Comparison of some European external quality assessment schemes in the field of occupational and environmental medicine

Jytte Molin CHRISTENSEN and Erik OLSEN

Department of Chemistry and Biochemistry, National Institute of Occupational Health, Copenhagen, Denmark

Summary. - In most European countries an increasing number of external quality assessment schemes (EQAS) are organized, and it seems appropriate to reinforce collaboration at the European level between organizers of EQAS related to occupational and environmental medicine. Since differences between these EQAS have been recognized, a collaborative project was set up focused on the ways the present occupational and environmental medicine EQAS evaluate results obtained by the same pool of laboratories analysing identical control samples for blood lead. The results confirmed that the samples delivered to the laboratories were homogeneous. Considering the performance as judged by five different schemes the study revealed that laboratories were not ranked identically. For laboratories, which either had a very bad or a very good performance, however, the ranking were comparable. The statistical design of the evaluated EQAS poses problems and requires attention.

Key words: inter-scheme comparison, laboratory performance evaluation, external quality control, blood lead, occupational medicine, environmental medicine.

Riassunto *(Confronto tra alcuni schemi europei di valutazione esterna di qualità nel settore della medicina ambientale ed occupazionale).* - Nella maggior parte dei paesi europei viene organizzato un numero crescente di schemi di valutazione esterna di qualità (VEQ) ed è pertanto opportuno rafforzare la collaborazione a livello europeo tra organizzatori di programmi di VEQ nel settore della medicina occupazionale e ambientale. Poiché esistono differenze tra gli schemi nazionali di VEQ, è stato organizzato un progetto collaborativo con lo scopo di confrontare i metodi usati nei vari programmi per valutare i risultati ottenuti da un gruppo di laboratori nella determinazione del piombo nel sangue in identici campioni di controllo. Questo studio ha dimostrato che i diversi schemi di VEQ non valutavano tutti i laboratori nello stesso modo. Tuttavia, laboratori con ottime o pessime prestazioni analitiche erano valutati analogamente da tutti gli schemi considerati. In conclusione, tutti gli schemi di VEQ considerati sono strutturati in modo simile. Tuttavia, l'approccio statistico e i criteri di valutazione delle prestazioni dei laboratori pongono dei problemi e devono essere considerati con attenzione.

Parole chiave: valutazione esterna di qualità, confronto tra schemi, piombo nel sangue, medicina occupazionale. ed ambientale.

Introduction

Reliable measurement and test results are of paramount importance for any ordered society. This holds true also with respect to unions of countries such as the European Union (EU). Consequently, in an internal market, harmonization of analytical performance is needed. An ongoing monitoring of analytical performance is provided by external quality assurance systems. External evaluation supplied by comparisons between analytical results of different laboratories (external quality assessment schemes, EQAS) measures the between-laboratory differences.

The application of legislation related to health and safety of workers at the workplace requires a spectrum of actions. These include the monitoring of workplace air for potential physical, chemical or biological hazards in order to assess and reduce the exposure. A complementary method is the examination of blood/urine specimens of the individual worker often related to biological monitoring programmes. For such biological measurements, data to be sound and reliable must have been obtained under a "good" quality assurance system including internal and external quality control [1, 2].

Participation in EQAS provide laboratories with an objective demonstration of the reliability of the data they produce. From a community point of view, it is important that such schemes evaluate laboratories identically. However, the main purpose of EQAS is not only to highlight repeatability and reproducibility performance between laboratories, but also to assess systematic errors [3-6]. Various terms may be used to describe schemes for external quality assessment, e.g., external quality control, performance schemes, proficiency testing, etc. Although

there are several types of schemes, they all share the common feature of comparison of a laboratory's results with those of other laboratories. A prerequisite for obtaining equal performance is that accreditation bodies are "judging EQAS test results equally" [7, 8].

In most European countries an increasing number of EQAS for biological measurements are organized [9-12]. However, differences between the EQAS have been recognized partly due to the use of different statistical methods, that potentially may give conflicting conclusions for different EQAS, even from the same raw data. Therefore, it seems appropriate to reinforce collaboration at the European level between organizers of EQAS related to occupational and environmental medicine.

The European Commission Directorate General XII - Science, Research and Development - is running a programme, the Standards, Measurements & Testing (SM&T) programme, former Bureau Communautaire de Référence (BCR) dedicated to the improvement of both ability to obtain and reliability of measurement results. Within the framework of this programme a project on "Collaboration between European organizers of EQAS in the field of occupational and environmental medicine" was initiated at a workshop organized by SM&T. Since differences between EQAS have been recognized, it was agreed that a collaborative effort should focus on the ways the present EQAS evaluate results reported for the circulated control samples, i.e. comparability of the classification of individual analytical performance.

Study design

The aim of the project was to compare the methods of evaluation of laboratory performance applied in different EQAS with a view to ensure that a specific laboratory performance should be judged and equally classified by the various schemes. To achieve this it was proposed that the same pool of analytical results produced by a group of selected European laboratories should be examined by each scheme according to their statistical procedures and criteria for evaluation of performance. In order to involve all EU member states and three EFTA countries it was proposed to carry out an intercomparison study on the determination of blood lead with the participation of two laboratories from each country and/or scheme. In total 32 high ranking European laboratories participated in the project and their respective contributions provided the pool of data for evaluation by each scheme organizer.

The EQAS run by the following organizations from six different countries were evaluated in the study:

- National Institute of Occupational Health (NIOH/DEQAS), Lersø Parkallé 105, DK-2100 Copenhagen, Denmark (DK);
- Institut für Arbeits - und Sozialmedizin (INAS), Universität Erlangen-Nürnberg, Schillerstrasse 25, D-8520 Erlangen, Germany (DE);
- Institut National de Recherche et de Sécurité (INRS), Av. de Bourgogne B.P. 27, F-54 501 Vandoeuvre Cedex, France (FR);
- Instituto Nacional de Seguridad e Higiene en el Trabajo (INSHT), Gabinete Técnico Provencial Zaragoza, B. Ramazzini s/n, E-50014 Zaragoza, Spain (ES);
- Robens Institute of Health and Safety (RIHS), Trace Element Laboratory, University of Surrey, Guildford, GU2 5XH, Great Britain (GB);
- Toxicology Laboratory (TOXL/SKZL), University Hospital of Leiden, P.O. Box 9600, NL-2300 RC Leiden, The Netherlands (NL).

Table 1 gives an overview of the main features of the different EQAS [13-20] compared with the requirements of the ISO 5725 [3-5].

The NIOH, Denmark, produced the control samples at five different concentrations from outdated human whole blood stabilized with glucose and a citrate/phosphate buffer. The five different blood lead concentrations were prepared by spiking a standard lead solution (Merck Chemical Tritisol No. 9969, 1 g/l) to the blood. The spiked amounts of lead to the blood were: 0 mmol/l; 0.19 mmol/l; 0.51 mmol/l; 1.02 mmol/l and 1.53 mmol/l for the samples B1001, B1002, B1003, B1004 and B1005, respectively. Samples were produced by accurately pipetting 3.00 ml of the blood controls directly into glass vials. The blood samples were then lyophilized using a Hetosic Freeze Dryer, type CD 12 with an ice capacity of 6 kg ice/24 hours and a maximum batch capacity of 600 vials. All blood handling steps were carried out under clean room conditions to avoid contamination.

Optimal homogeneity between reconstituted vials requires that the filling procedure before lyophilization gives identical amounts of blood in each vial. The difference accepted by weight in the preparation was ± 0.1%. Homogeneity between vials within the same batch was secured by well documented treatment at all steps of production [19]. Systematic errors and random events of the manufacturing procedure were statistically tested for by measuring lead in randomly selected vials from each of the five blood lead levels [10, 19].

One week before sample dispatch the participating laboratories received information and instructions for appropriate handling of the samples. The laboratories were asked to reconstitute the lyophilized samples in 3.00 ml Milli Q water pipetted into the vial and then closed tightly with the stopper and kept over night before

gently rolling for 10 min. The samples B1001 (coded 1 and 6), B1002 (coded 2 and 7), B1003 (coded 3 and 8), B1004 (coded 4 and 9) and B1005 (coded 5 and 10) were requested to be analysed and treated as normal routine samples, and all 10 results were reported on the enclosed report sheet. The EQAS samples (in total 320 vials) were distributed in a single delivery to the participating laboratories together with the reporting sheet. The laboratories were requested to analyse the ten EQAS samples (coded 1-10) within two weeks. The results were mailed to SM&T, where the result sheets were coded to provide anonymity, before the results were dispatched to scheme organizers. Each EQAS organizer processed the data according to their usual procedures. The participating laboratories received reports from their national scheme organizer as usual. Discussion meetings between EQAS organizers were planned and completed in Dublin, April 1994 and in Rome, December 1994. In order to compare the performance characteristics, as judged by different schemes, the participating laboratories were evaluated by five of the six different EQAS using their performance indices.

In DEQAS (NIOH) a performance index is calculated based on the results of $RMSE^{1/2}$ at five different concentration levels, i.e.

$$\text{Performance index DEQAS} = \sum_{i=1}^{5} \frac{RMSE^{1/2}_{Level(i)}}{RMSE^{1/2}_{Ideal(i)}}$$

The square root of the relative mean square error is defined as:

$$RMSE^{1/2} = \sqrt{\frac{MSE(\mu)}{\mu^2}}$$

where MSE is the sum of the systematic error and the overall standard deviation of the analytical method [13, 20].

In the German scheme (INAS) the participants receive a certificate if both results are within the tolerance range (3 s-range) [9, 14]. This scheme uses assigned values estimated by high ranking reference laboratories. The certification is valid for one year.

The Spanish scheme (INSHT) calculates "variance index", which is derived from the difference between the result returned by a participant and the consensus mean, expressed as a percentage of the consensus mean [16].

The Guildford (UK) scheme calculates a performance score based on differences between the consensus medians and reported concentrations, evaluated against inner and outer zones plotted on the chart. A point that falls within the inner zone, scores 2 while a point in the outer zone

Table 1. - Overview of EQA schemes

Scheme organizer	Presentation of data	No. of concentration levels	Outlier test	Consensus values	Assigned values	Performance index	Trials per year	Reference
NIOH (DK)	Ratio plot	5	Tolerance interval	Adopted	Adopted	RMSE [a]	2	[10, 13]
INAS (DE)	-	2	Tolerance interval	-	Adopted	Certificate	> 1	[9, 14]
INRS (FR)	Bar plot	2	± 3 SD	Adopted	-	Scores based on taget zones	6	[15]
INSHT (ES)	Bar plot	3	± 2 SD	Adopted	Adopted	Variance index	12	[16]
RIHS (UK)	Bar plot	3	± 3 SD	Adopted	-	Scores based on target zones	12	[17]
TOXL (NL)	Scatter plot	3	± 3 SD	Adopted	-	-	4	[18]
ISO 5725 (part 4/5)	Bar plot	5	Grubbs Cochran	-	Adopted	-	-	[3-5]

(a) Relative mean square error.

score 1 [17]. Scores for the five results are added together. Thus a good laboratory will achieve a high score and the maximum in this study is 10. From previous work with blood lead measurement it is suggested that a competent laboratory should achieve a performance score of 7 or more.

The French scheme (INRS) has criteria of acceptability which are also derived from target zones [15], where the limits to the zone are described as "limit of goodness" and "limit of acceptance". Differences between the mean value and the analytical result are compared to those limits and scored accordingly. In addition, the system evaluates the recovery of addition and between run precision.

Results and discussion

In the present interscheme comparison study on lead in human blood the analytical techniques were flame or electrothermal atomic absorption with standard addition or standard graph (standard curve) for calibration. Four laboratories applied flame-AAS, 28 laboratories used the ETA-AAS technique. For calibration 21 of 32 laboratories applied the standard graph method and 11 of 32 laboratories applied the method of standard addition.

Table 2 presents the original measured values. Three laboratories, n. 8, 53 and 66, clearly failed to carry out the study correctly and had either not completed the analyses as directed or had misreported their results. Scheme

Table 2. - Reported values of blood lead in the collaborative study

| Lab-code | Original data (µmol/l) | | | | |
	Level 1 (B1001)	Level 2 (B1002)	Level 3 (B1003)	Level 4 (B1004)	Level 5 (B1005)
2	0.237 0.250	0.490 0.503	0.913 0.953	1.595 1.630	2.293 2.213
3	0.243 0.251	0.463 0.434	0.759 0.753	1.197 1.178	1.665 1.639
6	0.180 0.300	0.400 0.440	0.790 0.850	1.340 1.430	2.040 2.020
8		0.180 0.190	0.540 0.530	0.970 1.110	1.840 1.880
9	0.160 0.150	0.360 0.330	0.960 0.730	1.260 1.210	1.780 1.700
15	0.193 0.170	0.386 0.425	0.772 0.796	0.965 0.917	1.544 1.593
22	1.038 1.062	0.820 0.796	1.284 1.236	1.361 1.406	1.882 1.931
25	0.140 0.160	0.300 0.320	0.710 0.710	1.240 1.250	1.760 1.740
28	0.200 0.210	0.410 0.390	0.800 0.740	1.380 1.380	1.960 1.950
31	0.299 0.328	0.526 0.521	0.795 0.812	1.298 1.327	1.786 1.804
32	0.170 0.170	0.350 0.340	0.780 0.800	1.310 1.370	1.930 1.830
34	0.087 0.111	0.275 0.275	0.603 0.656	1.095 1.149	1.583 1.665
35	0.150	0.299	0.666	0.888	1.467
36	0.160 0.150	0.350 0.380	0.680 0.710	1.250 1.280	1.730 1.740
39	0.050 0.050	0.130 0.140	0.510 0.480	0.880 0.930	1.430 1.370
40	0.155 0.149	0.301 0.302	0.606 0.612	1.122 1.114	1.617 1.617
41	0.220 0.160	0.400 0.300	0.700 0.650	1.170 1.120	1.630 1.630
42	0.266 0.261	0.442 0.457	0.692 0.721	1.055 1.041	1.510 1.520
49	0.480 0.480	0.480 0.528	0.624 0.624	1.248 1.200	1.824 1.824
53	2.200 1.900	5.900 6.200	13.10 13.50	21.60 21.90	31.80 28.50
57	0.168 0.176	0.387 0.393	0.766 0.777	1.269 1.264	1.825 1.824
59	0.190 0.160	0.350 0.250	0.670 0.670	0.970 1.120	1.710 1.690
65	0.150 0.140	0.360 0.320	0.710 0.660	1.020 1.050	1.410 1.380
66	0.184 1.747	0.197 1.805	0.190 1.765	0.163 1.752	0.172 1.759
67	0.210 0.210	0.420 0.440	0.860 0.850	1.440 1.460	2.040 2.030
71	0.080 0.090	0.270 0.310	0.630 0.690	1.210 1.230	1.770 1.730
77	0.292 0.332	0.961 1.007	1.123 1.167	2.199 2.339	2.477 2.635
83	0.194 0.149	0.379 0.385	0.812 0.875	1.297 1.445	1.843 2.006
87	0.157 0.152	0.356 0.359	0.751 0.773	1.245 1.260	2.120 2.110
90	0.162 0.169	0.343 0.336	0.691 0.688	1.054 1.063	1.638 1.600
94	0.197 0.203	0.375 0.365	0.718 0.742	1.337 1.303	1.881 1.919
96	0.190 0.190	0.330 0.370	0.680 0.740	1.170 1.230	1.730 1.750

Table 3. - Inter-scheme comparison (sample B1001 (µmol/l)). Lead in human blood

Scheme organizer	Spiked amount	Consensus mean	Assigned value	CV (%)	Number of outliers
NIOH (DK)	0	0.18	0.18	33	3
INAS (DE)	0	0.17	0.16	12	14
INRS (FR)	0	0.19	-	28	5
INSHT (ES)	0	0.19	0.16	31	8
RIHS (UK)	0	0.19	-	42	4
TOXL (NL)	0	0.18	-	27 [a]	3
ISO 5725, part 4 and 5	0	0.18	0.18	34	7

(a) All levels.

Table 4. - Inter-scheme comparison (sample B1003 (µmol/l)). Lead in human blood

Scheme organizer	Spiked amount	Spiked amount + consensus mean for B1001	Consensus mean	Assigned value	CV (%)	Number of outliers
NIOH (DK)	0.51	0.69	0.74	0.75	17	2
INAS (DE)	0.51	0.68	0.71	0.71	8	9
INRS (FR)	0.51	0.70	0.73	-	16	3
INSHT (ES)	0.51	0.70	0.72	0.69	13	6
RIHS (UK)	0.51	0.70	0.74	-	16	4
TOXL (NL)	0.51	0.69	0.73	-	27 [a]	2
ISO 5725 part 4/5	0.51	0.69	0.72	0.75	13	6

(a) All levels.

organisers received the data as given in Table 2 and followed their own protocol to determine which results should be excluded, without any direction from the study organisers.

Tables 3-5 show the mean values, assigned values and coefficients of variation for the samples B1001, B1003 and B1004, calculated by the different EQAS organizers. The comparable figures for samples B1002 and B1005 have been omitted to avoid this presentation becoming too long. Copies of all data can be obtained from the authors.

Further, for comparison the results have been evaluated according to ISO 5725 part 4 and part 5 [4, 5], using a Windows version of the statistical software package AMIQAS [20]. ISO 5725 part 4 provides basic methods for estimating "the bias of a measurement

Table 5. - Inter-scheme comparison (sample B1004 (µmol/l)). Lead in human blood

Scheme organizer	Spiked amount	Spiked amount + consensus mean for B1001	Consensus mean	Assigned value	CV (%)	Number of outliers
NIOH (DK)	1.02	1.20	1.20	1.27	14	1
INAS (DE)	1.02	1.19	1.28	1.27	6	15
INRS (FR)	1.02	1.21	1.19	-	14	3
INSHT (ES)	1.02	1.21	1.21	1.13	14	4
RIHS (UK)	1.02	1.21	1.21	-	14	1
TOXL (NL)	1.02	1.20	1.23	-	27 [a]	4
ISO 5725 part 4/5	1.02	-	1.22	1.27	13	4

(a) All levels.

method and laboratory bias when a measurement method is applied to an interlaboratory experiment" [4]. In ISO 5725 part 5 alternative methods for determining the repeatability and reproducibility standard deviations, namely the split-level design, are described [5]. In general, ISO 5725 requires all results to be presented in a table, with plots of the distribution at each concentration level, two sample plots to demonstrate bias of data (Youden Plot), and a regression line between measured (y) and target values (x), for each laboratory (at least three concentration levels). The ISO 5725 standard does not give any recommendation for a performance score system, but the z-score is preferred by an AOAC/IUPAC/ISO Working Party [7]. The z-score is calculated according to the following formula:

$$z = (Y_i - \mu)/\sigma_\mu$$

where Y_i is the i'th result and μ is a good estimate of the true value and σ_μ a good estimate of the standard deviation.

Because the performance indices used in the five EQAS are so different, ranking has been used, whenever possible to compare the position achieved by each laboratory according to the different evaluation systems. A comparison of the rankings obtained in the different schemes is demonstrated in Table 6. Altogether, some agreement between the scheme evaluations exist, approximately 40% of the results agreed on the performance evaluation although the score systems definitely varies.

Table 6 shows that a fairly good agreement in the ranking exists for laboratories with lab-code numbers 25, 36, 39, 41, 57 and 96, however different types of performance indices, i.e. ranks, scores and certificate achievement are compared, therefore full comparison is not possible.

The ISO 5725 evaluation (regression lines) for selected laboratories is presented in Fig. 1 and the regression line for all laboratories is shown in Fig. 2. The three laboratories (no. 8, 53 and 66) which, as mentioned above, reported clearly anomalous results were excluded before the statistical analysis of the results. In Fig. 1 the plots illustrate that good performance exists, i.e. the EQAS results are excellent for laboratory number 57, while the results clearly indicate an analytical bias at laboratory number 39. For laboratory number 90 a bias is evident at higher concentration levels, a situation resulting in great differences in the performance scores (Table 6). Laboratory number 9 had a fairly good score in most of the schemes, but the one which used the variance index for its judgement evaluated this laboratory differently. Furthermore, the results reported by laboratory number 22 clearly illustrate the need for several concentration levels in EQAS.

An analysis of variance (ANOVA) of all data, except the results from the three above mentioned outlying laboratories, demonstrated significant difference between laboratories (p < 0.05). In contrast, no significant differences between pairs of results were revealed at a five percent significance level, confirming that the samples delivered to the laboratories were homogeneous. Considering the performance of individual

Table 6. - Inter-scheme comparison of lead in blood. Rankings and other indices of performance scores

Lab-code number	Rankings in the schemes				
	NIOH (DK)	INAS [1] (DE)	INRS (FR)	INSHT (ES)	RIHS [2] (UK)
57 [*]	1	Y	1	3 [g]	10
96 [*]	2	Y	5	1 [g]	10
09	3	Y	2	7 [g]	10
36 [*]	4	Y	3	6 [g]	10
25 [*]	7	Y	4	8 [g]	10
32	5	Y	9	5 [g]	8
94	6	Y	6	2 [g]	8
28	9	Y	12	12	7
41 [*]	10	N	10	11	9
03	11	N	8	17	8
90	14	N	7	4 [g]	8
71	15	N	11	20 [p]	9
83	8	N	20	13	7
40	18	N	13	15	7
59	20	N	17	9	8
87	23	N	16	10	8
06	12	N	23	22 [p]	5
67	13	N	22	21 [p]	5
39 [*]	28	N	29	29 [p]	1

(*) Fairly good agreement between ranking and other indices of performance used in the various schemes.
(1) Certification: Yes (Y) or No (N); *(2)* Performance score: a competent laboratory should score > 7.
(g) good performance; *(p)* poor performance.

laboratories, occasionally mistakes occur, even though the laboratories were informed that samples were to be analysed and evaluated in a proficiency test. It is to be expected that such mistakes also occur during daily routine analysis. Some laboratories performed well at higher concentration levels but less satisfactory at lower concentrations, indicating that their method was not valid over the entire range necessary for measuring lead in blood for assessment of exposure in the working environment. The majority of the laboratories, however, had a fairly good performance.

Conclusions

Thirty-two laboratories reported results for lead in human blood at five different lead concentrations measured in duplicate, i.e. ten results were reported from each laboratory. Observations deviating as much as 1000% from the assigned values were seen for a single laboratory.

Considering the performance as judged by five of the six different EQAS compared, the study revealed that the schemes did not rank all laboratories identically. For laboratories, which either had a very good or a very bad
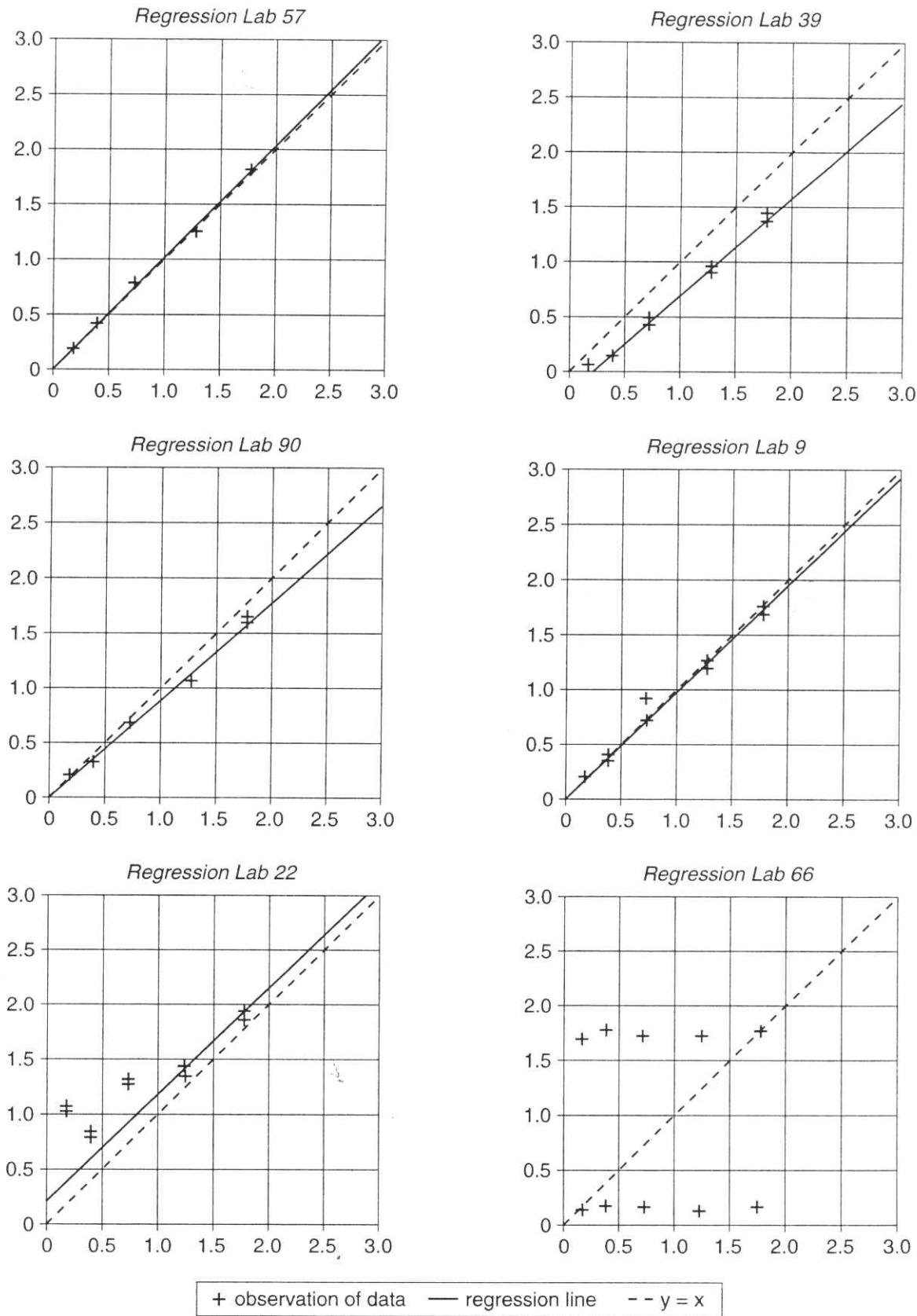
Fig. 1. - Regression lines for selected laboratories according to the ISO 5725 statistical design. Y-axes are measured values and X-axes are the assigned values for the five control samples analyzed in duplicate, concentration in µmo/l.
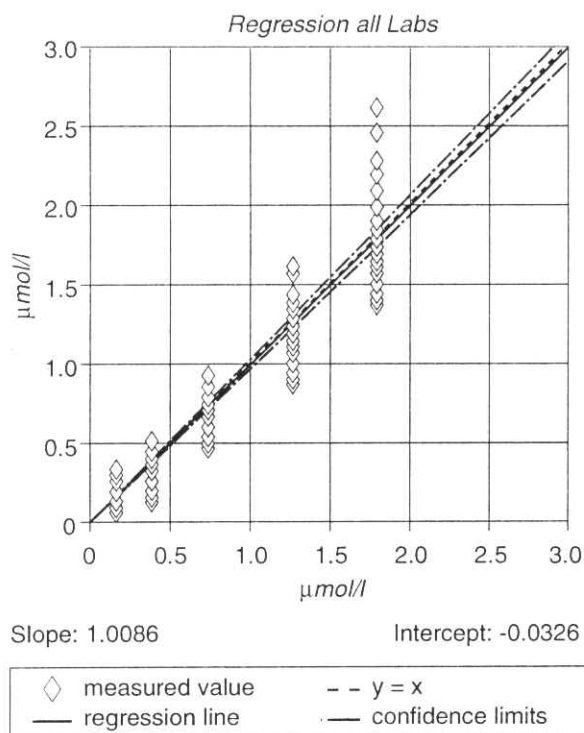
Regression all Labs

Slope: 1.0086   Intercept: -0.0326

◇ measured value	– – y = x
—— regression line	·—— confidence limits

Fig. 2. - Regression lines for all laboratories except those who reported anomalous results (outliers). Y-axis is measured values and X-axis is assigned values (DEQAS) for the five control samples analysed in duplicate, concentration in µmol/l.

performance, however, the rankings were comparable. It can be concluded that all the evaluated EQAS differ widely in statistical design and evaluation, and this poses problems and requires attention. It is important to stress that the selected statistical model and any outlier tests used in EQAS must be described in detail together with the underlying assumptions, i.e normal distribution, minimum number of results etc. A specific laboratory performance should be judged and equally classified by various EQAS. Generally applicable guidelines for selection of a statistical model cannot be provided, although a harmonized protocol for EQAS have recently been described and recommended by various international organizations (ISO, IUPAC, WELAC). In general, interlaboratory comparisons provide a mechanism to assess the analytical performance of groups of laboratories, and the obtained information allows sources of errors to be identified. As EQAS are applied as evidence for accuracy when traceability cannot be achieved, harmonized and internationally recognized statistical designs and procedures should be introduced in the future.

REFERENCES

1. CHRISTENSEN, J.M., POULSEN, O.M. & ANGLOV, T. 1994. Method evaluation, quality control and external quality assurance systems of analytical procedures. In: *Handbook on metals in analytical and clinical chemistry*. H.G. Seiler, A. Sigel & H. Sigel (Eds). Marcel Dekker, New York. p. 45-63.

2. DYBKAER, R., JORDAL, R., JØRGENSEN, P.J., HANSSON, P., HJELM, M., KAIHOLA, H-L., KALLNER, A., RUSTAD, P., ULDALL A. & DE VERDIER, C-H. 1993. A quality manual for the clinical laboratory including the elements of a quality system. *Scand. J. Clin. Lab. Invest.* **53** (Suppl. 212): 60-84.

3. INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. 1995. *Accuracy (trueness and precision) of measurement methods and results. Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement methods. Basic methods for estimating the trueness of a test method.* ISO, Geneva (ISO/IEC Standard 5725).

4. INTERNATIONAL ORGANIZATION FOR STANDARDIZA-TION. 1995. *Accuracy (trueness and precision) of measurement methods and results. Part 4: Basic methods for the determination of the trueness of a standard measurement method.* ISO, Geneva (ISO/IEC Standard 5725).

5. INTERNATIONAL ORGANIZATION FOR STANDARDIZA-TION. 1995. *Accuracy (trueness and precision) of measurement methods and results. Part 5: Alternative methods for the determination of the precision of a standard measurement method; draft.* ISO, Geneva (ISO/IEC Standard 5725).

6. BECKER, D., CHRISTENSEN, R., CURRIE, L., DIAMONDSTONE, B., EBERHARDT, K., GILLS, T., HERTZ, H., KLOUDA, G., MOODY, J., PARRIS, R., SCHAFFER, R., STEEL, E., TAYLOR, J., WATTERS, R. & ZEISLER, R. 1992. *Use of NIST standard reference materials for decisions on performance of analytical chemical methods and laboratories.* National Institute of Standards and Technology, Gaithersburg. (NIST special publication, no. 829).

7. THOMPSON, M. & WOOD, R. 1993. The international harmonized protocol for the proficiency testing of chemical analytical laboratories. *Pure & Appl. Chem.* **65**: 2123-2144.

8. WESTERN EUROPEAN LABORATORY ACCREDITATION CONFERENCE / EUROPEAN ASSOCIATION FOR ANALYTICAL CHEMISTRY. 1993. *Accreditation for chemical laboratories: Guidance on the interpretation of the EN 45000 series of standards and ISO/IEC Guide 25.* WELAC/EURACHEM, Teddington, GB.

9. SCHALLER, K.H., ANGERER, J. & LEHNERT, G. 1991. Internal and external quality control in the toxicological analysis of blood and urine samples in the Federal Republic of Germany. *Int. Arch. Occup. Environ. Health* **62**: 537-542

10. ANGLOV, T., HOLST, E. & CHRISTENSEN, J.M. 1993. Danish external quality assessment scheme: an interlaboratory comparison study on lead, cadmium and chromium in lyophilized human blood concentrate. *Int. Arch. Occup. Environ. Health* **64**: 431-438.

11. BULLOCK, D.G., SMITH, N.J. & WHITEHEAD, T.P. 1986. External quality assessment of assays of lead in blood. *Clin. Chem.* **32**: 1884-1889.

12. TAYLOR, J.K. 1988. The role of statistics in quality assurance. *Fresenius Z. Anal. Chem.* **332**: 722-725

13. CHRISTENSEN, J.M. & KRISTIANSEN, J. 1996. The Danish external quality assessment scheme. *Ann. Ist. Super. Sanità* **32**(2): 221-224.

14. HEINRICH-RAMM, R., LEHNERT, G. & ANGERER, J. 1996. The German external quality assessment scheme in occupational and environmental medicine. *Ann. Ist. Super. Sanità* **32**(2): 247-251.

15. BARUTHIO, J., ARNAUD, J. & PIERRE, F. 1996. The French interlaboratory quality assessment programme for copper, zinc and selenium in blood serum. *Ann. Ist. Super. Sanità* **32**(2): 233-239.

16. MARCUELLO, D. 1996. The Spanish external quality assessment scheme for lead in blood. *Ann. Ist. Super. Sanità* **32**(2): 277-280.

17. TAYLOR, A. & BRIGGS, R.J. 1996. The Guildford trace element quality assessment scheme. *Ann. Ist. Super. Sanità* **32**(2): 253-259.

18. WEYKAMP, C.W. & PENDERS, T.J. 1996. Trace elements in body fluids: external quality assessment scheme in The Netherlands. *Ann. Ist. Super. Sanità* **32**(2): 271-275.

19. ANGLOV, T., CHRISTENSEN, J.M. & HOLST, E. 1990. Lyophilized human whole blood for internal and external quality assurance in blood assays. *Fresenius' J. Anal. Chem.* **338**: 530-533.

20. CHRISTENSEN, S.L., ANGLOV, J.T.B., CHRISTENSEN, J.M., OLSEN, E. & POULSEN, O.M. 1993. Application of a new AMIQAS computer program for integrated quality control, method evaluation and proficiency testing. *Fresenius' J. Anal. Chem.* **345**: 343-350.