Clinical Chemistry 48:11
2000–2007 (2002)

Laboratory Management

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

Andrew Taylor,[1*] Jurgen Angerer,[2] Francoise Claeys,[3] Jesper Kristiansen,[4] Olav Mazarrasa,[5] Antonio Menditto,[6] Marina Patriarca,[6] Alain Pineau,[7] Ilse Schoeters,[8] Christophe Sykes,[3] Sinikka Valkonen,[9] and Cas Weykamp[10]

Background: **The different scoring methods used by eight European External Quality Assessment Schemes (EQASs) for occupational and environmental laboratory medicine were compared to develop suitable quality specifications as a step toward harmonization.**
Methods: **Real results for blood lead and serum aluminum assays, reported by participants in Italian and United Kingdom EQASs, were evaluated according to individual scheme scoring criteria. The same results were then used to produce *z* scores using scheme-based between-laboratory SDs as the estimate of variability to determine whether simple performance-derived quality specifications produced better agreement among schemes.**
Results: **The schemes gave conflicting assessments of participants' performance, and participants judged to be successful by one scheme could be defined as performing inadequately by another. An approach proposed by Kenny et al. (*Scand J Clin Lab Invest* 1999;59:585), which uses clinical inputs to set targets for analytical imprecision, bias, and total error allowable, was then used to elaborate quality specifications.**
Conclusions: **We suggest that the CLIA '88 recommendations for blood lead (± 40 μg/L or ± 10% of the target concentration, whichever is the greater) could be used as a quality specification, although a revision to ± 30 μg/L or ± 10% is recommended. For serum aluminum, a suitable quality specification of ± 5 μg/L or ± 20% of the target concentration, whichever is the greater, is suggested. These specifications may be used to compare laboratory performance across schemes.**

Undue exposure to chemicals, such as metals, pesticides, and organic solvents, can be harmful to health. Occupational and environmental exposure assessment by biological monitoring is, therefore, an essential activity and is prescribed by legislation in many countries and by the European Commission *(1)*. Procedures used to analyze biological specimens are often complex; nevertheless, the provision of accurate results is crucial. Decisions with far-reaching implications are made and actions taken that are entirely dependent on the laboratory data. For this

[1] Centre for Clinical Science and Measurement, School of Biomedical and Life Sciences, University of Surrey, Guilford GU2 7XH, United Kingdom.
[2] Institute of Occupational, Social and Environmental Medicine, University of Erlangen-Nuernberg, 91054 Erlangen, Germany.
[3] Unit of Epidemiology, Scientific Institute of Public Health, B-1050 Brussels, Belgium.
[4] National Institute of Occupational Health, 2100 Copenhagen, Denmark.
[5] Higiene Industrial, Centro de Seguridad y Salud en el Trabajo, Gobierno de Cantabria, 39012 Santander, Spain.
[6] Laboratorio di Biochimica Clinica, Istituto Superiore di Sanità, 00161 Rome, Italy.
[7] Laboratoire de Toxicologie, UFR de Pharmacie, Université de Nantes, 44035 Nantes, France.
[8] Association Européenne des Metaux, 1150 Brussels, Belgium.
[9] Biomonitoring Laboratory, Topeliuksenkatu 41aA, Finnish Institute of Occupational Health, FIN-02500 Helsinki, Finland.
[10] MCA Laboratory, Queen Beatrix Hospital, 7101 BN Winterswijk, The Netherlands.
All authors are members of the Thematic Network "European Organisers of External Quality Assessment/Proficiency Testing Schemes Related to Occupational and Environmental Laboratory Medicine".
*Address correspondence to this author at: Centre for Clinical Science and Measurement, School of Biomedical and Life Sciences, University of Surrey, Guildford, Surrey GU2 7XH, United Kingdom. Fax 44-1483-689979; e-mail A.Taylor@surrey.ac.uk.
Received April 3, 2002; accepted August 8, 2002.

reason, there is increasing awareness of the importance of using properly validated methods, with each laboratory going through a process of determining its own imprecision, bias, and uncertainty *(2)*. It is accepted that truly independent assessments of laboratory performance are provided by properly constructed and managed surveillance programs or external quality assessment schemes (EQASs).[11] Schemes not only measure performance of the participants but also provide a stimulus for improvements in accuracy. Schemes relating to occupational and/or environmental laboratory medicine are organized from at least nine countries of the European Union (EU). Most use similar protocols to monitor performance, but each has its own technique to define a satisfactory standard of performance, although clear guidelines for statistical evaluation of performance have been developed *(3–5)*. Divergence occurs when organizers use techniques that are complementary to those of other "clinical EQASs" within the same country to provide for harmonization within that nation. However, a consequence of these variations in evaluating performance is that different schemes have the potential to give conflicting conclusions, even from the same raw data *(6)*.

Even within the same country, multiple criteria for the judgment of laboratory performance may exist. The criteria set for blood lead in the US by the Occupational Safety and Health Administration for occupational monitoring purposes are ± 60 μg/L or ± 15%, whichever is the greater. Clinical laboratories, however, must comply with the performance criteria for blood lead set under CLIA '88, i.e., ± 40 μg/L or ± 10%, whichever is the greater. These more restrictive criteria, enacted in 1992, were designed to support a major change in public health practice in the US in 1991, i.e., the lowering of the pediatric blood lead threshold from 250 μg/L to 100 μg/L.

In 1999, nine European organizers of EQASs for monitoring assays relevant to occupational and environmental health began formal collaboration as an EU Thematic Network. One of the main objectives of the Network was to harmonize the goals of individual schemes with respect to setting common standards for laboratory performance.

Performance of a participant on a single test item within an EQAS may be evaluated by reference to specific scheme-devised targets, to a *z* score, or to E_n numbers:

- Specific scheme targets are likely, by definition, to vary among schemes and may not be equally sensitive at detecting poor performance
- The *z* score, however, does have a clear definition within the International Organization for Standardization *(4)* and other authoritative documentation:

$$z = \frac{x - X}{s},$$

where x is the laboratory result, X is the target concentration, and s is an appropriate estimate/measure of variability that is selected to meet the requirements of the scheme organizer. The s value can be determined from data derived from the results of the participants of a particular scheme or in other ways, e.g., the SD achieved by reference laboratories *(4)*. It is likely, therefore, that *z* scores may also vary among different schemes and not necessarily be equally effective when used to detect inadequately performing participants.

- E_n numbers are similar to the *z* score, but take into account the stated uncertainty of the participant's result and the uncertainty of the target value *(4)*:

$$E_n = \frac{x - X}{\sqrt{U_{\text{lab}}^2 + U_{\text{target}}^2}},$$

where U_{lab} is the uncertainty of a laboratory's result, and U_{target} is the uncertainty of the target concentration.

The objectives of this study were to look at European EQASs for blood lead and serum aluminum determinations to:

- Compare how performance scores used by the European schemes were developed
- Compare the *z* scores that would be given by each scheme
- Apply these scores to the same sets of blood lead and serum aluminum results to determine by how much the scores will vary
- Show whether the schemes would detect performance proposed to be inadequate
- Develop suitable quality specifications that could be included in a program of harmonization of European EQASs for these assays.

Materials and Methods

Eight EQASs within the EU include the measurement of metals in blood and serum as part of their testing programs. Details of how these schemes are organized are given in a special issue of *Annali dell'Istituto Superiore di Sanità (7)*. Additional information, including an indication of the typical between-laboratory SDs observed at blood lead concentrations of 100, 400, and 700 μg/L and at serum aluminum concentrations of 100 and 200 μg/L, were provided by the scheme organizers. These concentrations were chosen because they are critical for monitoring occupational and environmental exposure to lead and aluminum.

We used results reported by participants in the Italian and United Kingdom schemes for blood lead on control samples, where target concentrations were at or close to 100, 400, and 700 μg/L, to compare the different methods used by scheme organizers to evaluate laboratory performance for blood lead analysis. The same approach was

[11] Nonstandard abbreviations: EQAS, external quality assessment scheme; EU, European Union; and TEa, total error allowable.

applied to results reported by participants for serum aluminum concentrations of 100 and 150 μg/L. Of the nine schemes participating in the Network, only six monitor performance for serum aluminum.

The results reported by participants were assessed according to (*a*) the "performance limits" used by each scheme *(7)* or (*b*) *z* scores calculated using the typical between-laboratory SDs at the appropriate concentrations after outliers were excluded according to individual scheme rules, as the *s* value *(7)*. The performance limits vary among schemes. In Denmark, limits are defined as a *z* score of ± 3. At concentrations ≥100 μg/L, the *s* value is taken as 0.1 times the target concentration. Thus, the effective limits are at ± 30% of the target concentration (lead in blood only). In Belgium, France, Italy, and the United Kingdom, limits are set by the organizers to take into account clinical needs and analytical performance among experienced laboratories. In Germany, limits are 3 times the CV associated with the results of a group of reference laboratories. In Spain, deviations of ± 60 μg/L at target concentrations <400 μg/L and ± 15% at higher concentrations are considered acceptable (lead in blood only). The EQAS in The Netherlands provides results only and does not offer any judgments as to proficiency.

Results

LEAD IN BLOOD

The performance limits set by each scheme at blood lead concentrations of 100, 400, and 700 μg/L are given in Table 1. It is evident that there are large variations among the schemes with differences up to two- and threefold at 700 and 400 μg/L, respectively, the key concentrations for suspension of workers from occupational lead exposure. However, the differences at 100 μg/L, the key pediatric threshold that is defined as harmful, are much smaller, almost in agreement at ± 30 μg/L.

The typical SDs and CVs determined from the results reported by participants in the European EQASs are also shown in Table 1. Although most laboratories participating in these European schemes use similar equipment and methodologies, it is evident that the differences in performance among participants in the schemes are quite marked: the CVs range from 7% to 21.2% at 100 μg/L, from 7% to 16.0% at 400 μg/L, and from to 9.8% to 14.3% at 700 μg/L. It has been shown that variations in performance reflect the role and responsibilities of participants and that specialist laboratories reduce bias and imprecision compared with those where the trace element activity is a minor component of the work *(8)*. It is likely that the differences in SDs evident in Table 1 will reflect the nature of the participant laboratories in individual schemes and that this in turn will reflect the approach to occupational and environmental monitoring in the particular countries. The European Network "agreed limits" refer to the decisions taken at the Second Network Meeting (held in Rome during November 2000) as being indicative of minimum acceptable analytical performance. They were developed taking into account the performances achieved by participating laboratories and were proposed as a first attempt to demonstrate how harmonization among schemes could be obtained. These limits (Table 1) are effectively ± 20% of the target concentrations, and equivalent ranges were applied to the five control samples used in this study to identify results indicating less than acceptable performance.

The upper section of Table 2 summarizes blood lead results obtained in different exercises carried out in the

Table 1. Acceptance limits currently used for measurements of lead in blood with typical between-laboratory SDs and CVs, observed in the European EQAS.[a]

	Target concentration, μg/L								
	100			400			700		
Country (n)[b]	Limits	SD, μg/L	CV, %	Limits	SD, μg/L	CV, %	Limits	SD, μg/L	CV, %
Belgium (30)	±29.0	16.9	17	±41.4	64.0	16	±53.8	8.0	14
Denmark[c] (8)	±30.0	7.0	7.0	±120.0	28.0	7.0			
France (54)	±29.0	16.3	16	±41.4	63.6	16	±53.8	100.0	14
Germany (80)	±22.2	17.0	17	±57.1	42.5	11	±84.0	68.5	9.8
Italy (66)	±20.0	21.2	21	±45.0	56.7	14	±71.0	77.8	11
Spain[d] (70)				±60.0	47.5	12	±105.0	81.6	12
The Netherlands (20)	NA[e]	14.0	14	NA	51.6	13	NA	88.2	13
United Kingdom (65)	±29.0	12.6	13	±41.4	42.3	11	±53.8	71.9	10
European Network "agreed limits"	±20.0			±80.0			±140.0		
OSHA limits	±60.0			±60.0			±105.0		
CLIA '88 limits	±40.0			±40.0			±70.0		

[a] US acceptance limits (Occupational Safety and Health Administration and CLIA '88) are shown for comparison.
[b] n = approximate number of participants.
[c] Samples at concentrations >500 μg/L are not included in the scheme.
[d] Samples at concentrations <200 μg/L are not included in the scheme.
[e] NA, not applicable; OSHA, Occupational Safety and Health Administration.

Table 2. Number and percentage of results that would be viewed as consistent with poor performance for blood lead determinations according to the performance limits used by the European EQASs, as given in Table 1.

Sample code	596	A	593	B	538
Target concentration, μg/L	99.4	119.5	399.6	412.5	722.4
No. of results	62	67	60	67	63
Range of results, μg/L	54–1536	40–475	311–592	124–718	397–869
Country, n (%)					
Belgium, France, United Kingdom	6 (9.7)	14 (20.9)	9 (15.0)	26 (38.8)	23 (36.5)
Denmark	5 (8.1)	10 (14.9)	1 (1.7)	8 (11.9)	
Germany	8 (12.9)	21 (31.3)	7 (11.7)	21 (31.3)	11 (17.5)
Italy	10 (16.1)	26 (38.8)	8 (13.3)	25 (37.3)	17 (27.0)
Spain			6 (10.0)	19 (28.4)	7 (11.1)
European Network "agreed limits"	10 (16.1)	22 (32.8)	2 (3.3)	15 (22.4)	4 (6.3)

Italian and United Kingdom schemes and used here to compare the scoring systems. The target concentrations are the consensus medians after the removal of outliers (i.e., values outside the range "mean ± 3 SD" of all reported results). These outliers are, however, included in the assessments of performance.

The five data sets were examined to determine how many individual results would be reported as unacceptable according to the criteria used by each scheme organizer, as shown in Table 1. The number of unacceptable results is given in lower part of Table 2. There is little agreement among the schemes and little overlap with the European Network agreed limits. Table 2 also illustrates the varied performance associated with the laboratories of different schemes. The percentage of "poor results" for samples with approximately the same concentrations is much higher in the Italian compared with the United Kingdom scheme, which could be a consequence of the greater number of specialized laboratories that participate in the United Kingdom scheme *(8)*. If the European Network agreed limits were applied, it can be seen that there would be a concentration-related effect, with a greater percentage of unacceptable results at ~100 μg/L compared with higher concentrations.

The variation in identifying poor performers caused by the different scheme acceptance limits is illustrated in Fig. 1. Individual results reported for sample B (target lead concentration, 412.5 μg/L) are shown with the acceptance limits for each scheme superimposed. The huge differences among the concentrations at which the results begin to fall outside the limits associated with the schemes can be clearly seen. Similar trends were obtained for the results of the other specimens.

For each scheme, *z* scores based on the between-laboratories SDs given in Table 1 were determined for the results observed for each sample. The scheme-related *z* scores for sample B are plotted in Fig. 2, and again the variations among the schemes are obvious and the number of "unacceptable" results is quite different from one scheme to another (Table 3).

ALUMINUM IN SERUM OR PLASMA

Data for assessing laboratories measuring aluminum were treated as for blood lead. A description of the procedure and the results may be seen in the data supplement to this article at *Clinical Chemistry Online* (http://www.clinchem.org/content/vol48/issue11/).

Discussion

Taylor et al. *(9)* showed that schemes adopt similar procedures for preparing and distributing test items but differ in methods of evaluating performance. Our work here demonstrates that it is possible for a participant to be

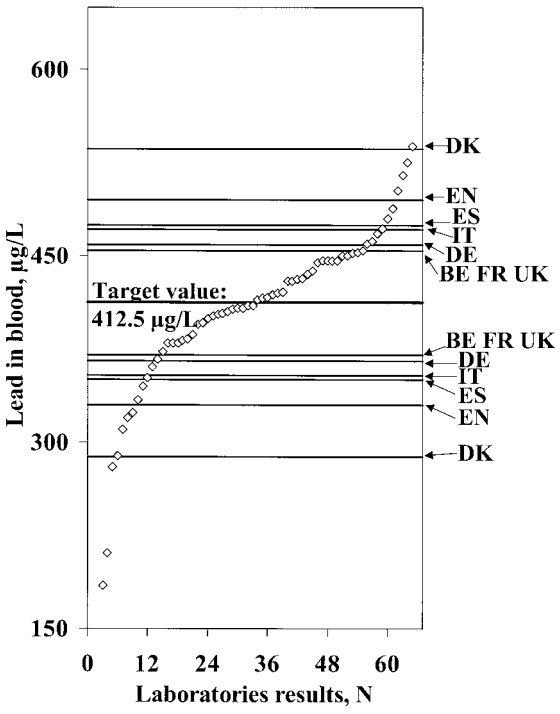

Fig. 1. Distribution of the laboratory results for blood lead in a sample with target value of 412.5 μg/L plotted against the acceptance limits of each scheme for that concentration.

DK, Denmark; *EN*, European Network; *ES*, Spain; *IT*, Italy; *DE*, Germany; *BE*, Belgium; *FR*, France; *UK*, United Kingdom. Note that some of the symbols overlap with a loss of resolution.

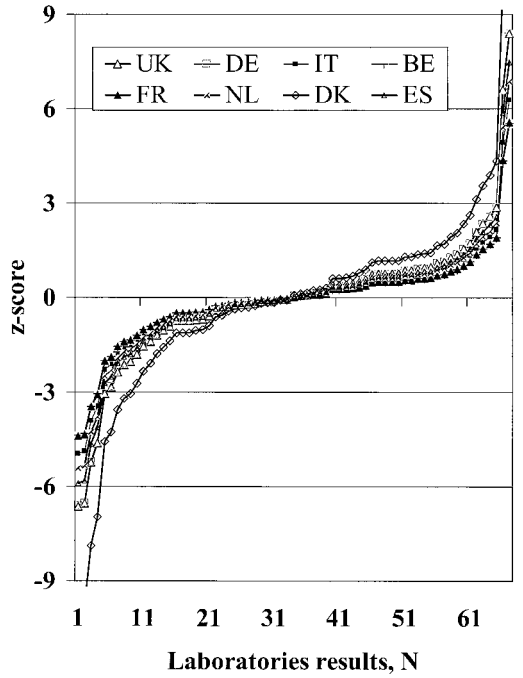

Fig. 2. Distributions of the *z* scores, determined on the basis of the SDs observed in each scheme, associated with results for blood lead in a sample with a target concentration of 412.5 μg/L.

UK, United Kingdom; *FR*, France; *DE*, Germany; *NL*, The Netherlands; *IT*, Italy; *DK*, Denmark; *BE*, Belgium; *ES*, Spain.

judged successful by one scheme but inadequate by another (see Figs. 1 and 2). These differences follow from the more exacting acceptance limits used by some organizers compared with others and are unacceptable when trying to ensure that the results of biological monitoring programs within the EU are comparable. Thus, harmonization of performance evaluation is needed to eliminate any equivocation.

One problem in setting common quality specifications is that the target value is established in different ways in different schemes, and the uncertainty of the target value may differ between schemes. In the *z* score system, it is possible only to formulate common quality specifications in terms of *s*, if it is assumed that the uncertainty of the target value is either negligible or the same in all schemes. An alternative is to evaluate performance based on the approach used in the field of calibration by use of the E_n numbers, where the uncertainty of the target value (U_{target}) is taken into consideration when evaluating performance. Hence, if analytical quality specifications were set in terms of a "laboratory uncertainty" (U_{lab}), then the E_n number will be a valid indicator of the fulfillment of the quality specifications. However, uncertainty of target values is not currently determined in most schemes.

Harmonization could be achieved if all schemes reported a *z* score as recommended by the International Organization for Standardization *(4)*, but if the *s* values in the *z* score formula were simply derived from the typical between-laboratories SDs observed in each scheme at the target concentration, this also would give dissimilar scores. As seen in Table 1, the results reported in some schemes are more closely grouped around the target concentration than in other schemes. Thus, *z* scores based on analytical data will give dissimilar outcomes unless there is agreement to use the same *s* values in the *z* score formula. When there is agreement to use the same *s* values, the *z* score will no longer represent the "consensus analytical performance" for certain schemes but will include a degree of compromise. For individual schemes, this compromise will reflect either a relaxation or an enhancement of analytical standards depending on whether the *s* values have been increased or decreased.

The analytical limits agreed on at the Second Network meeting in Rome bear little relationship to the performance targets in use by the majority of scheme organizers, who tend to recommend much smaller allowable deviation about the target concentration (Table 1). The Network limits were proposed with consideration to the minimum overall performance of all participants. It may be argued, therefore, that the EQAS performance limits are unrealistic for many of the participants to achieve and that they do not represent the "analytical situation" in real laboratories. Alternatively, it may be suggested that limits, which simply reflect the broad spectrum of performance, neither eliminate poor laboratories nor reward those who are competent. It has been demonstrated that performance

Table 3. Number and percentage of results for lead in blood that would have a *z* score >3 if typical between-laboratory SDs as observed in each scheme (Table 2) were used to determine *z* scores.

Sample code	596	A	593	B	538
Target concentration, μg/L	99.4	119.5	399.6	412.5	722.4
Country, n (%)					
Belgium	3 (4.8)	5 (7.5)	1 (1.7)	6 (9.0)	1 (1.6)
Denmark	8 (12.9)	20 (29.9)	2 (3.3)	15 (22.4)	
France	3 (4.8)	6 (9.0)	1 (1.7)	6 (9.0)	1 (1.6)
Germany	3 (4.8)	6 (9.0)	1 (1.7)	7 (10.4)	2 (3.2)
Italy	3 (4.8)	4 (6.0)	1 (1.7)	6 (9.0)	1 (1.6)
Spain			1 (1.7)	6 (9.0)	1 (1.6)
The Netherlands	4 (6.5)	9 (13.4)	1 (1.7)	6 (9.0)	1 (1.6)
United Kingdom	5 (8.1)	10 (14.9)	1 (1.7)	7 (10.4)	2 (3.2)

targets, which provide a realistic and achievable challenge to participants, have a positive effect on the overall standard of performance *(10)*.

An alternative to simple considerations of interlaboratory variance in setting quality specifications has been developed by Kenny et al. *(11)* and may be useful to this discussion. This uses a systematic hierarchical approach that takes account of analytical and clinical targets. To develop quality specifications that are suitable for harmonization of European EQASs for measurement of lead in blood and aluminum in serum, we have applied this approach to our data and other observations.

GOALS BASED ON CURRENT STATE OF THE ART

As demonstrated by data from EQASs. This information is given in Table 1, and there is some consensus with CVs at ~15%.

Other data. Good performance is shown by results from reference laboratories. For example, the United Kingdom trace element Supraregional Assay Service Centres and the specialist laboratories used by the German and New York State EQASs achieve between-laboratory CVs that are <10% at 100 μg/L and <5% at higher concentrations (data from scheme reports presented in 2000 and 2001).

QUALITY SPECIFICATIONS SET BY

Regulatory bodies. Many regulatory organizations state that laboratories undertaking measurements must meet the criteria for competence set by the EQAS organizer (see below). Specifications, with legislative authority, have been defined for certain applications, but these are usually established to eliminate very poorly performing laboratories rather than to stimulate improvement or with consideration to the clinical relevance of the assay. Thus, the German occupational biological monitoring program allows for limits that are three times the SD achieved by reference laboratories (equivalent to 22.2% at 100 μg/L, decreasing to 12.0% at 700 μg/L), whereas the US CLIA '88 *(12)* defines acceptable performance for lead in blood as being within ± 40 μg/L or ± 10% of the target concentration, whichever is the greater, and the criteria set by the US Occupational Safety and Health Administration for occupational monitoring purposes are ± 60 μg/L or ± 15%, whichever is greater.

Organizers of EQASs. Goals set by the organizers of the European schemes are indicated in Table 1. As noted previously, some are based only on analytical grounds, whereas others attempt to also account for the clinical usefulness of the test *(13, 14)*. These goals correspond to CVs of ~8–30%, even at the higher concentration of 700 μg/L.

PUBLISHED PROFESSIONAL RECOMMENDATIONS

The US NCCLS and the US CDC currently recommend that the specification for internal quality-control limits should be ± 20 μg/L or ± 10%, whichever is greater *(15, 16)*. To our knowledge, there are no other published recommendations from expert bodies, groups, or individuals for quality specifications in occupational and environmental laboratory medicine.

EFFECT OF ANALYTICAL PERFORMANCE ON CLINICAL DECISIONS

Data based on biological variation. The observed values of a biomarker of exposure, such as lead in blood, depend on the individual exposure and are expected to change to reflect changes in exposure. However, if the extent of exposure can be considered constant, some information can be obtained on the biological variability for such a biomarker. A few studies involving individuals not occupationally exposed to lead have investigated seasonal variability of blood lead concentrations by comparing the values obtained in samples collected over a period of time *(17–20)*. When seasonal variability can be excluded, these series of data can be used to derive information on inter- and intraindividual variability of blood lead. Delves et al. *(17)* have reported on the temporal stability of blood lead concentrations of 21 healthy adults (14 men and 7 women), exposed only to environmental lead, in the UK between 1981 and 1982. For these individuals, 9–17 blood samples were collected serially over a period from 7 to 11 months. The mean (SD) blood lead concentration was 120 ± 22.4 μg/L (18.7%), and the intraindividual variation ranged from 1.4% to 9.1%, with a mean of 4.5%. In a series of reports, Fraser et al. *(20)* have suggested desirable targets for analytical imprecision (CV_a), bias, and total error allowable (TEa) derived from biological variation and expressed by the following formulas:

$$CV_a\ (\%) = 0.5(CV_{intra})$$

$$\text{Bias}\ (\%) = 0.25(CV_{intra}^2 + CV_{inter}^2)^{1/2}$$

$$\text{TEa}\ (\%) < 0.25(CV_{intra}^2 + CV_{inter}^2)^{1/2} + z(0.5 \times CV_{intra})$$

where $z = 1.65$ for a 95% probability level. Similar formulas can be applied to derive optimal and minimal targets of performance. Using the data provided by Delves et al. *(17)*, we have calculated such targets for the determination of lead in blood (Table 4).

Data based on clinician opinions. We are unaware of any published opinions relating to lead in blood performance

Table 4. Optimal, desirable, and minimal performance targets for analytical imprecision (CV_a, %), bias, and total error for the determination of blood lead based on biological variation data.

	Optimal	Desirable	Minimal
CV_a, %	1.1	2.3	3.4
Bias, %	2.4	4.8	7.2
TEa, %	4.3	8.5	12.8

criteria for occupational and environmental laboratory medicine.

effect of analytical performance on clinical outcomes in specific clinical settings

Relationships between the concentration of lead in blood and the onset of clinical signs and symptoms have been presented by various authors *(21–23)*. Although there are differences among individuals, on a population basis it can be seen that the activities of the erythrocyte enzymes pyrimidine 5′-nucleotidase and δ-aminolevulinic acid dehydratase are inhibited at lead concentrations of ~100 and 200 μg/L, respectively, that erythrocyte porphyrins are increased and nerve conduction velocity is decreased at ~300 μg/L, that urinary δ-aminolevulinic acid is increased and sperm morphology and function are affected at ~400 μg/L, and that hemoglobin concentrations begin to decrease at ~500 μg/L. Thus, although it may be a relatively crude observation, most biochemical and health effects in adults do appear at apparently regular intervals of ~100 μg/L. Therefore, for this assay a suitable quality specification for total error should at least allow differentiation between blood lead concentrations that are 100 μg/L apart; therefore, the maximum TEa should not exceed ± 50 μg/L. However, more restrictive limits may be required at the lower concentrations, which are critical for the monitoring of pediatric populations. Better performance at these lower concentrations is easily achievable by most laboratories and is desirable to reduce misclassification.

proposed standard of performance for the measurement of lead in blood

From the previous discussion, it appears that a performance target for the determination of lead in blood could be established taking into account biological variability and the effect of analytical performance on clinical outcomes. From the former criterion, an indication for a minimum performance for TEa of 12.8% was derived from the data available at a concentration of 120 μg/L. From the latter, it can be stated that if the total maximum acceptable error should be ± 50 μg/L, the target SD for a group of laboratories performing blood lead analysis should be ≤16 μg/L if all of them were to achieve acceptable results. According to this, 95% of the laboratories would achieve results within ± 32 μg/L from the target value and 99.7% of them would provide results within ± 48 μg/L from the target value. Fraser *(20)* remarked that, for the purposes of clinical investigations, it is not necessary to achieve accuracy that is better than the TEa. On this basis, the performance target does not alter with time (unless there is a change in clinical constraints). It should be noted however that (*a*) there is evidence for modern analytical methods to allow better performance *(24)*, (*b*) many laboratories are already able to achieve better performance (Table 2), and (*c*) professional organizations have recommended more rigorous targets (for internal quality-control purposes). The CLIA ′88 criteria for acceptable results, i.e., ± 40 μg/L or ± 10% of the target concentration, whichever is the greater, correspond closely to the TEa calculated above, and we suggest that the CLIA ′88 targets could be used as an analytical goal, although a revision to ± 30 μg/L or ± 10%, whichever is the greater, is recommended.

proposed standard of performance for the measurement of aluminum in serum

A similar discussion applied to measurement of aluminum in serum is presented in the data supplement to this article at *Clinical Chemistry Online* (http://www.clinchem.org/content/vol48/issue11/). From that information we propose that for this assay, a suitable quality specification for TEa would be ± 5 μg/L or ± 20%, whichever the greater. This is equivalent to a TEa of ± 5 μg/L up to 25 μg/L and a variation of ± 20% for higher concentrations.

This work was supported by the Commission of the European Union under Contract No. SMT4-CT98-7522.

References

1. European Union. Council Directive 98/24/EC of 7 April 1998 on the protection of the health and safety of workers from the risks related to chemical agents at work [fourteenth individual Directive within the meaning of Article 16(1) of Directive 89/391/EEC]. Off J L 1998;131:11–23.
2. International Organization for Standardization. ISO guide to the expression of uncertainty in measurement. Geneva: ISO, 1993.
3. Thompson M, Wood R. International harmonized protocol for proficiency testing of (chemical) analytical laboratories. J AOAC Int 1993;67:926–36.
4. International Organization for Standardization. ISO guide 43-1. Proficiency testing by interlaboratory comparisons. Part 1: development and operation of proficiency testing schemes, 2nd ed. Geneva: ISO, 1997.
5. International Laboratory Accreditation Cooperation. ILAC-G13. Guidelines for the requirements for the competence of providers of proficiency testing schemes. Rhodes, Australia: ILAC, 2000.
6. Christensen JM, Olsen E. Comparison of some European external quality assessment schemes in the field of occupational and environmental medicine. Ann Ist Super Sanita 1996;32:285–94.
7. Morisi G, Menditto A, Patriarca M, Taylor A. European external quality assessment schemes in occupational and environmental medicine. Ann Ist Super Sanita 1996;32:191–316.
8. Taylor A. Reference materials for measurement of aluminium in biological samples. Fresenius Z Anal Chem 1988;332:616–9.
9. Taylor A, Patriarca M, Menditto A, Morisi G. Prospects of harmonisation of European external quality assessment schemes in occupational and environmental laboratory medicine. Ann Ist Super Sanita 1996;32:295–307.
10. Taylor A. Reference materials and analytical standards to stimulate improved laboratory performance: experience from the external quality assessment scheme for trace elements in biological samples. Mikrochim Acta 1996;123:251–60.
11. Kenny D, Fraser CG, Hyltoft Petersen P, Kallner A. Consensus agreement. Scand J Clin Lab Invest 1999;59:585.
12. US Department of Health and Human Services. Medicare, Medic-

aid and the CLIA programs; regulations implementing the Clinical Laboratory Improvement Amendments of 1988 (CLIA). Final rule. Fed Regist 1992;57:149–62.

13. Taylor A, Starkey BJ, Walker AW. Measurement of aluminium—quality of analytical results. In: Taylor A, ed Aluminium and other trace elements in renal disease. Eastbourne: Balliere Tindall, 1986;260–73.

14. Taylor A, Briggs RJ. An external quality assessment scheme for trace elements in biological fluids. J Anal Atomic Spectrom 1986;1:391–5.

15. National Committee for Clinical Laboratory Standards. Analytical procedures for the determination of lead in blood and urine; approved guideline. NCCLS Document C40-A. Wayne, PA: NCCLS, 2001.

16. Centers for Disease Control and Prevention. Screening young children for lead poisoning: guidance for state and local public health officials, Appendix C.1. Atlanta, GA: CDC, 1997.

17. 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.

18. Watanabe T, Fujita H, Koizumi A, Chiba K, Miyasaka M, Ikeda M. Baseline level of blood lead concentration among Japanese farmers. Arch Environ Health 1985;40:170–6.

19. Gulson BL, Mizon KJ, Korsch MJ, Mahaffey KR, Pederson D. Limited seasonality effects on blood lead for a small cohort of female adults and children. Sci Total Environ 2000;253:119–26.

20. Fraser CG. General strategies to set quality specifications for reliability performance characteristics. Scand J Clin Lab Invest 1999;59:487–90.

21. International Programme on Chemical Safety (IPCS). WHO Task Group on Environmental Health criteria for lead. Lead. Environmental Health Criteria 3. Geneva: WHO, 1977.

22. International Programme on Chemical Safety (IPCS). WHO Task Group on Environmental Health criteria for lead. Inorganic lead. Environmental Health Criteria 165. Geneva: WHO, 1995.

23. Goyer RA. Lead. In: Wermer M, ed. CRC handbook of clinical chemistry, Vol. IV. Boca Raton, FL: CRC Press, 1989:229–42.

24. Parsons PJ, Geraghty C, Verostek MF. An assessment of contemporary atomic spectroscopic techniques for the determination of lead in blood and urine matrices. Spectrochim Acta Part B 2001;56:1593–604.

25. Nordal KP, Dahl E, Thomassen Y, Brodwall EK, Halse J. Seasonal variations in serum aluminium concentrations. Pharmacol Toxicol 1988;62:80–3. (Cited in Supplemental Data.)