ANNALS OF LABORATORY MEDICINE

Working Towards Accreditation by the International Standards Organization 15189 Standard: How to Validate an In-house Developed Method an Example of Lead Determination in Whole Blood by Electrothermal Atomic Absorption Spectrometry

Carine Garcia Hejl, Pharm.D.[1], Jose Manuel Ramirez, Pharm.D.[2], Philippe Vest, Pharm.D.[1], Denis Chianea, Pharm.D.[1], and Christophe Renard, Pharm.D.[2]

Laboratory of Biochemistry, Percy Military Hospital[1], Clamart; Val de Grâce Military Hospital[2], Paris, France

Laboratories working towards accreditation by the International Standards Organization (ISO) 15189 standard are required to demonstrate the validity of their analytical methods. The different guidelines set by various accreditation organizations make it difficult to provide objective evidence that an in-house method is fit for the intended purpose. Besides, the required performance characteristics tests and acceptance criteria are not always detailed. The laboratory must choose the most suitable validation protocol and set the acceptance criteria. Therefore, we propose a validation protocol to evaluate the performance of an in-house method. As an example, we validated the process for the detection and quantification of lead in whole blood by electrothermal absorption spectrometry. The fundamental parameters tested were, selectivity, calibration model, precision, accuracy (and uncertainty of measurement), contamination, stability of the sample, reference interval, and analytical interference. We have developed a protocol that has been applied successfully to quantify lead in whole blood by electrothermal atomic absorption spectrometry (ETAAS). In particular, our method is selective, linear, accurate, and precise, making it suitable for use in routine diagnostics.

Key Words: Validation, Lead, Atomic absorption spectrophotometry

---

Accreditation is a tool for medical laboratories to demonstrate their competence and to ensure the delivery of accurate and reliable results. More than 70 accreditation bodies are recognized by the international laboratory accreditation cooperation (www.ilac.org). Accreditation for clinical reference laboratories can be voluntary like in Finland and Canada or mandatory like in France and Korea [1-5]. Clinical laboratories can be accredited by International Standards Organization (ISO) 17025 or ISO 15189. ISO 15189 specifies the requirements of the quality management system for medical laboratories. Validation or verification of each test method is one of the requirements for accreditation (ISO 15189 section 5.5.2) [1].

This manuscript compared various international recommendations to propose a validation protocol for in-house methods. This methodology was applied to quantify lead in whole blood samples by electrothermal atomic absorption spectrometry (ETAAS) [6-11].

The validation was performed on an M series electrothermal atomic absorption spectrometer (Thermo Scientific, Marietta, OH, USA). The temperature program consisted of four stages:

ANNALS OF LABORATORY MEDICINE

Hejl CG, et al.
Quantification of lead in whole blood

drying (95°C, 10 sec and then 110°C, 25 sec), pyrolysis (700°C, 30 sec), atomization (1,200°C, 4 sec), and cleaning (2,500°C, 4 sec). The equipment used a graphite tube with integrated L'VOV platform. The radiation source was a mono-elementary hollow cathode lamp ($\lambda = 283$ nm). The Zeeman effect was used for background correction. Argon was used as the purge gas. All the chemicals used were of analytical reagent grade. Atomic absorption spectrometry standard solutions for lead (titrisol grades from Merck; Fontenay-sous-Bois, France) were used to build the calibration curves. They were prepared from a stock solution of 1 g/L by successive dilutions with reverse-osmosis purified water (Merck-millipore, Fontenay-sous-Bois, France).

Each sample was ten-fold diluted in a matrix modifier (2.5 g/L of $NH_4H_2PO_4$ prepared in 0.5% Triton X-100 [v/v], 0.065% HNO_3 [v/v]). Matrix-spiked calibration was performed. Quantification was performed by the integration of absorbance over the time of the transient absorption signal. Internal control samples used were the ones that were commercially available at the Quebec National Public Health Department (INSPQ). Our laboratory participated in inter-laboratory comparison program organized by INSPQ. This program had 6 cycles per year.

The fundamental parameters tested were selectivity, calibration model, precision, accuracy (and uncertainty of measure-

ment), contamination, stability, reference interval, analytical sensitivity, and specificity. The latter two parameters were expressed as the detection limit and interference, respectively [6-9].

The acceptance or rejection criteria were set prior to validation according to published data, specific legislative and regulatory requirements, and professional judgment. To be validated, each parameter had to meet the acceptance criteria. Results and statistical analysis are detailed in Tables 1 and 2.

To test selectivity, at least six blank samples were analyzed to ensure that there was no interference [7]. The wavelength used was specific to lead and allowed the user to distinguish between it and other components in the sample prior to quantification. No interference was noted.

The calibration range was between 10 and 1,000 µg/L. To assess the relationship between concentration and response, five calibration standards were run in duplicate, as recommended by the Australian national association of testing authorities (NATA), US food and drug administration (FDA), French Normalization Association (AFNOR), and Feinberg [7-11]. Five rounds of calibration were performed [10, 11]. The model was found to be linear.

The least squares regression method was used to establish the relationship between the analyte concentration (x) and in-

Table 1. Statistical tests of calibration model

ANOVA				
Sources		Sum of squares (ss)	Degree of freedom (df)	Mean squares (ms): ss/df
Regression residual		SS_{reg}	1	ms_{reg}
		SS_{resid}	IJ-1	ms_{resid}
	Non linearity	$ss_{nl} = ss_{resid} - ss_r$	I-2	ms_{nl}
	Repeatability J=5 calibration sets I=6 calibration standards $y_{i,j}$ = signal measured	$ss_r = \Sigma(ssy_{i,j})$ (e.g. $SSy_{i,j}$ for each calibration standard level)	J (I-1)	ms_r
F Test				
	H0	H1	Inference	F-statistic
Slope existence	Slope = 0 i.e. does not exist	Slope ≠ 0 i.e. slope exists	Reject H0 i.e. the slope is something other than zero	$F = ms_{reg}/ms_{resid}$
Goodness of the model	Model fits the data	Model does not fit the data	Accept H0 i.e. the model is valid	$F = ms_{reg}/ms_r$
C Test				
	H0	H1	Inference	C-statistic
C-test	All variances are equal	At least, one variance value is significantly larger than the other variance values	Accept H0	$C = \dfrac{Max\ (var\ (y_{i,j}))}{\sum\limits_{i=1}^{N} (var\ (y_{i,j}))}$

Table 2. Statistical results

Linearity			
	Results	Acceptance criteria	References
Slope of the regression model	0.00543		
Intercept of the regression model	0.07776		
Correlation coefficient	0.989		
Slope existence (Fisher)	1053 (HS)	F (0.05; 1; 23) = 4.49	7, 10, 11
Goodness of the model (Fisher)	0.039 (NS)	F (0.05; 5; 20) = 3.26	7, 10, 11
Variances homogeneity (Cochran)	0.328 (NS)	C (0.05; 5; 4) = 0.616	7, 10, 11

Repeatability				
	Mean	CV (%)	Acceptance criteria	References
QC 41 µg/L	44	7.6	≤7.6%	9, 13
QC 114 µg/L	120	1.3	≤4.5%	9, 13
QC 236 µg/L	241	1.4	≤4.5%	9, 13
QC 532 µg/L	539	1.1	≤4.5%	9, 13

Reproducibility				
	Mean	CV (%)	Acceptance criteria	References
QC 41 µg/L	37	15.2	<20%	9, 12
QC 114 µg/L	116	3.3	<15%	9, 12
QC 236 µg/L	249	3.5	<10%	9, 12
QC 532 µg/L	545	2.9	<10%	9, 12

Stability				
	Mean	CV (%)	Acceptance criteria	References
QC 41 µg/L	34	9.2	<20%	7, 9
QC 114 µg/L	116	2.1	<15%	7, 9
QC 236 µg/L	254	2.6	<10%	7, 9
QC 532 µg/L	565	2.5	<10%	7, 9

Lower limit of quantification (LLOQ)				
	Mean	CV (%)	Acceptance criteria	References
10 µg/L	9.8	14.9	<20%	12, 15
15 µg/L	15.3	15.1	<20%	12, 15

Uncertainty (%)				
	Mean	CV (%)	Acceptance criteria	References
<50 µg/L	44	38.5	<40%	12
50-200 µg/L	124	10.7	<30%	12
>200 µg/L	343	8.8	<20%	12

Abbreviations: HS, highly significant; NS, not significant; F, Fischer value; C, Cochran value.

strument response (y), and to assess the validity of the calibration model. Slope, intercept, and coefficient of determination of the regression model were calculated. Since a high coefficient of determination did not guarantee that the model fit the data, the existence of a slope and the goodness of the model were

tested. These tests were based on a Fisher's statistical test, which compared the variance of two different set of values. In order to evaluate each hypothesis, a research hypothesis (H1) and a null hypothesis (H0) were framed. The null hypothesis was rejected if the F-distribution calculated from the data was

ANNALS OF
LABORATORY
MEDICINE

Hejl CG, et al.
Quantification of lead in whole blood

greater than the critical value of the F-distribution for the desired false-rejection probability (0.05).

The hypothesis and Formula are shown in Table 1. Two sources of variation were identified and used: regression error (ss_{reg}) and residual error (ss_{resid}), which was composed of non-linearity error and repeatability error (Table 1). Variances in homogeneity of the five curves were evaluated by using Cochran's C Test, which was a one-sided upper limit variance outlier test (Table 1). As described previously, the null hypothesis, H0 was tested against an alternative hypothesis, H1 (Table 1). The upper limit critical value (C_{theo}) was chosen according to the desired significance level (α), the number of data series considered (N), and the number of data points (n) per data series. If C exceeded C_{theo}, H0 was rejected, i.e variances were not equal (Table 1). These statistical tests confirmed the existence of a slope, the goodness of the model, and the homogeneity of the five calibrations curves. The deviation of standards from nominal concentration was also calculated and it did not exceed 15%. Hence, the model was acceptable.

The lower limit of quantification (LLOQ) was fixed at 10 µg/L. According to published data, a LLOQ between 10 and 15 µg/L was considered normal [12-15].

It was established using five samples independent of standards. Precision was within 20% of the coefficient variation. Precision included within-batch run or repeatability and between–batch run or reproducibility.

Analyte recovery lied within 80 and 120%, which was in the acceptable range. To assess repeatability, 4 controls (QC1=41 µg/L, QC2=114 µg/L, QC3=236 µg/L, and QC4=532 µg/L) were measured 30 times during a single analytical run. Reproducibility was studied for 20 days with 30 determinations per concentration of the same controls with a maximum of two runs per day. Intra-assay and inter-assay imprecision were evaluated as requested by French regulation authorities, and were less than 10% [12]. It was also possible to compare the performances of the method to published data or proficiency test results [13-15]. On the contrary, biological imprecision criteria published by Vassault et al. [16] and Minchinela et al. [17] could not be used, as exogenous substances like lead were not included in their databases.

Contamination was assessed using the protocol described in SH GTA 04 by the French Committee for accreditation (CO-FRAC) [11]. A high-level quality control (QCH 532 µg/L) was analyzed in triplicate (QCH1, QCH2, and QCH3) and a low-level quality control (QCL 12 µg/L) was also analyzed in triplicate (QCL1, QCL2, and QCL3). Five replicates were performed and

the percentage of contamination was calculated using the formula (mean [QCH/QCL3-QCL1]×100). It was less than 1% and within the acceptable range.

The stability was evaluated on eight aliquots of each QC maintained at the routine storage conditions (4°C) and during the intended storage period (14 days). The results were compared to the mean of the back-calculated values for the standards at the appropriate concentrations from the first day of long-term stability testing [7-9]. The mean values of analyte concentration were within ±10% of the nominal value and the CV did not exceed the acceptance criteria [9].

Uncertainty of the method was estimated via the ISO guides to the expression of uncertainty in measurement (GUM) approach based on modeling the measurement and evaluating all sources of uncertainty [18, 19]. The uncertainty in patient result included both pre-analytical and analytical variations. Sources of pre-analytical uncertainty such as sample collection, handling, transport, and storage were difficult to estimate. When standardized procedures, such as guidelines for sample collection were followed, pre-analytical variations might be minimized [20]. Hence, to estimate uncertainty, we mainly considered the analytical variations. Two components were identified as standard sources of uncertainty: long-term precision (intra-laboratory reproducibility) and uncertainty of the calibration (estimated using bias). All identified sources were combined by "error-propagation" formulas to yield the standard uncertainty of the measurement results, which was multiplied by a factor of 2 to obtain the uncertainty. Estimates of measurement uncertainty showed results that were acceptable by COFRAC standards [12] (Table 2). Reference interval for acceptable levels in human body was not verified since the acceptable levels were defined by various national authorities.

To conclude, all the parameters tested satisfied our acceptance criteria. The validity of this method was established. This protocol seemed appropriate to guide laboratories in performing in-house methods validation like ETAAS. Validation allowed ensuring the appropriateness of test results. After initial validation, quality standards had to be maintained. ISO 15189 recommended the development of a quality control system to verify the intended quality of results. Medical laboratories have to develop continuous monitoring of assay performances [1, 6-9]. The laboratory should not get out-of-range results and our results were acceptable. It could also include, for instance, an internal control system, preventive maintenance of equipment, staff education and training, and internal audits [2, 3, 6].

Hejl CG, et al.
Quantification of lead in whole blood

ANNALS OF
LABORATORY
MEDICINE

Authors' Disclosures of Potential Conflicts of Interest

Acknowledgments

REFERENCES

1. Guzel O and Guner EI. ISO 15189 accreditation: Requirements for quality and competence of medical laboratories, experience of a laboratory I. Clin Biochem 2009;42:274-8.
2. Li H and Adeli K. Laboratory quality regulations and accreditation standards in Canada. Clin Biochem 2009;42:249-55.
3. Plebani M. Appropriateness in programs for continuous quality improvement in clinical laboratories. Clin Chim Acta 2003;333:131-9.
4. Laitinen P. Laboratory and quality regulations and accreditation standards in Finland. Clin Biochem 2009;42:312-3.
5. Shin BM, Chae SL, Min WK, Lee WG, Lim YA, Lee do H, et al. The implementation and effects of a clinical laboratory accreditation program in Korea from 1999 to 2006. Korean J Lab Med 2009;29:163-70.
6. Nichols JH. Verification of method performance for clinical laboratories. Adv Clin Chem 2009;47:121-37.
7. US Department of Health and Human Services. Guidance for industry: Bioanalytical method validation. 1st ed. Rockville: Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Veterinary Medicine (CVM), 2001:25.
8. National Association of Testing Authorities (NATA). Guidelines for the validation and verification of quantitative and qualitative test methods. Technical Note 17. 6th ed. Sydney: NATA, 2013:32.
9. COFRAC. SH GTA 04 Guide technique d'accréditation de vérification (portée A)/validation (portée B) des méthodes en biologie médicale. 1st ed. Paris: COFRAC, 2011:46.
10. AFNOR. XP-T 90-210 Qualité de l'eau- Protocole d'évaluation des performances d'une méthode alternative d'analyse physico-chimique quantitative par rapport à une méthode de référence. 1st ed. Paris: AFNOR, 1999:58.
11. Feinberg M. Labo-Stat: Guide de validation des méthodes d'analyse. 2nd ed. Paris: Lavoisier, 2009:361.
12. COFRAC. SH REF 20 Exigence specifique et recommandations d'accréditation en plombémie. 1st ed. Paris: COFRAC, 2010:36.
13. Labat L, Olichon D, Poupon J, Bost M, Hautfroid V, Moesch C, et al. Variability of blood lead determination for low levels near target values of 100 µg/L: a multicentric study. Ann Toxicol Anal 2006;18:297-304.
14. Izquiero Álvarez S, Calvo Ruata ML, González López JM, García de Jalón Comet A, Escanero Marcén JF. Validation of determination of lead (Pb) in blood by electrothermal atomic absorption spectrometry (ETAAS) on the basis of interlaboratory comparison data. J Trace Elem Med Biol 2007;21:S26-8.
15. Olichon D, Labat L, Poupon J, Bost M, Haufroid V, Moesch C, et al. Analytical approach of the limit of quantification of blood lead testing: a multicentric study. Ann Toxicol Anal 2007;19:31-6.
16. Vassault A, Grafmeyer D, de Graeve J, Cohen R, Beaudonnet A, Bienvenu J. Analyses de biologie médicale: spécifications et normes d'acceptabilité à l'usage de la validation de techniques. Ann Biol Clin 1999;57:685-95.
17. Minchinela J, Ricós C, Perich C, Fernández-Calle P, Alvarez V, Domenech M, et al. Biological variation database, and quality specifications for imprecision, bias and total error (desirable and minimum). The 2012 update. http://www.westgard.com/biodatabase-2012-update.htm (Updated on 2012).
18. Kristiansen J. The guide to expression of uncertainty in measurement approach for estimating uncertainty: an appraisal. Clin Chem 2003;49:1822-9.
19. COFRAC. SH GTA 14 Guide technique d'accréditation pour l'evaluation des incertitudes de mesure en biologie medicale. 1st ed. Paris: COFRAC, 2011.
20. Cornelis R, Heinzow B, Herber RF, Molin Christensen JM, Poulsen OM, Sabbioni E, et al. Sample collection guidelines for trace elements in blood and urine. IUPAC Commission of Toxicology. J Trace Elem Med Biol 1996;10:103-27.