of an accuracy profile obtained for five concentration levels when using β and λ criteria as defined above is shown in Fig. 3.

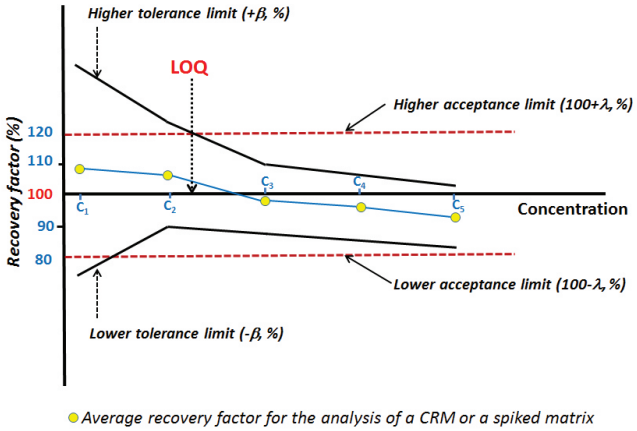

Fig. 3. Typical accuracy profile obtained for five concentration levels. Dotted lines correspond to the acceptance limits whereas the bold (continue) lines correspond to β expectation tolerance (*adapted from* [5]).

It is worth to underline that for the validation of a multi-matrix method, the levels (e.g., C_1 to C_5) in the accuracy profile correspond to different matrices, namely either CRM with different analytes' concentrations or matrices spiked at various levels. For the multi-element method (M_2) developed at ANSES [4], the accuracy profile for 31 analytes was constructed using five genuine (CRM) or spiked matrices at various levels depending on the analyte (see Tables 1a-b).

In order to check if the developed method can be applied for other types of (food) samples, method selectivity/specificity is assessed to ensure that there are no matrix effects. For this purpose, the recoveries factors of different foodstuffs samples spiked with the corresponding analytes at levels within the calibration range are measured. The method is validated only for the matrices that pass the specificity test [4].

Table 1a. Matrix and spiking levels employed for the construction of the accuracy profile for As and Pb in view of ANSES' multi-element method (M_2) validation.

Spiking level[a]	As	Pb
	Matrix	
C_1	Pure water (0.025)	Pure water (0.03)
C_2	Pure water (0.05)	Bovine whole milk (0.06)
C_3	NIST1548a-typical diet[b] (1.24)	NIST 1566b-oyster tissue[b] (1.873)
C_4	TORT-2-Lobster Hepatopancreas[b] (13.1)	Infant formula (5.0)
C_5	Tuna tissue (50.0)	Tuna tissue (10.0)

Table 1b. Matrix and spiking levels employed for the construction of the accuracy profile for Cd and Hg in view of ANSES' multi-element method (M2) validation.

Spiking level[a]	Cd	Hg
	Matrix	
C_1	Bovine whole milk (0.006)	Pure water (0.1)
C_2	Bovine whole milk (0.012)	Bovine whole milk (0.2)
C_3	IAEA-155-whey powder[b] (0.10)	TORT-2 Lobster Hepatopancreas[b] (1.64)
C_4	NIST 1566b-oyster tissue[b] (7.54)	Infant formula (5.0)
C_5	Tuna tissue (10.0)	Tuna tissue (10.0)

[a] spiking level ($\mu g \cdot L^{-1}$) or the genuine concentration (the case of CRM analysis) is reported in the brackets corresponding to each matrix ;

[b] certified reference material (CRM) (analysis without spiking);

Table 2 shows the recoveries factors obtained for the analysis of different matrices for selectivity/specificity assessment in view of validation of ANSES' multi-element method [4].

Table 2. Recovery factors (%) obtained for the measurement of As, Pb, Cd and Hg in 39 spiked food matrices at different levels.[a]

Sample	Recovery factor (%)			
	As	Pb	Cd	Hg
Honey all flowers	111	95	98	91
Acacia honey	108	95	97	89
Lavender honey	106	97	95	90
Porcine kidney	109	97	102	97
Infant formula	102	95	105	95
Breakfast cereals	99	93	107	98

Table 2 (continuation)

Tuna tissue	89	93	103	101
Infant cereals	117	94	113	96
Baby fruit juice	101	97	101	107
Carrot semolina	112	107	114	101
Growth milk powder (2nd age)	107	120	129	105
Plain yoghurt	103	107	110	123
Vegetables-chicken-rice	112	105	114	112
Salmon-spinach blend	109	106	113	100
Milk/fruits based dessert	113	100	102	111
Peaches pineapple blend	107	99	102	98
Growth milk powder (1st age)	105	100	102	97
Drinking water	98	111	112	-
Carrot	99	99	97	101
Multivitamin fruits juice	102	99	101	91
Fresh orange	102	98	102	97
Fresh white grapes	95	101	101	100
Cucumber	99	100	103	99
Chicken (raw)	115	98	107	72
Pork liver	112	99	107	97
Scallop nuts	109	83	122	100
Shrimps	78	97	95	119
Beef liver	114	95	98	85
Trout	81	99	99	101
Tuna tissue	108	95	105	96
Orange jus (no sugar)	110	94	94	104
Spareribs	110	92	96	92
Beef steak	112	88	96	101
Partly skimmed milk yoghurt	96	89	94	99
Semi-skimmed milk	94	90	93	90
Cutlet sautéed chicken	115	-	99	94
Mackerel	-	90	94	91
Ray	-	101	92	95
Whole bovine milk	-	-	92	95

[a] spiking levels were comprised between 1-50 $\mu g \cdot L^{-1}$ depending on analyte and matrix.

2.1. Construction of the accuracy profile

For constructing the accuracy profile, accuracy (bias or recovery factor) is assessed for each analyte at different levels (5) in natural matrices (CRM, if available) or a given spiked matrix. Further, a tolerance interval (β_{TI}) is calculated for each analyte at every level for each the accuracy is assessed. β_{TI} is the interval in which a prediction can be made for a known proportion of the results, which is represented on the accuracy profile (Fig. 3). β_{TI} is calculated using (1).

$$\beta_{TI} = k_{TI} \times s_{TI} \qquad (1)$$

where:

k_{TI} : coverage factor

S_{TI} : standard deviation of the tolerance interval.

For multi-series analysis with the same number of replicates carried out for each series, S_{TI} is calculated by means of (2).

$$S_{TI} = S_R \sqrt{\left(1 + \frac{1}{IJB^2}\right)} \qquad (2)$$

where:
S_R, standard deviation of the intermediate precision (see section 2.3);
I, number of series (days);
J, number of measurement replicates per series;
B, parameter depending on intra- and inter-series standard deviations (3):

$$B = \sqrt{\frac{A+1}{JA+1}} \qquad (3)$$

where A is calculated by (4):

$$A = \frac{s_B^2}{s_r^2} \qquad (4)$$

where:
S_r, S_B, intra- and inter-series standard deviation, respectively (see section 2.3);

k_{TI} is actually the Student statistical test depending on the number of freedom degrees and the expected probability (β) of the tolerance interval (5-6).

$$k_{TI} = t_{v,\frac{1+\beta}{2}} \qquad (5)$$

where:

$$v = \frac{(A+1)^2}{\frac{\left(A+\frac{1}{J}\right)^2}{I-1} + \frac{1-\frac{1}{J}}{IJ}} \qquad (6)$$

It is worth to note that the accuracy profile allows the assessment of several analytical parameter characteristics, such as:

(i) limits of quantification (LOQ) and detection (LD);
(ii) repeatability (S_r) and intermediate precision (S_R);
(iii) combined standard uncertainty (u_c).

The calculation of these parameters is briefly described in the sections 2.2-2.4.

2.2. Assessment of the limit of quantification by means of the accuracy profile. Comparison with the 10×s_0 criterion

Based on the accuracy profile approach, LOQ are assessed as being the lowest validated concentration. Concretely, LOQ is the concentration corresponding to the intersection of the tolerance intervals and the acceptability limits (see Fig. 3). This is conceptually different than the approach relying on the standard deviation (s_0) calculated for repetitive measurement of a number of analytical blanks (10×s_0 criterion)[12,13]. LD is further calculated as 3/10 of LOQ [13].

LOQ obtained for the ICP-MS multi-element analysis method (M_2) developed in ANSES laboratory by means of the two approaches, namely the accuracy profile (AP) and 10×s_0 criterion are given in Table 3.

Table 3. LOQ for ICP-MS determination of As, Pb, Cd and Hg obtained by means of the accuracy profile (AP) approach and 10×s_0 criterion.

Analyte	LOQ (µg·L^{-1}) (n=20)	
	AP	10×s_0
As	0.024	0.140
Pb	0.030	0.069
Cd	0.006	0.006
Hg	0.100	0.570

[a] s_0, standard deviation for the analysis of 20 analytical blanks under intermediate precision condition (one week).

As can be seen, LOQ obtained by the two approaches are significantly different, with higher values obtained by means of 10×s_0 criterion. Given that the same method sensitivity was used by both approaches, higher LOQ values obtained by 10×s_0 criterion is most likely arising from blank contribution. It is worth to underline that when the blank contribution is extremely low, LOQ obtained by both approaches must be comparable. This is confirmed for Cd, for which identical LOQ were obtained by both approaches; it is also worth noting for this analyte a very low LOQ (pg·L^{-1} level) was obtained.

2.3. Assessment of the intermediate precision

The intermediate precision (S_R) is calculated taking into account the intra- and inter-series variability (7).

$$S_R = \sqrt{S_r^2 + S_B^2} \qquad (7)$$

where:

S_r, intra-series standard deviation
S_B, inter-series standard deviation
S_r and S_B are in turn calculated by (8) and (9).

$$S_r^2 = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\left(x_{ij} - \bar{x}_i\right)^2}{I(J-1)} \quad (8)$$

$$S_B^2 = \frac{\sum_{i=1}^{I}\left(\bar{x}_i - \bar{\bar{x}}\right)}{I-1} - s_r^2 \quad (9)$$

where:
I, J, were defined above
$\bar{x}_r^2$, $\bar{x}_B^2$, intra- (x_{rl}) and inter-series (x_{BJ}) average.

S_r and S_R obtained by our multi-element method (M_2) for As, Pb, Cd and Hg for two concentration levels are summarized in Table 4.

Table 4. Repeatability (S_r, %) and intermediate precision (S_R,%) for As, Pb, Cd and Hg determination by ANSES multi-element method (M_2) obtained for two concentration levels, namely < 2×LOQ (L_1) and ≥2×LOQ (L_2), respectively.

Analyte	S_r		S_R	
	L_1	L_2	L_1	L_2
As	16	5	20	12
Pb	10	5	15	8
Cd	12	5	12	10
Hg	10	5	15	10

As can be seen in Table 4, S_R is consistently higher than S_r and their values vary considerably with the measured concentration. These findings confirm the necessity to assess the method intermediate precision for at least two measurement levels for a realistic determination of the expanded uncertainty as described in section 2.4.

2.4. Combined standard uncertainty

Apart from the general purpose of method validation, accuracy assessment has a broader importance in food chemistry. In fact, a food control method must provide measurement uncertainty lower than the maximum standard measurement uncertainty (u_{max}) as calculated by (10) [14].

$$u_{max} = c\left(\alpha^2 + \frac{1}{b^2}\right)^{1/2} \quad (10)$$

where:

c, target concentration (mg·kg^{-1}), e.g., the regulated maximum admissible level of an analyte in a given food matrix;
α, factor depending on the value of c (see Table 5)
b, factor also dependent on c but taking two values, namely 10 for c <0.1 mg·kg^{-1} or 20 for c ≥0.1 mg·kg^{-1}.

Table 5. Values of α to be used for the calculation of u_{max} (10) depending on the concentration level.

c (mg·kg^{-1})	α
≤ 0.05	0.20
0.051-0.5	0.18
0.501-1.0	0.15
1.001-10	0.12
> 10	0.10

The expanded uncertainty (U) is calculated as twice (k=2) the combined uncertainty. Based on the accuracy profile approach, u_c is calculated taking into account two main parameters, namely the bias and S_R. This approach leads to u_c equal to the standard deviation of the tolerance interval (S_{TI}) [6]. As can be seen in (2), for a large number of intra- (*I*) and inter-series (*J*) analyses, S_{TI} can be approximated with S_R. Hence, in practice, S_R is used for the calculation of the expanded uncertainty of a given result (11).

$$U = k \times u_c = 2 \times \frac{S_R(\%)/100}{\sqrt{n}} \times \bar{Y} \quad (11)$$

where:

S_R(%), relative standard deviation in terms of intermediate precision;
$\bar{Y}$, average concentration of *n* replicates measurement.

It is important to note that u_c is dependent on the concentration level used for the construction of the accuracy profile. As previously mentioned in section 2.3, u_c values obtained for levels < 2×LOQ as well as for levels ≥ 2×LOQ (for a given analyte) are considered for calculating the uncertainty by means of (11). It is also worth to underline that u_c calculated based on the accuracy profile does not take into account the uncertainty related to the determination of the recovery factor. Nevertheless, when CRM are analyzed, the associated uncertainty could be used for the assessment of the combined uncertainty and in such a case S_{TI} rather than S_R must be used for u_c determination. Additionally, the uncertainty of other steps of the analytical procedure such as the sampling, samples

preparation, dilution, etc. are not accounted for individually.

A comparison of expanded uncertainty (U, %) calculated by means of the accuracy profile (AP) (11) and GUM recommendations [7] was carried in this study (see Table 6). In the latter case, U was calculated by using Wincert software (V.3.11.2002.0115, Implex, France). For simplicity, only the uncertainties corresponding to the middle level used for construction of the accuracy profile for each analyte were determined (different matrices were analyzed for As, Pb, Cd, Hg determination).

As can be seen from Table 6, the expanded uncertainties calculated using GUM approach are consistently lower compared to those obtained by accuracy profile method. This could be explained by the fact that GUM takes into account individual sources of uncertainty hence avoiding their overestimation. An exception was obtained for Cd, where ANSES' approach to calculate the uncertainty led to a significantly lower value. This is in agreement with the very low LOQ obtained for the same analyte; in this case, GUM approach apparently led to an overestimation of the expanded uncertainty.
More experiments should be pursuit in order to comprehensively compare the two approaches for uncertainty determination, especially for analytes characterized by a very low LOQ.

Table 6. Comparison of expanded uncertainty determination for the measurement of As, Pb, Cd and Hg in various CRM by using the accuracy profile and GUM approaches, respectively.

Analyte ($\mu g \cdot L^{-1}$)[a]	Matrix (CRM)	U (k=2, %)[b]	
		AP	GUM
As (1.21)	NIST1548a CRM (typical diet)	13.0	6.8
Pb (1.77)	NIST 1566b CRM (oyster tissue)	12.7	2.6
Cd (0.11)	IAEA-155 CRM (whey powder)	0.98	15.1
Hg (1.41)	TORT-2 CRM (lobster hepatopancreas)	12.6	7.2

[a] concentration measured in the sample (CRM) extract;
[b] ratio of the absolute uncertainty to the average concentration value (n=5).

3. CONCLUSIONS

Adopting the appropriate validation method for routine multi- (trace) element and multi-matrix analysis is still a matter of debate. The largely employed approach by the analytical chemistry community is based on the assessment of accuracy and precision (mostly repeatability) by the analysis of a limited number of CRM and/or spiked matrices. This approach is not fully acceptable for national reference laboratories which in most cases need to be accredited for the measurement of a large number of major, minor and trace analytes (up to 40-50) in a large panel of matrices. This makes the method validation extremely challenging either from cost and time-efficiency point of view. One realistic method validation approach that can be used by NRL and also by routine laboratories is based on the construction of an accuracy profile. The key point of this approach is the accuracy assessment by measuring various (natural or spiked) matrices for analytes' levels spanning from the LOQ up to the maximum of the linear range. Additionally, the profile is constructed by different analysts during a time span of at least three months. Several analytical parameters such as the quantification limit, precision (both repeatability and reproducibility) as well as the method uncertainty can be calculated out of the accuracy profile. Nevertheless, if a method must be validated for a very large number of analytes and matrices, constructing an accuracy profile per analyte and per matrix is practically impossible. In such a case, the recovery factors for the analysis of all matrices of interest that are not comprised in the accuracy profile are assessed for a full method validation. Along with an appropriate laboratory quality control system, this ensures a long term data reliability, which is a key requirement for a food quality control laboratory.

ACKNOWLEDGMENTS

The authors thank Caroline Oster (LNE) for the help with Wincert calculations of the expanded uncertainty.

REFERENCES

[1] Eurachem Guide (2014), "The Fitness for Purpose of Analytical Methods A Laboratory Guide to Method Validation and Related Topics". https://www.eurachem.org/images/stories/Guides/pdf/MV_guide_2nd_ed_EN.pdf

[2] M. Thomson, S. L. R. Ellison and R. Wood, "Harmonized guidelines for single laboratory validation of methods of analysis", Pure Appl. Chem., vol. 74, no. 5, pp. 835-855, 2002.

[3] S. Sannac, P. Fisicaro, G. Labarraque, F. Pannier, M. Potin-Gautier, "Development of a reference measurement procedure for the determination of methylmercury in fish products", *Accred. Qual. Assur.*, 14, 263-267, 2009.

[4] E. Chevallier, R. Chekri, J. Zinck, T. Guérin, L. Noël, "Simultaneous determination of 31 elements in foodstuffs by ICP-MS after closed-vessel microwave digestion: Method validation based on the accuracy profile", *J. Food Comp. Anal.,* 41, 35-41, 2015.

[5] M. Feinberg, "Validation des méthodes d'analyse quantitatives au moyen du profil d'exactitude", *Techniques de l'Ingénieur*, 2012. http://www.techniques-ingenieur.fr/base-documentaire/mesures-analyses-th1/qualite-au-laboratoire-42497210/validation-des-methodes-d-analyse-quantitatives-au-moyen-du-profil-d-exactitude-p224/

[6] M. Feinberg (2009). Guide de validation des méthodes d'analyse, Lavoisier, France.

[7] Bureau International des Poids et Mesures (BIPM), *GUM: Guide to the Expression of Uncertainty in Measurement.* http://www.bipm.org/en/publications/guides/gum.html

[8] Commission regulation (EU) 2015/1006 amending Regulation (EC) No 1881/2006 as regards maximum levels of inorganic arsenic in foodstuffs : http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32015R1006&from=FR

[9] Commission regulation (EU) 2015/1005 amending Regulation (EC) No 1881/2006 as regards maximum levels of lead in certain foodstuffs. http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32015R1005&from=EN

[10] Commission regulation (EC) No 629/2008 amending Regulation (EC) No 1881/2006 setting maximum levels for certain contaminants in foodstuffs. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:173:0006:0009:EN:PDF

[11] Commission regulation (EU) No 420/2011 amending Regulation (EC) No 1881/2006 setting maximum levels for certain contaminants in foodstuf. http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32011R0420&from=FR

[12] F. Ruggieri, A. Alimonti, B. Bocca, "Full validation and accreditation of a method to support human biomonitoring studies for trace and ultra-trace elements", Trends Anal. Chem. vol. 80, pp. 471-485, 2016.

[13] Commission regulation (EU) 2016/582 amending Regulation (EC) No 333/2007 as regards the analysis of inorganic arsenic, lead and polycyclic aromatic hydrocarbons and certain performance criteria for analysis. http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0582&from=FR

[14] French regulation NF EN 13804 "Foodstuffs - Determination of elements and their chemical species - General considerations and specific requirements", May 2013.