Ann. Ist. Super. Sanità, vol. 32, n. 2 (1996), pp. 295-307

Prospects of harmonisation of European external quality assessment schemes in occupational and environmental laboratory medicine

Andrew TAYLOR *(a)*, Marina PATRIARCA *(b)*, Antonio MENDITTO *(b)* and Gino MORISI *(b)*

(a) Trace Element Laboratory, Robens Institute of Health and Safety, University of Surrey, Guildford, Great Britain
(b) Laboratorio di Biochimica Clinica, Istituto Superiore di Sanità, Rome, Italy

Summary. - Since 1979, twelve external quality assessment schemes (EQAS) in occupational and environmental laboratory medicine (OELM) have been developed in nine European countries. These schemes cover a broad range of analytes and matrices and have developed differing procedures for the evaluation of laboratories' performances. Collaboration among EQAS in OELM and harmonisation of methods for evaluation of laboratory performance would be advantageous to cope with the demand for assessment of quality of analytical results for an increasing number of analytes and the need to guarantee comparability and consistency of conclusions on health and regulatory issues in different countries and at an international level. This paper compares the different features of the existing EQAS, to highlight critical points for further discussions, and to make some suggestions for possible further collaboration.

Key words: external quality assessment schemes, occupational and environmental laboratory medicine, harmonisation, Europe.

Riassunto (*Prospettive di armonizzazione degli schemi europei di valutazione esterna di qualità in medicina occupazionale ed ambientale).* - Dal 1979, dodici schemi di valutazione esterna di qualità (VEQ) in medicina occupazionale ed ambientale sono stati attivati in nove paesi europei. Questi schemi coprono un vasto numero di analiti e matrici e hanno sviluppato procedure diverse per la valutazione delle prestazioni di laboratorio. La collaborazione tra schemi di VEQ in medicina occupazionale ed ambientale e l'armonizzazione di metodi per la valutazione delle prestazioni di laboratorio sarebbe vantaggiosa per far fronte alla richiesta di valutazione della qualità dei risultati per un crescente numero di analiti e la necessità di garantire la confrontabilità e la coerenza delle conclusioni su problemi di salute pubblica in paesi diversi e a livello internazionale. In questo lavoro, vengono confrontate le diverse caratteristiche degli schemi di VEQ esistenti, mettendo in luce i punti critici da discutere, e facendo alcuni suggerimenti per una successiva collaborazione.

Parole chiave: valutazione esterna di qualità, Unione Europea, armonizzazione, medicina occupazionale ed ambientale.

Introduction

Harmonisation of methods and procedures for the assessment of the quality of analytical results is an important issue to ensure comparability of scientific conclusions and the activities of regulatory bodies at an European and international level. Laboratories seeking accreditation according to EN 45001 are required to demonstrate participation in external quality assessment schemes (EQAS) [1, 2]. In view of a possible use of EQAS results for purposes of accreditation, methods of evaluation of laboratory performance in various countries should be comparable and, wherever appropriate, harmonized. To this aim, international organizations have developed guidelines, which can be applied to the organization of EQAS [3-5]. In the field of occupational and environmental laboratory medicine (OELM), EQAS have been developed in various European countries over the course of the years [6-16]. These EQAS, which often extend beyond national borders, cover a broad range of analytes and matrices, but differ in their organization and their procedures for evaluation of laboratories' performances. Therefore, a given laboratory could be judged differently by different EQAS [17]. Organisers of schemes in OELM have to cope with complex problems, such as: the demand for EQA for an increasing number of new chemicals, which may pose a risk to human health when present in the environment; in-lab development and production of appropriate control materials; design of strategies for evaluation of performance within small groups of participants; definition of relevant standards of performance. These issues were raised at recent meetings of European EQAS organisers which were concluded with a general agreement for further collaboration [18].

Comparison of different solutions to these problems is a first step towards harmonisation and optimization of procedures. The aim of this paper is to provide such a comparison for the main features of the existing European EQAS, to highlight their differences and similarities and make some suggestions of possible strategies for improved collaboration and harmonisation.

European EQAS in OELM

General overview of the European EQAS in OELM

Table 1 reports some general features of the existing European EQAS in OELM [6-16], such as organising institutions, aim of the schemes, areas of activity, in terms of geography and number of participants, and sources of financial support.

To our knowledge, there are twelve EQAS from nine European countries operating in the field of OELM. The organisers are mainly institutes and societies of hygiene, health, occupational (and environmental) health. Two schemes are run by national institutes for quality assessment (the French scheme for lead in blood and the Dutch EQAS). The French EQAS for the determination of Cu, Se and Zn in serum is organised by the French Society for Clinical Biochemistry and is included here because of close affinities with other schemes for trace element determination. Essential trace elements are traditionally determined in laboratories which also measure toxic metals and therefore most EQAS devoted to trace elements include both categories.

The first EQAS in OELM were initiated in 1979 by the Finnish Institute of Occupational Health (FIOH, FI), for metabolites of organic substances, and the Robens Institute of Industrial and Environmental Health and Safety (RIIEHS, GB), for trace metals. Six out of twelve schemes have been active for more than ten years.

All schemes aim to improve the quality of analytical results produced in OELM. However, in two countries (Germany and France), there are legal requirements on laboratories wishing to perform analyses of given substances in body fluids to provide evidence of their technical competence by successful participation in EQAS. The French scheme for lead in blood and the German EQAS have therefore been established by law to meet these aims and participation is compulsory for any laboratory seeking permission to perform those analyses. All other European schemes have educational objectives and participation is voluntary. Besides providing laboratories with the means for self-assessment of their analytical performance, some of the EQAS have additional purposes, such as surveying laboratory performances for a given analysis (IHE, BE; ISS, IT), assessing the transferability of analytical results (SFBC, FR) and the continuous development and optimisation of procedures for EQA (ISS, IT).

All schemes are active at least at a national level and seven schemes include a large number of foreign participants. The size of the schemes (number of participants) varies according to number of analytes made available for EQA and the area of interest of the scheme, i.e. size of the country, local regulations and demand for the analysis of toxicologically relevant compounds, extension to an international level. The number of participants per scheme ranges from 10 to 172. The total number of laboratories reported to be involved in EQAS of some kind in OELM is 856, of which about 650 are located in Europe.

Financial support is from participants' fees in most schemes with contribution from the Government (IHE, BE) and the Institute funds (FIOH, FI) in two cases. Only the Italian and Spanish schemes are fully funded from the organising institutes and participation is free of charge.

Table 2 reports the analytes and matrices covered by the various schemes. Eleven are involved with trace elements analysis and only two with quality assessment for organic substances relevant to occupational and environmental exposure. As regards to trace metals analysis, the schemes include from one to 23 analyses. Due to previous European activities for the biological screening of the general population and legislation for the protection of workers [19, 20], Pb in blood is included in all schemes except the Finnish. Five schemes offer EQA for Cd in blood and four for serum Al, Cu, Se and Zn. Schemes for urinary As, Cd, Co, Cr, Cu, Hg, Ni and Zn exist in at least three countries and there are large overlaps among the elements determined in blood and urine in the British, German and Dutch schemes. As regards organic compounds, the German scheme covers a wide range of substances in blood and urine, but mainly for purposes of laboratory certification. The need for an educational EQAS in this field is partly covered by the FIOH, which offers EQA for five metabolites of organic solvents in urine.

Preparation of control samples

The procedures used in the various schemes for the preparation of control samples are outlined in Table 3. All schemes use control materials prepared in-lab by the organisers, either for each round or at the beginning of a cycle, i.e. annually or bi-annually. The methods for sample preparation for the same analytes are quite similar among different EQAS. The matrices of interest are: blood, serum or plasma, urine, dialysis fluids and water. Animal blood and serum are preferred in some schemes for economic, ethic and safety reasons, since their analytical behaviour is similar to that of analogous human matrices. Methods for blood haemolysis include sonication [12], addition of Triton X-100 [11] or deep-freezing [6, 13]; however, microclots have been reported to occur in blood samples haemolysed by addition of

Table 1. - Main features of the existing European EQAS in OELM (1995)

Country/acronym of scheme	Organiser	Started in	Aim	Extension	Participation	Support	Participants no.
Belgium (BE)/ QCB [6]	Institute of Hygiene and Epidemiology (IHE), Epidemiology Unit	1993	Education; survey of laboratory performance	National	Voluntary	Participants' fees Government funds	37 (*)
Denmark (DK) /DEQAS [7]	Danish National Institute of Occupational Health (AMI)	1990	Provide EQA	International	Voluntary		10
Finland (FI) [8]	Finnish Institute of Occupational Health (FIOH), Biomonitoring Laboratory	1979	Education	World-wide	Voluntary	Participants' fees FIOH funds	31
France (FR) [9]	Société Française de Biologie Clinique (SFBC), Group for Quality-Assurance for Trace Elements	1988	Improvement of accuracy; survey of transferability of results	European	Voluntary	Participants' fees	64
France (FR) [10]	Laboratoire de Biochimie-Toxicologie (LBT), Hôpital Jean-Bernard	1983	Self-evaluation of the quality of analytical work	World-wide	Voluntary	Participants' fees; free to foreign participants	80
France (FR) [10]	Agence du Médicament - - Unité Contrôle National de Qualité (AM-UCNQ)	1992	Ministerial approval	National	Compulsory	Participants' fees	66
Germany (DE) [11]	German Society of Occupational and Environmental Medicine (GSOEM)	1982	Certification	National/open to foreign participants	Compulsory	Participants' fees	120
Great Britain (GB)/TEQAS [12]	Robens Institute of Industrial and Environmental Health and Safety (RIIEHS), University of Surrey	1979	Education	International and national	Voluntary	Participants' fees	172
Italy (IT)/METOS [13]	Laboratorio di Biochimica Clinica, Istituto Superiore di Sanità (ISS)	1983	Development of EQA procedures; education; survey of laboratory performance	National/open to foreign participants	Voluntary	ISS funds	120
The Netherlands (NL) [14]	Foundation for Quality Assessment (SKZL), Section multi component analysis	1989	Education	National/open to foreign participants	Voluntary	Participants' fees	25
Spain (ES)/PICC-PbS [15]	Instituto Nacional de Seguridad e Hygiene en el Trabajo (INSHT), Gabinete Técnico Provincial de Zaragoza	1985	QA; assessment of performance of analytical procedures	International	Voluntary	INSHT funds	79
Spain (ES)/PICC-HgU [16]	Instituto Nacional de Seguridad e Hygiene en el Trabajo (INSHT), Gabinete Técnico Provincial de Cantabria	1986	QA; assessment of performance of analytical procedures	International	Voluntary	INSHT funds	52

(*) Largest number of participants per single analysis.

Table 2. - Analytes and matrices covered by the existing European EQAS in OELM (1995)

Organiser	Inorganics				Organics		
	Blood	Serum	Urine	Dialysis fluids/water	Blood	Serum	Urine
IHE, BE	Cd, Pb, Se						
AMI, DK	Pb						
FIOH, FI							metabolites of organic solvents
SFBC, FR		Cu, Se, Zn					
LBT, FR		Al		Al			
AM-UCNQ, FR	Pb						
GSOEM, DE	Cd, Cr, Co, Hg, Mn, Ni, Pb		Al, As, Cd, Co, Cr, Cu, F, Hg, Mn, Ni, Pb, Sb, Tl, Zn		aromatic and chlorinated hydrocarbons	organochlorine compounds, PCBs, penta-chlorophenol	metabolites of organic compounds
RIIEHS, GB	As, Cd, Hg, Mg, Mn, Pb, Zn	Al, Au, Cu, Se, Zn	As, Cd, Co, Cr, Cu, Fe, Hg, Mn, Ni, Zn	Al			
ISS, IT	Cd, Pb	Al, Cu, Se, Zn	Cr, Ni				
SKZL, NL	Cd, Co, Hg, Se, Pb	Al, Co, Cu, Li, Mg, Se, Zn	As, Cd, Co, Cu, Hg, Mg, Pb, Tl, Zn				
INSHT/PICC--PbS, ES	Pb						
INSHT/PICC--HgU, ES			Hg (Cr)				

lysing agents such as Triton-X 100 or saponine, if stored for more than few weeks [21]. Serum is used as such, but in two schemes the endogenous content of essential elements is reduced by treatment with Chelex-100 (in RIIEHS, GB, for Zn) or activated sepharose (in ISS, IT, for Zn and caeruloplasmin-Cu). Urine is deep frozen to increase sedimentation in some schemes (FIOH, FI; GSOEM, DE; ISS, IT), but used as such or after filtration in others (RIIEHS, GB; PICC-Hg, ES). Procedures for sample preservation include one or more of the following methods: deep-freezing; freeze-drying; additions of stabilisers; gamma-irradiation. In schemes where new batches of samples are prepared for each distribution, liquid control samples are sent to the participants soon after preparation.

In most cases, control samples with different concentrations are obtained by addition of the metals or organic compounds to the matrix. However, there are some exceptions, which are adopted to improve commutability between control and real samples. In the Finnish EQAS, urine samples with different concentrations of metabolites of organic solvents are preferably obtained from exposed workers. In the Italian EQAS, Cu is added to serum as caeruloplasmin and samples at different concentrations of Se are prepared by pooling, in different ratios, specimens with high and low Se content. All schemes offering more than one analyte in the same matrix use multielemental samples, with the exception of the Finnish and the Italian schemes.

Plastic vials are generally preferred for liquid samples, but glass containers are necessary for control samples to be freeze-dried, to avoid variations in the water content of the lyophilised sample, due to water absorption through plastic vials. However, whether these variations are large enough to affect analytical results is open to debate.

Organization of schemes

The general features of the organization of the schemes are reported in Table 4. Trials for EQA take place from 1 to 12 times per year. The GSOEM (DE) organises at least one EQA exercise per year for the purpose of laboratory certification. Three schemes organise EQA trials 12 times per year and three schemes six times per

Table 3. - Methods for preparation of control samples

Organiser	Matrix			Concentration levels, obtained by means of	Mono-, multi-elemental samples	Stabilizers	Physical status and storage	New batches prepared	Vials
	Blood	Serum	Urine						
IHE, BE	human			addition of metal	multi-		liquid, frozen	once a year	plastic
AMI, DK	human			addition of metal	mono-		freeze-dried		glass
FIOH, FI			human	samples from exposed subjects; addition of organic compounds	mono-		liquid, frozen	each round	plastic
SFBC, FR		bovine		addition of metals	multi-		liquid	each round	plastic
LBT, FR		human (*)		addition of metal	mono-		liquid		plastic
AM-UCNQ, FR	human			addition of metal	mono-		liquid	each round	
GSOEM, DE	animal	animal (*)	human	addition of chemicals	multi-	azide	liquid	each round	glass/plastic
RIIEHS, GB	human	bovine/ equine	human	addition of metals	multi-	γ-rays	liquid/frozen	every six-months	plastic
ISS, IT	bovine	bovine/ equine	human	addition of metals; pooling of specimens	mono- (except for Cu/Zn)	gentamycin	freeze-dried	once a year	plastic
SKZL, NL	human	human	human	addition of metals	multi-		liquid, frozen	once a year	plastic
INSHT/PICC-PbS, ES	human			addition of metals	mono-		liquid	each round	plastic
INSHT/PICC-HgU, ES			human	addition of metals	mono-	K persulphate	freeze-dried	each round	glass

(*) plasma.

Table 4. - Main features of the organisation of the existing European EQAS in OELM

Organiser	Sample distribution				Communication				
	Trials per year	Samples per trial	Samples per year	Replicates	Results returned within	Results reported by	Cumulative report	Advice	Meetings
IHE, BE	4	3	12	yes	1 month	mail	yearly	on request	
AMI, DK	2/4	5	10-20	no	2 weeks	mail	-	in the report	
FIOH, FI	4	2	8	no		mail/fax	yearly in progress	on request	
SFBC, FR	6	2	12	yes	6 weeks	mail	yearly	individual, on personal request	
LBT, FR	5/6	3	15-18	no	1 month	mail/fax	yearly		
AM-UCNQ, FR	3	3	9	no	2 weeks	mail/fax	yearly		
GSOEM, DE	≥ 1	2	≥ 2	no	5 weeks	mail	-	personal communication and contacts	
RIIEHS, GB	12	3	36	yes	1 month	mail/fax/ phone	six-month	personal visits; training; teaching	
ISS, IT	3/4	6-8	18-32	yes	1 month	mail/fax/ modem	-	on request	yes
SKZL, NL	12	1	12	yes	4 times per year	mail	yearly	in annual meetings	yes
PICC-PbS, ES	12	3	36	no	1 month	mail	-	none	
PICC-HgU, ES	6	3	18	no	1 month	mail	-	none	

year. The group of samples for EQA include from one to eight samples, with the most frequent value being three samples per trial. Unknown duplicate samples are included as part of the scheme in five EQAS. In total, participants analyze from 2 to 36 samples per year, and at least 12 samples per year in eight of the twelve EQAS considered here. In most schemes, EQA samples are sent to the participants before each trial. However, in the Dutch scheme, twelve samples are sent at the beginning of the year and one of them is analyzed each month.

Results are returned within two to six weeks, with the exception of the Dutch scheme where results are returned four times per year.

Results are generally reported by mail or fax. In the Italian scheme, dedicated software has been developed which allows participants to input their results, via modem, directly into the database and, thereafter, to have direct access to target values.

In all EQAS a report is issued to the laboratories after each trial. Cumulative reports are issued once a year in most schemes and twice a year in the scheme run by RIIEHS (GB). Annual meetings for the discussion of results are organised in The Netherlands and Italy. Some of the EQAS organisers also provide professional advice and training on request.

Evaluation of results

In all schemes, descriptive statistics are obtained at the end of each trial and the results, in numerical or graphical form, communicated to the participants. A slightly different approach is used in DEQAS (AMI, DK), where distributions of z-scores ($|y_i-\mu_i|/SD_i$) and ratios y_i/μ_i are reported instead of plain distributions of results (y_i: laboratory result for sample i; μ_i: target value; SD_i: standard deviation of the distribution of laboratories results after exclusion of outliers). In addition, EQAS organisers may provide: an evaluation of the laboratory result in comparison with target values; a more thorough evaluation of the overall laboratory performance; the laboratory score according to an agreed scoring system; a comparison with standards of performance; certification. Details of the methods used in the various EQAS are given in Table 5.

Target values for the distributed control samples are chosen either as consensus values derived from the results provided by the laboratories or assigned by the organisers (assigned values) (Table 6). Consensus values are chosen as the mean of participants' results after exclusion of outliers, with the exception of the Italian scheme which uses the median of all participants' results without any exclusion. Results are considered outliers if

Table 5. - Evaluation of results and overall performance

Organiser	Evaluation of laboratory result	Evaluation of overall performance	Standards of performance
IHE, BE	comparison with target value	graphical presentation of regression lines (of the reference laboratory and of all participants) and laboratory results (annual)	none
AMI, DK	z-score (*), ratio y_i/μ_i	linear regression of laboratory results *vs* target values (method evaluation function), relative mean square error, Youden plot (each trial)	none
FIOH, FI	comparison with target values	comparison of laboratory results with target zones - annual - in progress	none
SFBC, FR	comparison with target values	average annual score based on target zones for proximity, recovery and between run precision expressed in percentage of the maximum possible score (100% if all results are within the inner limits of acceptability)	Annual global score: (max. score 100) good laboratories: > 70% acceptable laboratories: 50-70% inadequate laboratories: < 50%
LBT, FR	comparison with target values and recovery	average annual score based on target zone for proximity to target value and recovery	excellent score: > 100 (max. score 200)
AM-UCNQ, FR	comparison with target values and recovery	average annual score based on target zone for proximity to target value and recovery	good performers score: >100 max. score 200
GSOEM, DE	comparison with target values	Youden plot/comparison to tolerance intervals (3 x SD of the reference laboratories)	both results within tolerance limits
RIIEHS, GB	scores based on target zones	monthly cumulative score based on target zones for proximity, recovery and between run precision	in progress
ISS, IT	acceptable results defined by comparison with target zone, z-score	linear regression of laboratory results *vs* target values (per trial); % of acceptable results (per year)	% of acceptable results: good performers: > 80% acceptable performers: 79-50% poor performers: < 50%
SKZL, NL	comparison with target value	annual: linear regression *vs* added amounts; precision; recovery	none
PICC-PbS, ES	index of variance	mean of IVs (IVM)	IVM: good performers: < 60 acceptable performers: 60-100 poor performers: >100
PICC-HgU, ES	index of variance	mean of IVs (IVM)	IVM: good performers: < 50 acceptable performers: 50-90 poor performers: > 90

(*) Standard deviation obtained from laboratory results after outlier exclusion.

they lay outside the interval mean ± n SD, where n is 2 or 3 (Table 6). Outliers are excluded according to the Cochran and Grubbs tests in the Danish scheme. In some schemes (AMI, DK; ISS, IT), as an alternative procedure, the values determined by the organisers are used as target values, when the number of participants is low or most of the participants have little experience of the given analysis. Assigned values are determined by different methods in the various schemes, i.e.: estimated from the regression line between the results of an independent reference laboratory and the spiked amounts (IHE, BE); calculated as the sum of the endogenous concentration of the analyte, determined by the organisers, and the spiked amount (FIOH, FI); chosen as the mean of the values reported by a group of independent German and European reference laboratories after exclusion of outliers (data exceeding the 95% range) (GSOEM, DE); chosen as the mean of the values obtained by selected participants, i.e. laboratories who showed high performance in the last three exercises (PICC-HgU, ES).

Table 6. - Target values and methods for outliers detection

Organiser	Target value	Outliers
IHE, BE	Assigned values [a]	eye inspection t-test Dixon's tests
AMI, DK	Consensus mean [b] Assigned value [c]	Cochran and Grubbs tests
FIOH, FI	Consensus mean [b] (samples from exposed subjects)/Assigned value [d]	± 3 SD
SFBC, FR	Consensus mean [b]	± 3 SD
LBT, FR	Consensus mean [b]	± 2 SD
AM-UCNQ, FR	Consensus mean [b]	± 2 SD
GSOEM, DE	Assigned value [e]	
RIIEHS, GB	Consensus mean [b]	± 3 SD
ISS, IT	Median/Assigned values [f]	
SKZL, NL	Consensus mean [b]	± 3 SD
PICC-PbS, ES	Consensus mean [b]	± 2 SD
PICC-HgU, ES	Assigned value [g]	

(a) estimated from the regression line between the results of an independent reference laboratory and the spiked amounts; *(b)* mean of participants results after exclusion of outliers; *(c)* determined by the organisers when < 10 labs participate; *(d)* endogenous concentration determined by the organisers and added to spiked amount; *(e)* determined by the values reported by reference laboratories after exclusion of outliers (data exceeding the 95% range); *(f)* determined by the organisers when the number of participants is low or there is little experience; *(g)* mean of values obtained by selected participants (laboratories with high performance).

Participants can assess their own performance by comparing their results with the target values and the dispersion of the results of all other participants.

According to the scheme design, the organisers can provide additional information on the overall performance of the laboratory, either at the end of each trial or of a series of trials (six-months, one year). For example, reproducibility can be estimated by the results obtained on unknown duplicate samples (IHE, BE; SBFC, FR; RIIEHS, GB; ISS, IT; SKZL, NL) and the presence of systematic errors can be unveiled by the evaluation of recovery of the added amounts (SFBC, FR; LBT, FR; AM-UCNQ, FR; RIIEHS, GB; SKZL, NL) or linear regression of laboratory results (y) against the target values or the spiked amounts (x) (AMI, DK; ISS, IT; SKZL, NL).

The international harmonised protocol for proficiency testing of analytical laboratories [5] recommends that individual results should be classified according to their deviation from the target value and a performance score derived. Among the existing European EQAS in OELM, some (three) do not provide at present any judgement on the laboratory results, although they may do so in future. In the German scheme, scores are not assigned as such, but certification is awarded only to laboratories who obtained both results within assigned tolerance limits. In the Danish scheme, laboratory performance is evaluated by means of the software package AMIQAS [22] which uses a statistical procedure based on linear regression between laboratory results and target values (method evaluation function) and laboratories are ranked according to their $RMSE^{1/2}$, a quantity which takes into account both random and systematic errors [7, 23].

Among the other schemes, indices of the overall performance of the laboratory are derived from the comparison of the laboratory results with acceptable ranges and combined to give a cumulative score per: each trial, a definite number of trials, a period of time, the time since the laboratory joined the scheme (Table 5).

Acceptable deviations from target values are chosen as 15% in the Spanish schemes, whereas in the other schemes, acceptable deviations, for each analysis, depend

on the level of concentrations and are defined in a graphic way (FIOH, FI; SFBC, FR; LBT, FR; AM-UCNQ, FR; RIIEHS, GB; ISS, IT) to allow for larger per cent errors at lower concentrations (Figs 1, 2). Three schemes (FIOH, FI; SFBC, FR; RIIEHS, GB) include inner limits to allow the distinction between "good" and "acceptable" results.

Scores are calculated for each result as z-score multiplied by 100 (variance index, IV) in the Spanish schemes. The mean of the previous IVs (IVM) is taken as an index of the overall performance of the laboratory and updated at each trial.

Another type of score, defined in detail elsewhere [10], is calculated in the French schemes for Pb in blood and Al in serum and dialysis water, for both the deviation from target concentration values and the deviation from expected recovery of the added amounts. At the end of one year, laboratories are classified according to the sum of their scores for the two criteria, obtained as the mean of all individual scores.

In the Italian scheme, results are individually judged as acceptable or unacceptable, according to their compliance with acceptable deviation from target values. The overall performance of the laboratories over periods of about one year is described in terms of percentage of acceptable results obtained. However, as an additional measure, z-scores are also computed for each result. The value chosen for the target deviation depends on the concentration and is derived from the graphical criterion, to maintain correspondence between what is judged unacceptable in both cases.

The schemes organised by the RIIEHS (GB) and the SFBC (FR) adopt the same scoring system and assign scores of 2, 1 and 0 to results within inner limits, outer limits or beyond the outer limits of acceptable deviations, respectively. Cumulative scores for each trial are calculated adding up individual scores. At the end of a six trial-cycle, cumulative scores are computed for the proximity to the target value, recovery and precision on replicate specimens sent in different occasions.

Standards of performance

Standards of performance, i.e. the level of performance which should be achieved by a competent laboratory and the level of performance that would be judged unacceptable, are defined in some of the schemes. The introduction of standards of performance is reckoned useful because they stimulate laboratory improvement by setting goals and stating clearly what is expected to be achieved; establish limits beyond which intervention of advisory panels of professionals should be sought; establish limits beyond which a limitation of the laboratory activity may be required.

Since the ultimate goal is to achieve comparable judgement of the performance of OELM laboratories within Europe, the establishment of harmonised standards of performance has to be considered when discussing harmonisation of procedures for EQAS.

At present, eight EQAS include in their protocols the definition of standards of performance. The simplest method is used by the German scheme, where evaluation of laboratory performance consists of a Youden plot, comparing the results of the laboratory on two samples with assigned tolerance intervals. Certification is achieved if both results fall within the assigned tolerance intervals. In other six schemes (SFBC, FR; LBT, FR; AM-UCNQ, FR; RIIEHS, GB; PICC-PbS, ES; PICC-HgU, ES) standards of performance are set for the global performance score to be achieved by a laboratory. In the Italian scheme, competent laboratories are expected to obtain acceptable results for at least 80% of the control samples examined in one year.

Comments and suggestions

In the field of OELM, a high level of collaboration could be developed among EQAS, since the activities of the existing schemes already extend beyond national borders in most cases. Therefore, a fertile ground already exists for the development of supernational initiatives, which would be best developed within a network of EQAS organisers, supported by the European Commission. The need for assessment of analytical results in this field is enhanced by the increasing demand for measurements of indices of biological exposures to a variety of chemicals on one hand, whereas, on the other hand, the availability of a new generation of computer-controlled instruments, which are easier to set up and to use, makes it possible for a larger number of less specialised laboratories to undertake these measurements.

Collaboration among EQAS should aim to provide consistency of assessment of analytical performance for all relevant determinands and for all European laboratories. However, schemes cannot cover absolutely every measurement and should be regarded as providing representative assessments. Preference should be given to those analytes for which determination is required by European legislation [24], to tests carried out more frequently according to national surveys and taking into account local needs. Specific problems could be dealt with and experience shared. Pilot schemes could be set up at a European level for new analytes, for less common analytes (i.e. analytes which are determined by less than 10 laboratories in each country), and for analyses related to a specific problem or emergency.

An harmonised protocol for EQAS in OELM need to be agreed among EQAS organisers. This should define agreed minimum requirements for the conduct of EQAS, which could be adopted in all schemes, in addition to the present organization, and which could be developed further with increasing experience. Work in this direction was initiated by international organizations [5], whose recommendations are synthesised in Table 7.

Table 7. - Minimal requirements for proficiency testing schemes according to the international harmonised protocol for the proficiency testing of analytical laboratories [5]

Structure	Preparation and evaluation of test material (coordinator); sample distribution (according to regular schedule); participants analyse samples and report results within a given time; results subjected to statistical treatment by the coordinator; participants promptly notified of their performance; advice available to poor performers; participants identified in reports by code only.
Organisation	All practices and procedures to be documented in a quality manual; overall direction of the scheme overseen by a small advisory panel having representatives (i.e. practising laboratory scientists) from coordinator, contractors, professional bodies, participants and end-users of analytical data.
Test material	Similar to the material to be analysed (matrix and concentration); homogeneous - recommended between samples standard deviation should be less than 0.3 times the target value for the standard deviation; stable under transport conditions and laboratory environment; ideally, the quality checks on the test materials should be performed by a different laboratory from that which prepared the sample; target value unknown to the participants; precautions to be taken for potential hazards of the test material involved.
Assigned values	Consensus value from expert laboratories; formulation; direct comparison with certified reference materials; consensus of participants.
Test materials per round and analyte	Upper limits of six.
Duplication in the test	Possible but not required.
Frequency of rounds	Min: once every two weeks; Max: once every four months.
Report	Distribution of results from all laboratories, e.g. histograms; participant performance score; test results; sent before next group of samples.
Evaluation of results	Bias estimate; target value for standard deviation: by perception, prescription, reference to validated methodology, reference to a generalised model; z-score and its interpretation and classification: $\|z\| \leq 2$ satisfactory; $2 < \|z\| < 3$ questionable; $\|z\| > 3$ unsatisfactory; q-score, as an alternative to z-score; combination of scores within one round or a trial: sum of scores, sum of squared scores, sum of absolute values of the scores; ranking; running score.
Liaison with participants	Periodic and open meetings; communication via a newsletter or annual report; advice available to poor performers; possibility of referring to the coordinator by participants to consider that their performance assessment is in error.

According to these recommendations [5], all types of proficiency testing schemes "share the common feature of the comparison of test results obtained by one testing laboratory with those obtained by one or more other testing laboratories". All existing EQAS in OELM share the structure described in Table 7. Initiatives to provide advice to poor performers are available in most schemes. EQAS organisers agreed to implement a quality manual by 1999 [18]. The organization of EQAS at a European level should be overseen by an advisory panel including coordinators, contractors, professional bodies, participants and end-users of data.

Preparation of control samples is one of the tasks of EQAS organisers and there is fairly good agreement among procedures for the preparation of samples for OELM schemes. Exchanges of experiences on this issue are important and collaborative projects for the development of new control materials could be organised. The harmonised protocol indicates consensus values from expert laboratories as the best method to assign target values to the test materials. For practical reasons most schemes use the consensus value of participants' results as the target value. However, EQAS organisers could act together as a selected group of reference laboratories for the examination of control samples used in similar schemes and the assignment of target values. Frequency and number of samples distributed to participants depend on the aims of the scheme and the procedures of evaluation of results, and there is no imperative for this aspect to be harmonised.

Procedures for the evaluation of results pose the biggest problems to harmonisation, as all procedures are well established and valid. However, in all EQAS where they have established, scoring systems are either based on z-scores or can be transformed in z-scores, by means of relatively simple arithmetics. Harmonisation of the

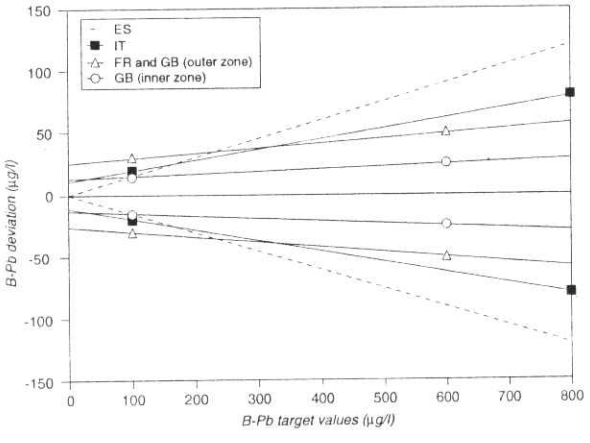

Fig. 1.- Examples of acceptable limits over a range of concentration adopted in the existing European EQAS for lead in blood. Full lines: acceptable limits (%) vary with concentration. Dotted line: acceptable limits (%) do not vary with concentration.

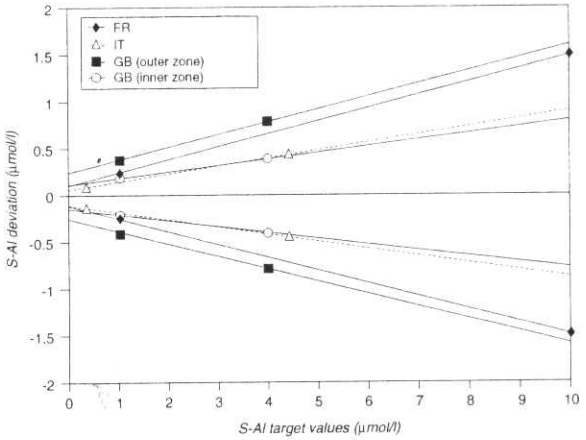

Fig. 2.- Examples of acceptable limits over a range of concentration adopted in the existing European EQAS for aluminium in serum.

"target standard deviation" and the "acceptability ranges" used by different schemes is all that is required to reach comparable scores. Figures 1 and 2 show the acceptable deviation in different schemes for blood lead and serum aluminium, respectively. The numerical values of the target standard deviation are determined by the organisers according to different methods (Table 7). It would be counterproductive to set the target standard deviation too small as the final objective is not to achieve a quality of data unnecessarily high for their application, but to satisfy stated or implied needs of the end-users. Choices need to be made and considerations on efficacy of the schemes and cost/benefit ratio will play a major part. In the existing EQAS, target standard deviations were chosen by the organisers according to their best estimate of the clinical needs and present analytical capabilities for the given analysis. Therefore large differences may arise from different local experiences of the organisers, e.g. overestimate of what is achievable in terms of interlaboratory dispersion due to limited experience of the group of laboratories considered. With improvement of technology and knowledge of biochemical effects of exposure, target values for standard deviation may need to be revised. Acceptability limits for results or expected standard deviation could be agreed at a European/ international level for analytes of major concern in environmental and occupational medicine, e.g. those for which biological exposure indices or similar values have been established.

Standards of laboratory performance have been established in most schemes in terms of scores to be achieved by a competent laboratory and it has been shown that their imposition can stimulate dramatic improvement of laboratory performance [21]. This feature should become part of all schemes and harmonised for the same analyses among different schemes.

In terms of evaluation of overall laboratory performance a number of schemes use linear regression or at least an estimate of recovery of added quantities. This approach, which is purely educational, provides essential information to the laboratory to identify the sources of error. However, this is not a minimal requirement for the assessment of laboratory performance. Inclusion of this type of procedure should be decided according to whether the main aim of the scheme is certification or education of laboratories.

The report returned to the participants should be provided within a short time and on-line transmission of data, to allow real time assessment of performance, is probably the ideal solution. Inclusion of scores provides an easy means to compare performance while standards of performance set goals to be achieved.

The ultimate objective of an EQAS is the improvement of laboratory performance. To achieve this objective, some further important issues have to be considered and, in addition to harmonisation of existing schemes, a network of EQAS organisers could undertake complementary activities.

Collaborative studies could be implemented for analyte(s) which are either particularly important or for which specific EU legislation exists [24]. A protocol, in agreement with the international harmonised protocol [5] should be agreed for sample preparation (choice of matrix, formulation, homogeneity testing, target value assessment, stability), sample distribution (number of samples, replicate samples, procedures to avoid identification of samples) and methods of evaluation of laboratory performance. The procedures tested in the pilot studies would be transferable to other analytes/ matrices.

Collaborative studies for the comparison of already existing EQAS could be envisaged.

To avoid the problems highlighted by the collaborative project to compare EQA schemes on blood lead [17], EQAS organisers should first agree on classification of laboratories within each scheme according to level of performance (e.g. good, acceptable, poor). Each scheme could use its own criteria to assign laboratories to classes of performance, but comparing classes of performance could be easier and more meaningful than comparisons of ranks.

Concerted training and education activities could also be undertaken, including:

- exchange of information among members on analytical methods and performance for new analytes; new analytical techniques; information to be spread to national laboratories;

- creation of an European database of laboratories and activities in OELM;

- agreement on a list of common themes to be developed in national courses for laboratory operators (i.e. analytical techniques, quality assurance, laboratory certification and accreditation). In same cases, exchange of speakers could be arranged;

- work in the area of distance education, training and information exchange.

Submitted on invitation.
Accepted on 15 March 1996.

REFERENCES

1. EUROPEAN COMMITTEE FOR STANDARDIZATION / EUROPEAN COMMISSION FOR ELECTROTECHNICAL STANDARDIZATION. 1991. *General criteria for the operation of testing laboratories.* 2nd ed. CEN/CENELEC, Bruxelles. (EN 45001).

2. 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.

3. INTERNATIONAL ORGANISATION FOR STANDARDISATION / INTERNATIONAL ELECTROTECHNICAL COMMISSION. 1994. *ISO/IEC Guide 43. Development and operation of laboratory proficiency testing.* ISO, Geneva.

4. INTERNATIONAL UNION OF PURE AND APPLIED CHEMISTRY. 1988. Protocol for the design, conduct and interpretation of collaborative studies. *Pure Appl. Chem.* **60**: 856-864.

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

6. DE PLAEN, P., CLAEYS-THOREAU, F. & DE BOECK, R. 1996. External quality assessment scheme in Belgium. Heavy metal analysis. *Ann. Ist. Super. Sanità* **32**(2): 215-220.

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

8. VALKONEN, S., ENGSTRÖM, K., AHONEN, I., MUTANEN, P. & AITIO, A. 1996. FIOH external quality assurance scheme for organic solvent metabolites. *Ann. Ist. Super. Sanità* **32**(2): 225-231.

9. BARUTHIO, F., 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.

10. GUILLARD, O. PIRIOU, A. & PINEAU, A. 1996. French external quality assessment schemes for lead in blood and aluminium in plasma and dialysis water. *Ann. Ist. Super. Sanità* **32**(2): 241-245.

11. 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.

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

13. MENDITTO, A., PATRIARCA, M., CHIODO, F. & MORISI, G. 1996. The Italian external quality assurance scheme for trace element analysis in body fluids. *Ann. Ist. Super. Sanità* **32**(2): 261-270.

14. 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.

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

16. QUINTANA, M.J. & MAZARRASA, O. 1996. The Spanish external quality assessment scheme for mercury in urine. *Ann. Ist. Super. Sanità* **32**(2): 281-284.

17. CHRISTENSEN, J.M. & OLSEN, E. 1996. Comparison of some European external quality assessment schemes in the field of occupational and environmental medicine. *Ann. Ist. Super. Sanità* **32**(2): 285-294.

18. PATRIARCA, M., MENDITTO, A. & TAYLOR, A. 1996. First steps toward harmonization of European EQAS in occupational and environmental laboratory medicine: from Dublin to Rome. *Ann. Ist. Super. Sanità* **32**(2): 193-198.

19. EUROPEAN ECONOMIC COMMUNITIES. 1977. Council Directive of 29 March 1977 on biological screening of the general population for lead. *Off. J. Eur. Comm.* (L 105 of 28/4/77, p. 10-17).

20. EUROPEAN ECONOMIC COMMUNITIES. 1982. Council Directive of 28 July 1982 on the protection of workers from the risk related to exposure to metallic lead and its ionic compounds at work (First individual Directive within the meaning of Article 8 of Directive 80/1107/EEC). *Off. J. Eur. Comm.* (L 247 of 28/7/82, p. 12-17).

21. TAYLOR, A. & BRIGGS, R.J. 1995. External quality assessment schemes and improvements in the measurement of trace elements in biological fluids. *J. Anal. Atom. Spectrom.* **10**: 1033-1037.

22. 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.

23. 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.

24. EUROPEAN ECONOMIC COMMUNITIES. 1994. Amended proposal for a Council Directive on the protection of the health and safety of workers from the risk related to chemical agents at work. *Off. J. Eur. Comm.* (C 191 of 14/7/94, p. 7-29).