Josiane Arnaud, Cas Weykamp*, Ross Wenzel, Marina Patriarca, Montserrat González-Estecha, Liesbeth Janssen, Ma'atem Beatrice Fofou-Caillierez, Montserrat Ventura Alemany, Valeria Patriarca, Irene de Graaf, Renaud Persoons, Mariona Panadès, Bernard China, Marieke te Winkel, Hans van der Vuurst and Marc Thelen

Analytical performance specifications for trace elements in biological fluids derived from six countries federated external quality assessment schemes over 10 years

https://doi.org/10.1515/cclm-2024-0551
Received May 2, 2024; accepted July 7, 2024;
published online July 22, 2024

Abstract

Objectives: This article defines analytical performance specifications (APS) for evaluating laboratory proficiency through an external quality assessment scheme.
Methods: Standard deviations for proficiency assessment were derived from Thompson's characteristic function applied to robust data calculated from participants' submissions in the Occupational and Environmental Laboratory Medicine (OELM) external quality assurance scheme for trace elements in serum, whole blood and urine. Characteristic function was based on two parameters: (1) β – the average coefficient of variation (CV) at high sample concentrations; (2) α – the average standard deviation (SD) at low sample concentrations. APSs were defined as 1.65 standard deviations calculated by Thompson's approach. Comparison between OELM robust data and characteristic function were used to validate the model.
Results: Application of the characteristic function allowed calculated APS for 18 elements across three matrices. Some limitations were noted, particularly for elements (1) with no sample concentrations near analytical technique limit of detection; (2) exhibiting high robust CV at high concentration; (3) exhibiting high analytical variability such as whole blood Tl and urine Pb; (4) with an unbalanced number of robust SD above and under the characteristic function such as whole blood Mn and serum Al and Zn.
Conclusions: The characteristic function was a useful means of deriving APS for trace elements in biological fluids where biological variation data or outcome studies were not available. However, OELM external quality assurance scheme data suggests that the characteristic functions are not appropriate for all elements.

Keywords: trace elements; biological fluids; external quality assessment; analytical performance specifications; standard deviation for proficiency assessment

The authors are members of the Occupational and Environmental Laboratory Medicine network (OELM) and/or participate in the OELM external quality assessment scheme.

***Corresponding author: Cas Weykamp**, MCA Laboratory, Queen Beatrix Hospital, Beatrixpark 1, Winterswijk 7101 BN, The Netherlands, E-mail: c.w.weykamp@skbwinterswijk.nl
Josiane Arnaud, Member of French Society for Clinical Biology (SFBC), and French Speaking Society for Trace Elements, Vitamins and Biofactors (SETViB), Paris, France
Ross Wenzel, Pathology NSW, Royal North Shore Hospital, St Leonards, NSW, Australia. https://orcid.org/0000-0002-4395-0239
Marina Patriarca and Valeria Patriarca, Department of Food Safety, Nutrition and Veterinary Public Health, Istituto Superiore di Sanità, Rome, Italy
Montserrat González-Estecha, Department of Laboratory Medicine, Hospital General Universitario Gregorio Marañón, Madrid, Spain
Liesbeth Janssen, Irene de Graaf, Marieke te Winkel and Hans van der Vuurst, MCA Laboratory, Queen Beatrix Hospital, Winterswijk, The Netherlands
Ma'atem Beatrice Fofou-Caillierez, Biochemistry, Molecular Biology and Nutrition Laboratory, CHRU Nancy, Nancy, France
Montserrat Ventura Alemany and Mariona Panadès, External Quality Assessment Schemes, Spanish Society of Laboratory Medicine, Barcelona, Spain
Renaud Persoons, University of Grenoble Alpes, CNRS, Grenoble INP, CHU Grenoble Alpes, TIMC, Grenoble, France
Bernard China, Department of Quality of Laboratories, Sciensano, Brussels, Belgium
Marc Thelen, Foundation of Quality Assurance in Laboratory Medicine (SKML), Nijmegen, The Netherlands; and Department of Laboratory Medicine, Radboud University Medical Center, Nijmegen, The Netherlands. https://orcid.org/0000-0003-1771-669X

Introduction

Clinical laboratories play a significant role in the diagnosis, treatment and follow-up of diseases and must provide the highest possible quality of results. To ensure reliable and relevant results, laboratories use internal and external quality controls. Internal quality controls allow to validate patient's results daily and to follow the imprecision. In contrast, external quality assessment schemes (EQAS) allow to verify accuracy retrospectively and occasionally. This latest verification is based on standard deviations (σ_{EQAS}) and analytical performance specifications (APS) defined by the organizer. Organizers of EQAS provide a report to the participants that compares the submitted results with an assigned value and indicate whether the deviation from the assigned value achieved by the laboratory is acceptable. APS define upper and lower acceptable limits [1] and σ_{EQAS} is used to grade participant's results.

Participant's scoring

As indicated above, laboratory performance can be graded by evaluation against acceptability criteria set by EQAS organizers. The most common method for scoring quantitative results in an EQAS is through the use of a z-score. They compare the difference between a participant's result (x_i) and a value assigned to the sample (X) with a standard deviation (SD) defined by the EQAS organizers (σ_{EQAS}), according to Eq. (1):

$$Z - \text{score} = (x_i - X)/\sigma_{EQAS} \quad (1)$$

Determination of standard deviation for EQAS

ISO 13528 [2] presents various approaches to determine the standard deviation (SD) for proficiency assessment (σ_{EQAS}). These include prescribed values (i.e. from legislation); performance levels that experts in the field wish to achieve; SD derived from general models of interlaboratory variability.

The European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) [3] have recommended setting analytical performance specifications (APS) as a percentage of sample concentration based on either clinical, biological or analytical information.

Ideally, performance requirements should reflect clinical needs derived from outcome studies. This approach is however difficult to apply [4, 5]. It requires cut-off concentrations that generate a modification in patient follow-up or treatment with limited risk of error [5–8]. Currently, insufficient information is available to permit a clinical outcome approach for most trace elements. Thresholds have been proposed for Cd related renal damage [9] and for Pb neuro- and repro-toxicity [10]. The American Heart Association has recently reviewed the link between As, Cd and Pb exposures and cardiovascular risk [11].

APS based on components of biological variation, namely, within- and between-subject variation, have been considered the most appropriate approach [4, 12, 13]. However, since reliable and relevant biological variation data are available for only a few essential trace elements [14–16], the application of this model is limited. In addition, results from different studies may be discrepant. These differences can often be attributed to study design, analytical technique used and statistical treatment of results [17, 18]. Considering these difficulties, the EFLM have proposed a standard protocol for assessing biological variation and the criteria applied for their critical appraisal [14, 18–21]. Studies following these protocols have been published using steady state conditions in healthy populations [15, 20]. These studies confirmed previous estimates for serum Cu, Se and Zn [15, 22] though high individual variability was observed for Se and Cu [15]. Data on biological variation are regularly updated on the EFLM website [14].

The state-of-the-art approach to defining APS considers the relationship between interlaboratory variability (as coefficient of variation (CV) or SD) and mass fraction [4, 23, 24]. These models are useful when no robust biological variation data are available [25]. A deficiency of this approach occurs when the CV expands asymptotically at concentrations approaching the limit of detection. To alleviate this, Thompson [23] proposed a characteristic function with empirically determined variables which model the CV both at low and high concentrations. This function was applied to Cd, Hg, Pb and Mn in whole blood and urine using data from interlaboratory comparisons [26, 27].

Objective of the article

The objective of this paper was to investigate the suitability of Thompson's characteristic function [23] to set APS for the six countries federated Occupational and Environmental Laboratory Medicine (OELM) EQAS [28]. Data used were that had been collated from this scheme in the years 2011–2021. Criteria for the validation of Thompson's function were proposed. Performances of the federative scheme were compared to biological variation data where available and to outcome approach for some toxic elements.

Materials and methods

OELM Federate EQAS design

To overcome difficulties linked to limited numbers of participants and differences in performance evaluation, the federated OELM EQAS was established in April 2011, providing determinations of eight trace elements in serum. Subsequent expansion of the program added whole blood, urine and additional serum trace elements. Currently, the OELM EQAS offers 19 elements in urine, 11 in whole blood, and 11 in serum [28].

All samples were prepared at the MCA Laboratory (Queen Beatrix Hospital, Winterswijk, Netherlands) according to standard procedures [29]. Human blood and urine were collected from donors under informed consent. Serum samples were prepared from bovine serum. For each matrix, two primary pools were supplemented with a complementary set of inorganic salts of selected elements. Then, these two primary pools were mixed in different proportions to obtain 12 individual secondary pools of various trace element concentrations. Each secondary pool was halved into two aliquots that were assigned different codes to form 24 samples that were shipped to national organizers for distribution to their participants.

Results of sample pairs and analytical technique were collected each month by electronic submission. Robust means, SD and CV were calculated according to algorithm A [2] for all participants and grouped by analytical technique i.e. electrothermal atomic absorption spectrometry (EAAS), flame atomic absorption spectrometry (FAAS), inductively coupled plasma – mass spectrometry (ICP-MS), inductively coupled plasma – optical emission spectroscopy (ICP-OES) or colorimetry.

Laboratory performances were graded using APS initially agreed by the organizers using previously available information.

Thompson's characteristic function parameters

The characteristic function of Thompson [23] was applied to evaluate fitted CV and SD data ($CV_{R(fit)}$ and $s_{R(fit)}$) based on laboratory capability for each element and sample matrix. The calculation of parameters required several steps. First, "Algorithm A" [2] was applied to the participants' results submitted for each sample in order to determine the robust mean (C) and robust standard deviation (s_R). Second, three parameters (CV_R, β, and α) were calculated based on C and s_R. Coefficients of variation (CV_R) were calculated by dividing s_R by C. The parameter β represents the CV_R at high concentrations and was determined by classic mean CV_R of samples with relatively high levels. The parameter α describes s_R at low concentrations characterized by asymptotically increase in CV_R where available. The individual α values were calculated using Eq. (2):

$$a = \sqrt{s_R^2 - \beta^2 C^2} \tag{2}$$

The overall α value was determined by calculating classic mean of α values at lower concentrations. After the empirical determination of all parameters, the fit for s_R vs. C was determined using Eq. (3):

$$s_{R\,(fit)} = \sqrt{\alpha^2 + \beta^2 C^2} \tag{3}$$

The fit for CV_R vs. C was determined by dividing $s_{R(fit)}$ by C to obtain $CV_{R(fit)}$.

Analytical performance specifications

APS define the upper and lower acceptance limits around the assigned value [1]. An acceptable participant's result must fall within the range defined by the value assigned plus or minus APS. APS was calculated with a 90 % probability according to Eq. (4) [22, 30–32]:

$$APS = 1.65\, s_{R\,(fit)} \tag{4}$$

Validation of Thompson's characteristic function

Characteristic functions are suitable when s_R of the investigated samples are close to s_{Rfit} and equally distributed around the curve. In contrast, characteristic functions may be inappropriate because of systematic (bias) and random (imprecision) variations of s_R.

In cases of bias (systematic error), there is an unbalance between s_Rs under and above the characteristic functions. If too many s_Rs are under the curve, derived APS are too lenient: too many laboratories receive satisfactory z-scores. If too few s_Rs are under the curve, derived APS are too stringent: too many laboratories receive unsatisfactory z-scores. The percentage of s_R below the characteristic functions were used as an index of bias. Validation criteria were chosen empirically: a percentage between 40.0 and 60.0 % was considered satisfactory, between 30.0 and 39.9 or 60.1 and 70.0 questionable, and lower than 29.9 or higher than 70.1 unsatisfactory.

Table 1: Number of participants and concentration ranges of trace elements in the OELM EQAS from cycle 2011/12 to cycle 2021/22.

Element	Unit	Number of samples			Number ofparticipants (range)			Sample concentration (range)		
		Blood	Serum	Urine	Blood	Serum	Urine	Blood	Serum	Urine
Al	µg/L		240			65–131			5.08–280	
As	µg/L	216		216	12–40		20–50	0.51–104		0.92–180
Cd	µg/L	216		216	39–61		29–59	0.29–8.43		0.06–10.9
Co	µg/L	216	240	216	23–36	31–56	29–52	0.21–56.5	0.17–50.4	0.08–56.9
Cr	µg/L	192	240	216	16–37	35–62	38–59	0.34–55.8	0.19–59.4	0.22–49.6
Cu	µg/L		240	216		110–207	63–99		403–3,110	19.1–909
Fe	µg/L			192			13–31			0.01–4.97
Hg	µg/L	216			23–49			0.38–148		
I	µg/L			193			23–41			23.2–827
Li	mg/L		230			31–44			0.10–21.4	
Mg	mg/L	192	240	216	15–38	23–33	21–27	16.4–71.0	10.3–48.2	22.4–158
Mn	µg/L	216		216	18–50		23–49	4.91–42.2		0.27–31.1
Ni	µg/L			216			28–55			0.49–59.2
Pb	µg/L	216		216	73–105		44–69	9.09–777		0.19–142
Se	µg/L	216	240	216	9–33	62–117	10–38	49.5–390	16.3–293	4.04–255
Tl	µg/L	216		216	11–28		15–37	0.04–9.64		0.05–19.3
V	µg/L			216			8–36			0.17–20.9
Zn	µg/L	216	240	192	20–32	111–214	48–68	3,040–34,800	445–3,033	130–3,300

In instances of imprecision (random error), there are too many s_Rs far above or far under the characteristic functions. For a sample with an s_R far above the curve there will be too many laboratories with a poor z-score and vice versa. The percentage of absolute difference between s_R and s_{Rfit} lower than 0.5 s_{Rfit} was used as an index of imprecision. Validation criteria were chosen empirically: a percentage higher than 90.0 % was considered satisfactory, between 80.0 and 89.9 % questionable and lower than 79.9 % unsatisfactory.

Influence of analytical techniques on performances

Comparison of performance from different groups of analytical techniques (ICP-MS, EAAS, FAAS and colorimetry) were evaluated using a Mann–Whitney U test when participant numbers in the various groups exceeded 15 (Supplementary Material, Table S1).

Results

Data used for analysis

In this study, data for all participants from 2011 to 2021 were used for those elements and matrices with a sufficient number of participants (average ≥15) that had been submitted over at least eight cycles. The number of participants and trace element concentrations varied according to matrix and element as shown in Table 1.

Over ten years, participant laboratories and analytical techniques used varied. ICP-MS users increased while EAAS users decreased. Participant numbers varied after leaving the United Kingdom National External Quality Assessment Service in 2012 and with the inclusion of federate partners, the Spanish Society of Laboratory Medicine, and the Royal College of Pathologists of Australasia Quality Assurance Program that joined the OELM in 2012 and 2014, respectively.

The effect of analytical technique on robust means and robust CV could only be assessed on whole blood Cd and Pb, serum Al, Cr, Cu, Se and Zn, and urine Cr, Cu, Pb and Zn. CV_R were significantly lower using ICP-MS compared to AAS or colorimetry (Mann–Whitney U test, Supplementary Material, Table S1); excepted for Cr in serum due to the small number of samples (n=6) and cycles involved. In contrast, robust means were not affected by analytical technique.

Thompson's characteristic function parameters

The β parameter was determined by taking the classic mean CV_R for C values higher than the concentrations shown in Table 2. The final β values were multiplied by 100 and the

Table 2: Thompson's characteristic function parameters for trace elements in whole blood, serum and urine samples including the concentration (C) used for the estimate of the fitted parameter β that represents the CV_R at high concentrations and parameter α that describes s_R at concentrations close to the LOD where available and the ratio α/β.

Element	Whole blood					Serum					Urine				
	β estimate		α estimate			β estimate		α estimate			β estimate		α estimate		
	C[a], µg/L	β[b], %	C[c], µg/L	α, µg/L	α/β, µg/L	C[a], µg/L	β[b], %	C[c], µg/L	α, µg/L	α/β, µg/L	C[a], µg/L	β[b], %	C[c], µg/L	α, µg/L	α/β, µg/L
Al						100	14.6	10	3.93	26.9					
As	30.0	11.0	1.50	0.81	7.44						90	8.53	1.00	0.58	6.81
Cd	7.00	9.10	0.90	0.20	2.15						4.5	9.22	0.20	0.09	0.96
Co	10.5	8.60	0.50	0.22	2.50	10.0	7.70	0.50	0.18	2.29	10.0	9.41	0.50	0.15	1.57
Cr	12.0	9.32	0.90	0.49	5.25	8.0	9.78	0.80	0.36	3.67	10.5	10.5	0.50	0.32	3.08
Cu						2,000	7.32	500	31.8	435	200	10.6	75.0	7.46	70.5
Fe											1,120	11.7	30.0	17.5	149
Hg	20.0	12.2	0.95	0.35	2.84										
I											550	8.70	50.0	5.29	60.8
Li						10,000	6.57	1,600	164	2,497					
Mg	16,000	7.46				20,000	6.20	12,000	466	7,516	2,800	9.10			
Mn	25.0	10.7	10.0	1.52	14.2						6.00	10.5	0.90	0.37	3.55
Ni											15.0	9.96	1.70	0.85	8.58
Pb	200	8.34	30.0	3.51	42.1						41.0	11.8	2.00	0.74	6.29
Se	300	9.15	97.0	8.30	90.7	200	9.24	25	4.87	52.7	180	9.93	6.00	1.82	18.3
Tl	4.00	8.00	0.30	0.11	1.32						5.00	9.30	0.14	0.08	0.85
V											10.0	11.1	0.25	0.24	2.14
Zn	3,100	8.12	3,100	252	3,107	1,500	7.81	600	51.1	654	2,050	8.74	150	21.5	246

[a]Lowest sample concentration in µg/L empirically used for β determination. β, mean of CV_R for sample concentrations higher than C. [b]β was multiplied by 100. [c]Highest sample concentration in µg/L empirically used for α determination. α, mean of s_R for sample concentration lower than C.

classic mean α was determined for C values lower than the concentrations shown in Table 2. These cut-off concentrations were empirically selected according to the distribution of data, the percentage of CV_R above and below Thompson's characteristic function, the lowest possible β value. Characteristic functions for serum, whole blood and urine samples are presented in Figures 1 to 3.

Low sample concentrations were difficult to obtain as they necessitate blood from deficient animals or humans. This was particularly the case for Cu, Mg, Se, Zn in serum, Mg, Mn, Se, Zn in whole blood, and I, Zn, Mg, Se in urine. This lack of low sample concentrations did not allow to calculate properly α and was a limiting factor for Thompson's approach. In addition, increased CV_R at the highest concentrations were noticed for serum Al, Co, Cr and whole blood Co and Cr. This was a limiting factor for the calculation of β.

Thompson's characteristic function suitability

Thompson's characteristic function was considered to overestimate results when at least 60 % of CV_R or s_R values were lower than $CV_{R(fit)}$ and $s_{R(fit)}$. Overestimation was noted for serum Al, Se, Zn; whole blood Mn, Zn; urine I, Mn and Ni (Figure 4). Overestimation was particularly important for serum Al and Zn in serum and whole blood Mn. For serum Al, there was no CV_R decrease at the highest sample concentration, which remained a limiting factor for the calculation of β. For serum Zn and whole blood Mn, calculation of α was difficult due to the lack of sample with low concentrations.

Thompson's characteristic function was considered inappropriate for defining APS when less than 80 % of absolute difference between s_R and s_{Rfit} were lower than 0.5 s_{Rfit}. This was the case for whole blood Tl and urine Pb due to dramatically high variability of robust data at concentrations around α/β and consequently remained a limiting factor for the calculation of β and α. Interestingly, percentages between 80.0 and 89.9 % were only observed in urine (Figure 4). For these elements, a careful follow-up seems useful.

Impact of APS change on laboratory evaluation

Compared to our previous APS, those already based on biological variation approach were unchanged (serum Cu, Mg, Se, and Zn, whole blood Pb and Se). APS based on analytical

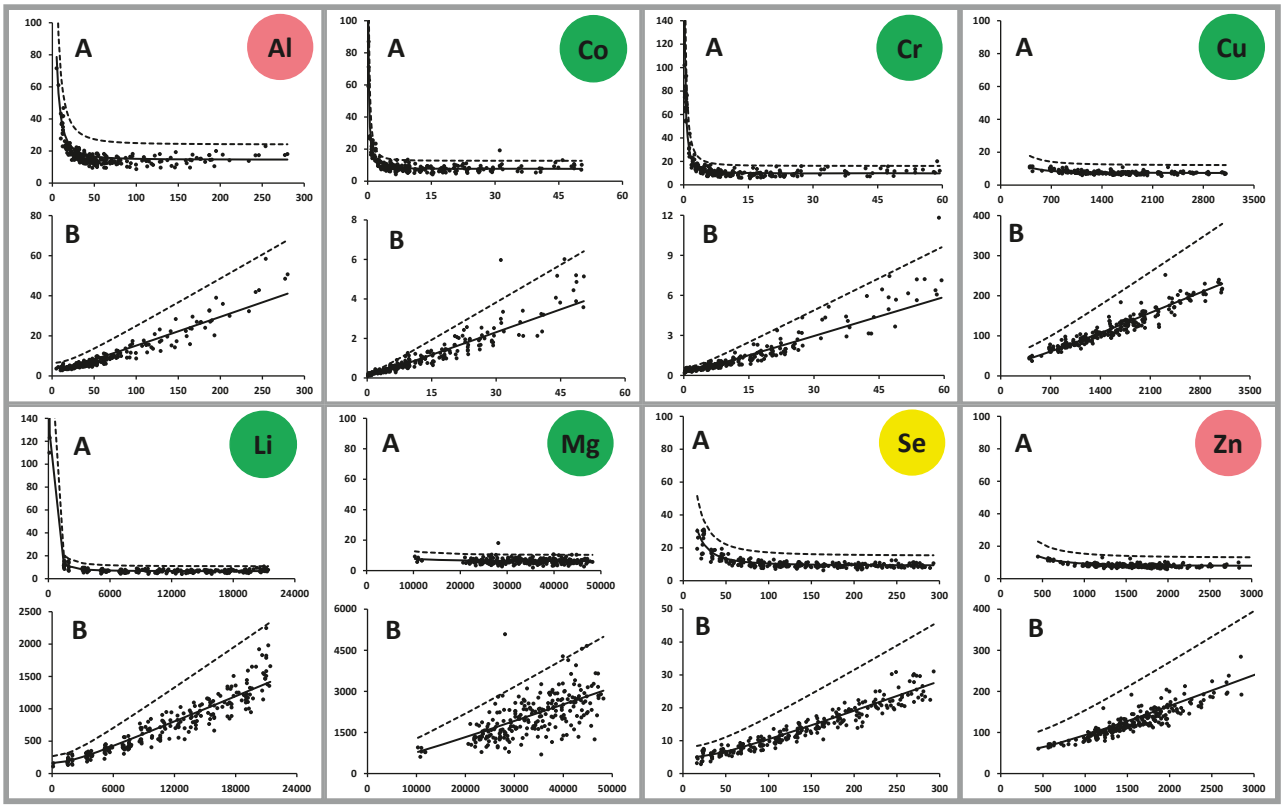

Figure 1: Plots of Thompson's characteristic function (black line) and APS (dashed black line) applied to OELM Federate EQAS participant data for serum trace elements (black circles). Y scale: (A) robust coefficient of variation (CV_R) in percent; (B) robust standard deviation (s_R) in µg/L. X scale: sample concentration in µg/L. Coloured circle key: green – acceptable bias and imprecision; yellow – questionable bias or imprecision; red – unsatisfactory imprecision; pale red – unsatisfactory bias.

performances remained similar for serum Al and Li, whole blood As, Mg, Mn and Zn as well as urine Cd, Co, Fe, Mg, Ni, Pb and Zn. In contrast, APS for serum Co, Cr, whole blood Cd, Co, Cr, Hg and urine As, Cr, Cu, I, Mn, Se, Tl and V were more stringent. Consequently, a higher number of results may be qualified as questionable or unsatisfactory.

Discussion

Laboratory results must be reliable for clinical decision-making and patient management. Therefore, APS must be objectively defined [25]. Limited data is available regarding APS for trace elements in biological fluids, whichever model has been applied.

Difficulties encountered in estimating the Thompson's characteristic function

We have reported the estimation of 34 characteristic functions for 18 elements in three biological fluids using a state-of-the-art model. Some difficulties were noted. Parameters such as analytical technique, calibration procedure, interferences and/or erroneous values may explain the high variability of CV_R observed particularly for whole blood Tl and urine Pb. Indeed, CV_Rs were significantly lower using ICP-MS (Supplementary Material, Table S1), consistent with previous papers [15, 16]. In contrast, robust means were not affected by analytical technique (FAAS, EAAS, colorimetry, ICP-AES or ICP-MS), suggesting comparability of analytical results and lack of specific biases associated with technique. However, in some samples, differences were noted between analytical techniques due to specific interferences (not shown). In addition, some participating laboratories did not submit results for all samples which may also contribute to CV_R variability. Some EQAS rely on results from reference laboratories to set σ_{EQAS} [33] and reduce CV_R [34]. Sample matrix was another parameter that influenced our results. One possible explanation for poorer performance in whole blood may be related to sample properties such as viscosity and adhesiveness of blood making it more difficult to reliably pipet. Urinary matrix is particularly prone to interferences due to its chemical composition both for EAAS and ICP-MS.

In addition to experimental CV_R variability, another difficulty was observed for essential elements. For serum Cu, Mg, Zn, whole blood Mg, Se, Zn, and urine Mg, Zn, no sample with low concentrations were available as noted by CV_R values at the lowest concentrations under 20 % which represented a limiting factor for the calculation of α.

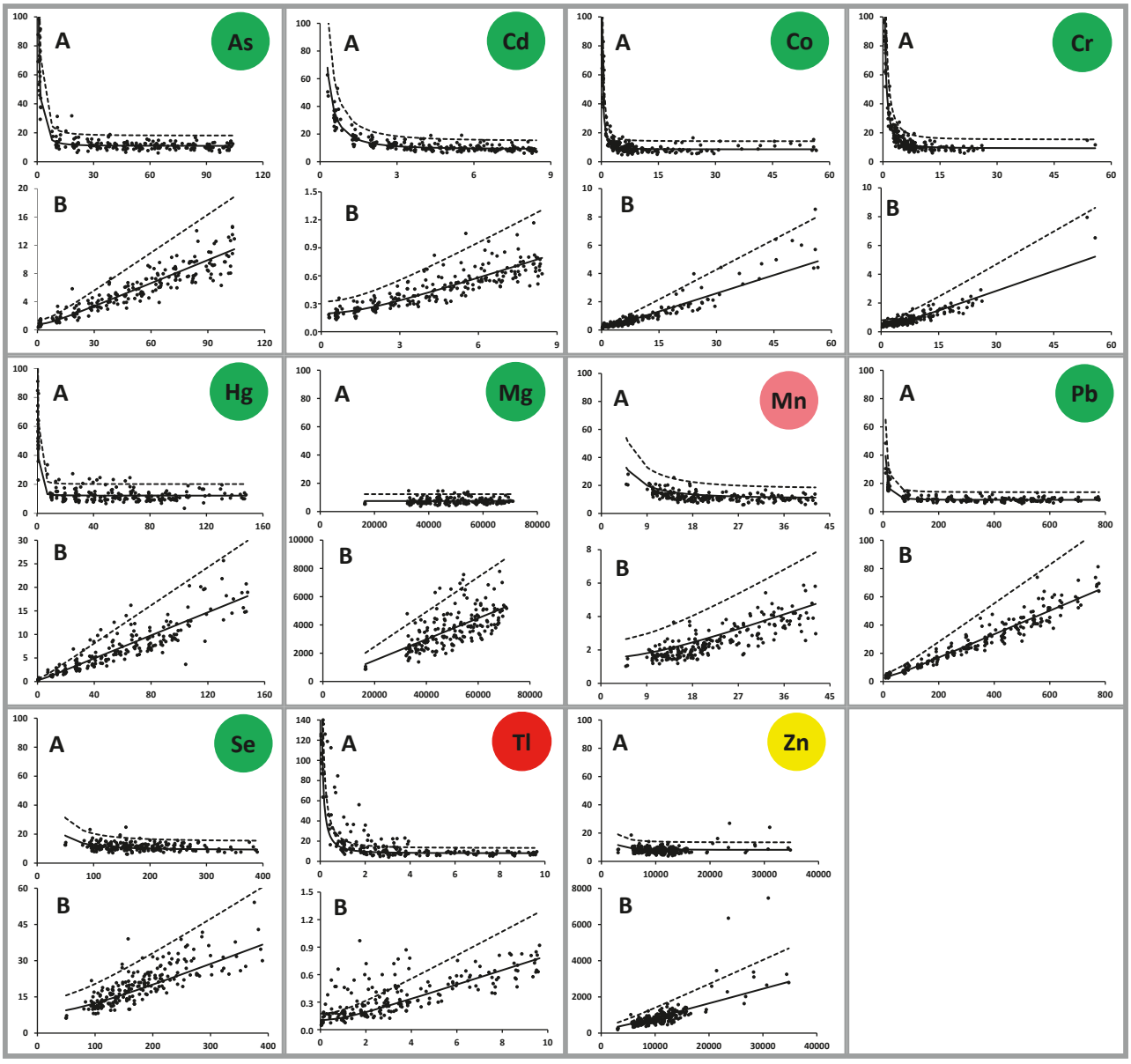

Figure 2: Plots of Thompson's characteristic function (black line) and APS (dashed black line) applied to OELM Federate EQAS participant data for whole blood trace elements (black circles). Y scale: (A) robust coefficient of variation (CV_R) in percent; (B) robust standard deviation (s_R) in µg/L. X scale: sample concentration in µg/L. Coloured circle key: green – acceptable bias and imprecision; yellow – questionable bias or imprecision; red – unsatisfactory imprecision; pale red – unsatisfactory bias.

Finally, the characteristic function overestimated APS for serum Al, Zn, and whole blood Mn suggesting that these APS can be too lenient.

Comparison of our characteristic functions to previous reports

Two reports have been published to establish Thompson's characteristic functions for trace elements in biological fluids. Cote et al. [26] reported results for whole blood and urine Cd, Hg, Pb (Supplementary Material, Table S2). They compared three different EQAS and reported differences in β and α/β according to EQAS. Our β were higher than those reported by Cote et al. [26] whereas α/β were similar for whole blood Cd, Pb, and lower for whole blood Hg and urine Cd, Pb. For Mn in urine and blood, Praamsma et al. [27] calculated β and α/β using robust data from four EQAS. In urine, β and α/β values were comparable to our results. In whole blood, β was lower than our result whereas α/β

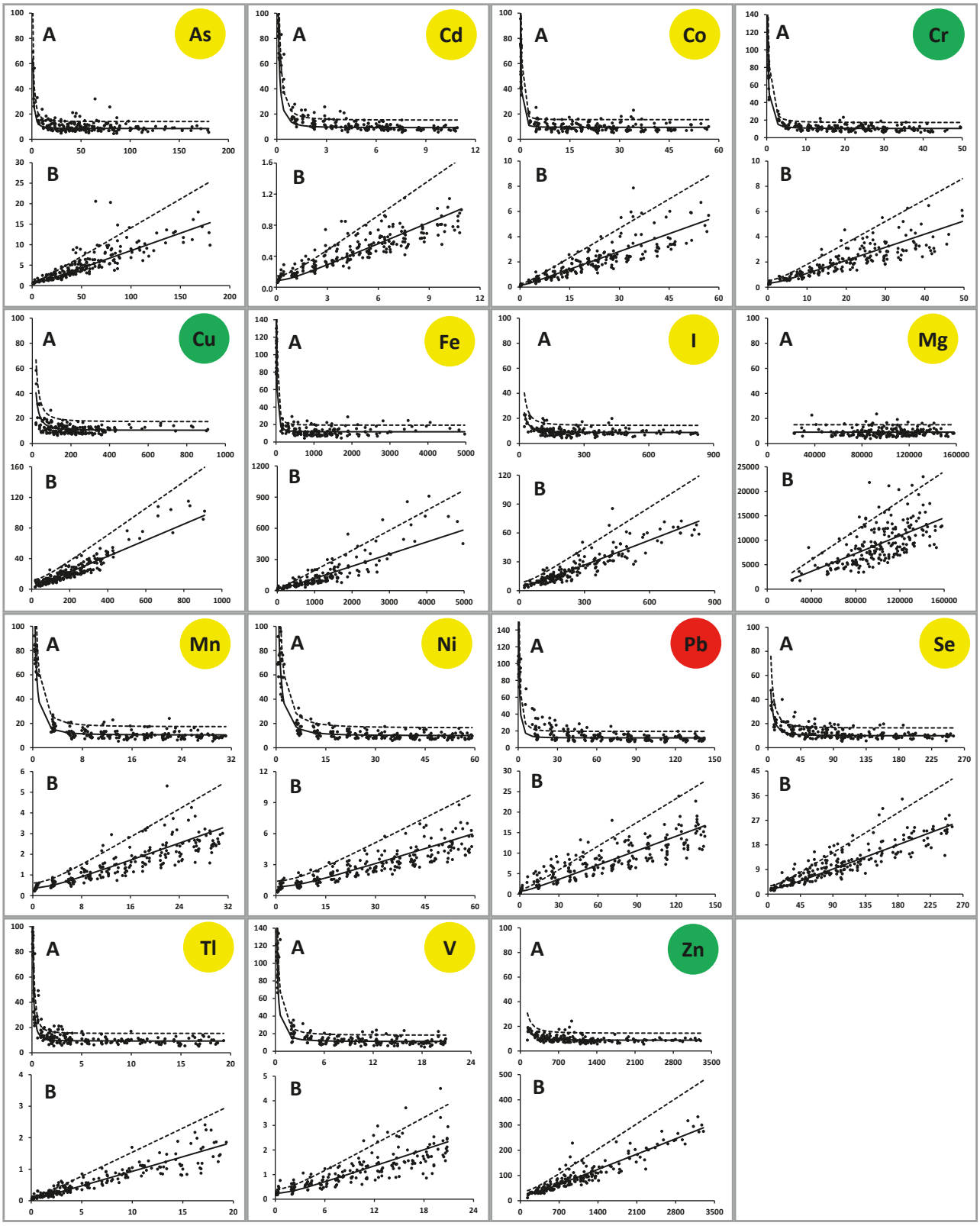

Figure 3: Plots of Thompson's characteristic function (black line) and APS (dashed black line) applied to OELM Federate EQAS participant data for urine trace elements (black circles). Y scale: (A) robust coefficient of variation (CV_R) in percent; (B) robust standard deviation (s_R) in µg/L. X scale: sample concentration in µg/L. Coloured circle key: green – acceptable bias and imprecision; yellow – questionable bias or imprecision; red – unsatisfactory imprecision; pale red – unsatisfactory bias.

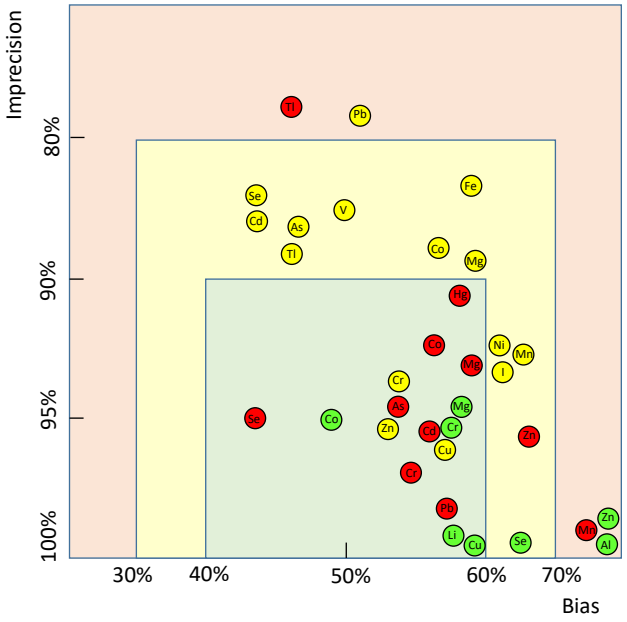

Figure 4: Estimation of Thompson's characteristic function suitability for setting analytical performance specifications (APS) based on two indexes: percentage of s_R lower than s_{Rfit} and percentage of absolute difference between s_R and s_{Rfit} lower than 0.5 s_{Rfit}. Y scale – imprecision evaluated by the percentage of $Is_R - s_{Rfit}| < 0.5 s_{Rfit}$. Imprecision was considered acceptable (green background) when between 90.0 and 100 %, questionable (yellow background) when between 80.0 and 89.9 %, unsatisfactory (red background) when <80.0 %. X scale – bias evaluated by the percentage of $s_R < s_{Rfit}$. Bias was considered acceptable (green background) when between 40.0 and 60.0 %, questionable (yellow background) when between 30.0 and 39.9 or 60.1–70.0 %, unsatisfactory (red background) when <30.0 or >70.0 %. Red dots – whole blood trace elements; green dots – serum trace elements; yellow dots – urine trace elements.

was similar. In agreement with previous reports [26, 27], characteristic function did not fit experimental data for all elements studied.

Comparison of APS derived from characteristic function and from biological variation models

Compared to a biological variation model, our APS determined by a state-of-the-art method model could only achieve the minimum requirement for serum Cu (15.0 %), and Zn (15.1 %) whereas the minimum requirement for serum Mg (6.02 %) was far from attainable and our serum Se APS was slightly higher than the minimum requirement (14.8 %) [14]. This finding could be a consequence of variability in analytical techniques used by our participants. In whole blood, Mn and Se APS defined by a state-of-the-art method model were higher than the desirable biological variation (14.4 % and 14.1 %, respectively) but lower than the minimum requirement (21.7 % and 21.2 %, respectively) when calculated using non-evaluated EFLM studies [27, 35, 36]. For whole blood Pb, our APS was one point higher than the minimum requirement evaluated by Taylor et al. (12.8 %) [30, 37]. These results suggest that further efforts must be made to improve our education to OELM EQAS participants. Of note, biological variation APS can currently only be applied to healthy populations. In addition, biological variation approach remains difficult to apply in urine because intra- and inter-individual variability in urinary excretion are higher than analytical performance precision [7], particularly in spot urine samples that are not suitable to accurately reflect circadian variations. For trace elements, it would be worthwhile evaluating populations with steady-state low or high trace element concentrations. However, even in the general population, there is an urgent need to evaluate trace element biological variations, particularly for toxic elements. Such data would help establish APS based on this model.

Comparison of APS derived from characteristic function and outcome needs

There is insufficient published information to allow for a clinical outcome approach to trace elements. This approach could be useful for toxic elements of public health concern identified by the World Health Organization (WHO), namely As, Cd, Hg, Pb [38]. Currently, the WHO has only proposed a threshold for whole blood Pb. For those individuals with a whole blood Pb concentration ≥50 µg/L, the WHO strongly recommends identification of the source(s) of exposure with appropriate action taken to terminate exposure [10, 39]. When compared to 35 µg/L as the 97.5th percentile of whole blood Pb for adults and children in the United States [40], our APS is slightly higher than the 17.6 % required to separate these two thresholds. We can assume that ICP-MS users are able to separate these two physio-pathological conditions. For whole blood Cd, an expert panel has proposed a critical value of 5 µg/L based on nephrotoxicity [9]. In the general adult population, the whole blood Cd concentration is less than 0.7 µg/L in non-smokers, increasing up to 3 µg/L in smokers [41]. Our APS is able to separate smokers from an exposed population. In contrast, for urine Cd, our APS do not allow separation from the 95th percentile of the general population (0.91 µg/L or 0.78 µg/g creatinine, [40]), and a threshold of 1 µg/g creatinine proposed for the prevention of tubular nephropathy in the general population [9]. However, based on the threshold of 2 µg/g creatinine proposed by the Scientific Committee on Occupational Exposure Limits [42]

or 3 µg/g creatinine proposed by the Occupational and Safety Health Administration [11], our APS separates the general and at-risk populations. Regarding As in urine, our APS do not allow separation at the 95th percentile for inorganic and methylated As in urine in the general population (27.8 µg/L or 30.1 µg/g creatinine, [40]) with the threshold of 35 µg/g creatinine based on biological exposure index [11]. Moreover, the American Conference of Governmental Industrial Hygienist has decreased this threshold in 2023 to 15 µg/g creatinine, a value within the range of the United States general population between 2015 and 2016 [40, 43]. Regarding whole blood Hg, the Environmental Protection Agency has calculated the lower limit of the benchmark dose for pregnant and nursing women, women considering pregnancy and children of 5.8 µg/L for blood methyl-Hg which corresponds to 6.4 µg/L of total mercury [44]. However, the Centers for Disease Control and Prevention defines high level exposure to Hg as a blood level greater than 10 µg/L [45]. When compared to 4.25 µg/L as the 97.5th percentile of whole blood Hg for adults and children in the United States [40], our APS only separates the general population from the threshold of 10 µg/L. These observations suggest that further efforts must be made to improve our education to OELM EQAS participants.

Conclusions

The application of Thompson's characteristic function to past OELM EQAS data indicate that this model could be applied to derive APS for trace elements in biological fluids where biological variation has not been defined. However, limitations were observed in this study with fitness for purpose mainly dependent upon concentration range and inter-laboratory performances. Indeed, this approach required sample with low concentrations, a steady decrease in coefficient of variation as concentration increased, and low variability in the coefficient of variation at similar concentrations. Although further efforts are needed to improve performance and sample commutability, these APS estimations can be implemented in our EQAS for whole blood As, Cd, Co, Cr, Hg, Mg, serum Co, Cr, Li and urine Cr, Cu, Zn but are not fit for purpose for whole blood Mn, Tl, serum Al and urine Pb. For other elements, we recommend a regular follow-up of participant performances.

Acknowledgments: The authors would like to thank participating laboratories who submitted their results.
Research ethics: Not applicable.
Informed consent: Not applicable.
Author contributions: The authors have accepted responsibility for the entire content of this manuscript and approved its submission.
Competing interests: The authors state no conflict of interest.
Research funding: None declared.
Data availability: Not applicable.

References

1. Jones GRD, Albarede S, Kesseler D, MacKenzie F, Mammen J, Pedersen M, et al. Analytical performance specifications for external quality assessment – definitions and descriptions. Clin Chem Lab Med 2017;55:949–55.
2. ISO 13528:2022. Statistical methods for use in proficiency testing by interlaboratory comparisons. https://www.iso.org/standard/78879.html [Accessed 6 Mar 2024].
3. European Federation of Clinical Chemistry and Laboratory Medicine (EFLM). https://www.eflm.eu [Accessed 17 Feb 2024].
4. Sandberg S, Fraser CG, Horvath AR, Jansen R, Jones G, Oosterhuis W, et al. Defining analytical performance specifications: consensus statement from the 1st Strategic Conference of the European Federation of Clinical Chemistry and Laboratory Medicine. Clin Chem Lab Med 2015;53:833–5.
5. Horvath AR, Bossuyt PM, Sandberg S, John AS, Monaghan PJ, Verhagen-Kamerbeek WD, et al. Setting analytical performance specifications based on outcome studies – is it possible? Clin Chem Lab Med 2015;53:841–8.
6. Miller WG, Jones GR, Horowitz GL, Weykamp C. Proficiency testing/external quality assessment: current challenges and future directions. Clin Chem 2011;12:1670–80.
7. Panteghini M, Ceriotti F, Jones G, Oosterhuis W, Plebani M, Sandberg S. Task Force on performance specifications in laboratory medicine of the European Federation of Clinical Chemistry and Laboratory Medicine (EFLM). Strategies to define performance specifications in laboratory medicine: 3 years on from the Milan Strategic Conference. Clin Chem Lab Med 2017;55:1849–56.
8. Ceriotti F, Cobbaert C. Harmonization of external quality assessment schemes and their role – clinical chemistry and beyond. Clin Chem Lab Med 2018;56:1587–90.
9. Lamkarkach F, Ougier E, Garnier R, Viau C, Kolossa-Gehring M, Lange R, et al. Human biomonitoring initiative (HBM4EU): human biomonitoring guidance values (HBM-GVs) derived for cadmium and its compounds. Environ Int 2021;147:106337.
10. Lead poisoning. World Health Organization; 2023. https://www.who.int/news-room/fact-sheets/detail/lead-poisoning-and-health [Accessed 6 Mar 2024].
11. Lamas GA, Bhatnagar A, Jones MR, Mann KK, Nasir K, Tellez-Plaza M, et al. Contaminant metals as cardiovascular risk factors: a scientific statement from the American Heart Association. J Am Heart Assoc 2023;12:e029852.
12. Keles M. Evaluation of the clinical chemistry test analytical performance by using different models and specifications. Turk J Biochem 2020;45:11–8.
13. Ricos C, Fernandez-Calle P, Perich C, Sandberg S. External quality control in laboratory medicine. Progresses and future. Adv Lab Med 2022;3:221–31.

14. Aarsand AK, Fernandez-Calle P, Webster C, Coskun A, Gonzales-Lao E, Diaz-Garzon J, et al. The EFLM biological variation database. https://biologicalvariation.eu/ [Accessed 17 Feb 2024].
15. Coşkun A, Carobene A, Aarsand AK, Aksungar FB, Serteser M, Sandberg S, et al. European Federation of Clinical Chemistry and Laboratory Medicine Working Group on Biological Variation and Task Group for the biological variation database. Within- and between-subject biological variation data for serum zinc, copper and selenium obtained from 68 apparently healthy Turkish subjects. Clin Chem Lab Med 2022;60:533–42.
16. Coşkun A, Aarsand AK, Braga F, Carobene A, Díaz-Garzón J, Fernandez-Calle P, et al. Systematic review and meta-analysis of within-subject and between-subject biological variation estimates of serum zinc, copper and selenium. Clin Chem Lab Med 2022;60:479–82.
17. Plebani M, Padoan A, Lippi G. Biological variation: back to basics. Clin Chem Lab Med 2015;53:155–6.
18. Bartlett WA, Braga F, Carobene A, Coşkun A, Prusa R, Fernandez-Calle P, et al. Biological variation Working Group, European Federation of Clinical Chemistry and Laboratory Medicine (EFLM). A checklist for critical appraisal of studies of biological variation. Clin Chem Lab Med 2015;53:879–85.
19. Aarsand AK, Røraas T, Fernandez-Calle P, Ricos C, Díaz-Garzón J, Jonker N, et al. European Federation of Clinical Chemistry and Laboratory Medicine Working Group on Biological Variation and Task and Finish Group for the Biological Variation Database. The biological variation data critical appraisal checklist: a standard for evaluating studies on biological variation. Clin Chem 2018;64:501–14.
20. Carobene A, Aarsand AK, Bartlett WA, Coskun A, Diaz-Garzon J, Fernandez-Calle P, et al. The European Biological Variation Study (EuBIVAS): a summary report. Clin Chem Lab Med 2021;60:505–17.
21. Sandberg S, Carobene A, Bartlett B, Coskun A, Fernandez-Calle P, Jonker N, et al. Biological variation: recent development and future challenges. Clin Chem Lab Med 2022;61:741–50.
22. Arnaud J, Weber JP, Weykamp CW, Parsons PJ, Angerer J, Mairiaux E, et al. Quality specifications for the determination of copper, zinc, and selenium in human serum or plasma: evaluation of an approach based on biological and analytical variation. Clin Chem 2008;54:1892–9.
23. Thompson M. Uncertainty functions, a compact way of summarising or specifying the behaviour of analytical systems. TrAC, Trends Anal Chem 2011;30:1168–75.
24. Horwitz W, Albert R. The Horwitz ratio (HorRat): a useful index of method performance with respect to precision. J AOAC Int 2006;89:1095–109.
25. Panteghini M. Redesigning the surveillance of *in vitro* diagnostic medical devices and of medical laboratory performance by quality control in the traceability era. Clin Chem Lab Med 2023;61:759–68.
26. Côté I, Robouch P, Robouch B, Bisson D, Gamache P, LeBlanc A, et al. Determination of the standard deviation for proficiency assessment from past participant's performances. Accred Qual Assur 2012;17:389–93.
27. Praamsma ML, Arnaud J, Bisson D, Kerr S, Harrington CF, Parsons PJ. Network of Organisers of External Quality Assurance Schemes in Occupational and Environmental Laboratory Medicine. An assessment of clinical laboratory performance for the determination of manganese in blood and urine. Clin Chem Lab Med 2016;54:1921–8.
28. OELM. Occupational and environmental laboratory medicine. http://www.trace-elements.eu/ [Accessed 29 Oct 2023].
29. ISO 13485. Dispositifs médicaux – Systèmes de management de la qualité – Exigences à des fins réglementaires; 2016. https://www.iso.org/fr/standard/59752.html [Accessed 6 Mar 2024].
30. Taylor A, Angerer J, Claeys F, Kristiansen J, Mazarrasa O, Menditto A, et al. Comparison of procedures for evaluating laboratory performance in external quality assessment schemes for lead in blood and aluminum in serum demonstrates the need for common quality specifications. Clin Chem 2002;48:2000–7.
31. Oosterhuis WP, Theodorsson E. Total error vs. measurement uncertainty: revolution or evolution? Clin Chem Lab Med 2016;54:235–9.
32. Haeckel R, Wosniok W, Streichert T. Optimizing the use of the "state-of-the-art" performance criteria. Clin Chem Lab Med 2015;53:887–91.
33. Göen T, Schaller KH, Drexler H. External quality assessment of human biomonitoring in the range of environmental exposure levels. Int J Hyg Environ Health 2012;215:229–32.
34. Taylor A. Quality assessment of measurement. J Trace Elem Med Biol 2011;25:S17–21.
35. Baker MG, Simpson CD, Sheppard L, Stover B, Morton J, Cocker J, et al. Variance components of short-term biomarkers of manganese exposure in an inception cohort of welding trainees. J Trace Elem Med Biol 2015;29:123–9.
36. Ricós C, Alvarez V, Cava F, García-Lario JV, Hernández A, Jiménez CV, et al. Current databases on biological variation: pros, cons and progress. Scand J Clin Lab Invest 1999;59:491–500.
37. Delves HT, Sherlock JC, Quinn MJ. Temporal stability of blood lead concentrations in adults exposed only to environmental lead. Hum Toxicol 1984;3:279–88.
38. 10 chemicals of public health concern. World Health Organization; 2020. https://www.who.int/news-room/photo-story/photo-story-detail/10-chemicals-of-public-health-concern [Accessed 6 Mar 2024].
39. WHO guideline for clinical management of exposure to lead: executive summary. World Health Organization; 2021. https://www.who.int/publications/i/item/9789240036888 [Accessed 6 Mar 2024].
40. Fourth national report on human exposure to environmental chemicals: updated tables. National Center for Environmental Health (U.S.). Division of Laboratory Sciences; 2019, Vol 1. https://stacks.cdc.gov/view/cdc/75822 [Accessed 6 Mar 2024].
41. Becker K, Kaus S, Krause C, Lepom P, Schulz C, Seiwert M, et al. German Environmental Survey 1998 (GerES III): environmental pollutants in blood of the German population. Int J Hyg Environ Health 2002;205:297–308.
42. SCOEL/OPIN/336. Cadmium and its inorganic compounds. Scientific Committee on Occupational Exposure Limits; 2017. https://echa.europa.eu/documents/10162/35144386/182_cadmium_and_its_inorganic_compounds_oel_en.pdf/19700d5a-2c65-1827-15ca-463f612096fb?t=1699953898709 [Accessed 5 Mar 2024].
43. TLV and BEI Documentation. Arsenic and soluble inorganic compounds BEI. Cincinnatti: ACGIH; 2023.
44. U.S. EPA. Reference dose for methylmercury (external review draft). Washington, D.C.: U.S. Environmental Protection Agency; 2000. https://cfpub.epa.gov/si/si_public_record_report.cfm?Lab=NCEA&dirEntryId=20873 [Accessed 3 Apr 2024].
45. Centers for Disease Control and Prevention. Case definitions for chemical poisoning. MMWR (Morb Mortal Wkly Rep) 2005;54:12–3.

Supplementary Material: This article contains supplementary material (https://doi.org/10.1515/cclm-2024-0551).