Ann. Ist. Super. Sanità, vol. 32, n. 2 (1996), pp. 221-224

The Danish external quality assessment scheme

Jytte Molin CHRISTENSEN and Jesper KRISTIANSEN

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

Summary. - Participation in proficiency testing schemes is generally recognized as an important part of quality control in analytical chemistry. The design and conduct of proficiency testing schemes is therefore a matter of great importance. A proficiency testing scheme should: a) cover a broad concentration range of the substance of interest; b) enable the detection and estimation of each laboratory's systematic errors; c) enable estimation of each laboratory's random error; d) provide each laboratory with a rank or score, and e) present the evaluation of each round to the participating laboratories in a comprehensible way, e.g. in a report with carefully designed text, tables and figures. This paper presents the design of the Danish external quality assessment scheme (DEQAS), which is based on the method evaluation function concept.

Key words: quality assessment, proficiency testing, method validation.

Riassunto *(Lo schema danese per la valutazione esterna di qualità)*. - La partecipazione in programmi di valutazione esterna di qualità è generalmente riconosciuta come una parte importante del controllo della qualità in chimica analitica. La pianificazione e l'organizzazione di tali programmi è quindi materia di grande importanza. Uno schema per la valutazione esterna della qualità dovrebbe: a) coprire un largo intervallo di concentrazioni della sostanza di interesse; b) consentire l'identificazione e la stima degli errori sistematici in ciascun laboratorio; c) consentire la stima dell'errore casuale di ciascun laboratorio; d) assegnare un punteggio a ciascun laboratorio; e) presentare i risultati di ciascun esercizio di valutazione esterna di qualità ai laboratori partecipanti in modo comprensibile, per esempio includendo nel rapporto testo, tabelle e figure appositamente studiate. Questo articolo descrive la struttura del programma danese per la valutazione esterna di qualità (DEQAS), che è basato sul concetto della "method evaluation function".

Parole chiave: controllo di qualità, valutazione delle prestazioni analitiche, validazione di metodi analitici.

Introduction

The Danish external quality assessment scheme (DEQAS), initiated in 1990, is developed and managed by the National Institute of Occupational Health (Arbejdsmiljøinstituttet, AMI) in Denmark. It is designed to evaluate the quality of results of measurements of a group of laboratories. The scheme, which evaluates performance at five different concentrations of the analyte, can be used universally to evaluate laboratories in the fields of clinical, biological, environmental or occupational chemistry. The scheme may also be used in validating other variable parameters such as the performance of technicians, efficiency of apparatus, etc.

Currently, DEQAS comprises proficiency testing schemes for lead in blood, iron, manganese and titanium in welding fume, and organic solvents on charcoal tubes [1, 2]. However, this paper will focus mainly on the EQAS for lead in blood.

Procedure of DEQAS

Blood samples are prepared from fresh human blood obtained from a blood bank. All blood used in the scheme has been tested and found negative for antibodies against HIV and hepatitis B and C. EDTA has been added as anticoagulant. The blood is filtered, spiked with an aqueous solution of lead nitrate, divided into vials and freeze-dried. Lyophilization vials and stoppers have been checked for lead contamination.

Five samples for analysis are dispatched in each round, and two to four EQAS rounds are performed every year. Together with the samples a laboratory report sheet is sent, which must be returned to AMI with the laboratory's results within two weeks of receiving the samples. To check that the registered results are identical with the data actually submitted, a control sheet, with the received and registered data, is mailed to the participating laboratory. Within two weeks of

receiving the control sheet the evaluation report produced by the AMIQAS computer programme is sent to the laboratories.

Laboratories are enrolled on a voluntary basis. The fee includes EQAS samples, statistical evaluation of results and reports.

Evaluation of results

Tests for outliers and normal distribution

To exclude extreme outliers a tolerance interval (TI) is calculated for each of the five concentrations using a relative standard deviation appointed to be 15% of the true value. TI is defined as follows:

$$TI = \bar{x} \pm (a \cdot \bar{x})\ 1.96\ (1 + 1/n^{1/2})$$

where a = 15 and n is the number of participating laboratories. The TI is calculated for 95% of the distribution with 95% confidence [1]. Thus there will be a probability of 95% that the TI will include 95% of the distribution. As the proficiency testing scheme is based on the assumption, that the results are normally distributed, Kolmogorov-Smirnov test for goodness of fit to the normal distribution is carried out and Cochran and Grubbs outlier tests are used to identify and exclude outliers [3].

True concentrations

In DEQAS the target values are established either as: a) consensus mean values, which are the mean values of the results of all participating laboratories after outlier exclusion, or b) reference values, which are values obtained by reference laboratories. The latter procedure is used when only a few (< 10) laboratories participate in the scheme.

Distribution plots

For each sample, the ratio between the obtained result Y_i and the target value μ_i is computed. The distribution of the results of the participating laboratories is visualized by plotting the (Y_i/μ_i)-ratios for the five samples (Fig. 1). All ratios within the interval (0.6 - 1.4) are shown, and the number of laboratories with ratio outside this interval is indicated.

Further, the z-value is computed for each laboratory at a selected true concentration, and the ranked z-values from all participants are plotted in increasing order (Fig. 2). The z-value is defined as [4]:

$$z = (Y_i - \mu_i)/SD_i,$$

where Y_i is the obtained result at the concentration level i; μ_i is the estimated consensus value or reference concentration at the concentration level i, and SD_i is the

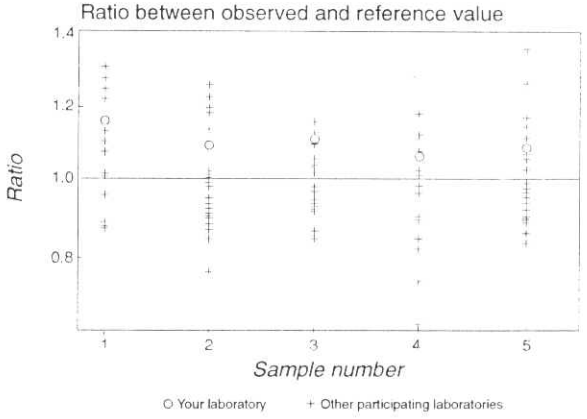

Fig 1. - The ratio plot between observed measurement results and reference values.

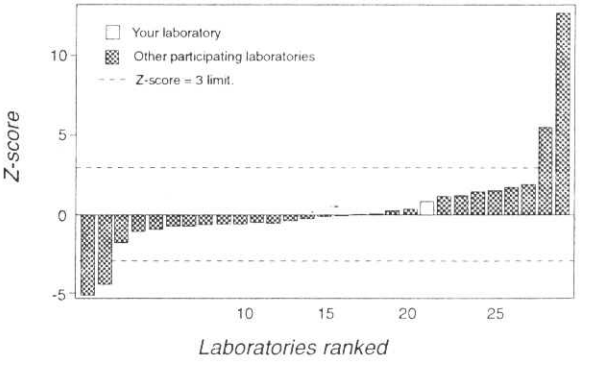

Fig. 2. - A ranked z-score plot of all laboratory results plotted in increasing order.

standard deviation at this concentration after outlier exclusion. If μ_i and SD_i are good estimates of the true concentration and standard deviation the z-values are approximately normal distributed with a zero mean and the unit standard deviation. In this case the evaluation of the laboratories are based on the z-score:

- $|z| < 2$ would be very common.
- $|z| > 3$ would be very rare in well-behaving systems.

Hence, values of $|z| > 3$ may with great confidence be considered due to poor performance. In the evaluation report the values of $|z| > 3$ are not included in the z-score plot, but the number of laboratories with $|z| > 3$ is indicated [5].

Random and systematic errors

Estimates of the experimental standard deviation σ_i of the method used by each laboratory, the slope β and intercept α of the method evaluation function (MEF), is obtained by statistical evaluation (least square regression analysis) of the linear relationship between the results obtained by the laboratory and the target values [1, 5, 6].

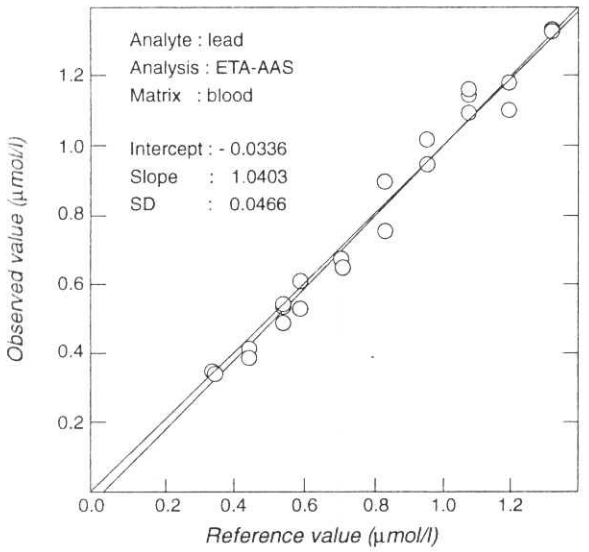

Fig. 3. - An example of the estimated method evaluation function (MEF) in DEQAS.

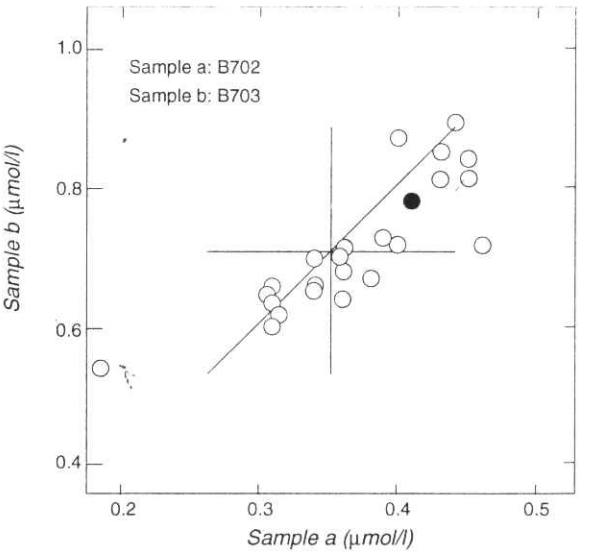

Fig. 4. - An example of a two-sample Youden plot.

shown in Fig. 3. In the MEF-plot the 0.95 confidence interval of the ideal MEF is indicated using σ_i-values assigned by AMI (0.048 µmol/l for $\mu_i \leq 0.48$ µmol/l, and 10% of μ_i for $\mu_i > 0.48$ µmol/l. The estimated MEF is accepted if the function values are within the 0.95 confidence intervals of the ideal MEF at all five true concentrations. As the number of samples in each round of proficiency testing is too low to allow a reliable method evaluation the estimated values of σ_i, α and β may only be used as an indication about the size of random and systematic errors.

The EQAS evaluation also contains an evaluation of a Youden-plot (Fig. 4) to analyze random and systematic errors of each laboratory. The target values of two samples are represented by the vertical and horizontal lines drawn throughout the point (μ_X, μ_Y), thereby dividing the chart into four quadrants. If the variation in results from different laboratories is due to random errors only, the points will be scattered with approximately equal numbers in each of the four quadrants. If a method produces results which are systematically to high, a predominance of points in the top right quadrant is expected. Similarly, if the results are systematically too low, a predominance of points in the low left quadrant is expected.

Evaluation of performance

In the DEQAS evaluation reports, the values of the square root of the relative mean square error ($RMSE^{1/2}$) are given for all participants at all true concentrations. The $RMSE^{1/2}$ is a measure of the sum of the laboratory's systematic and random errors, and it shows how exactly the laboratory in general measures the estimated true value [1, 6]. The smaller the $RMSE^{1/2}$ value, the closer the measurement is to the estimated true value. $RMSE^{1/2}$ is defined as:

$$RMSE^{1/2} = \sqrt{\left(\frac{\delta}{\mu}\right)^2 + \left(\frac{\sigma}{\mu}\right)^2}$$

where δ is the systematic error, σ is the standard deviation and μ is the conventional true value. The ideal $RMSE^{1/2}$ to a given concentration level i is calculated using the appointed standard deviation (σ_{app}) and a zero systematic error ($\delta = 0$) [1]. A robust tolerance interval is computed by assuming normal distribution, $TI = \mu_i \pm 1.96\ \sigma_{app}$.

$RMSE^{1/2}$ values are calculated at different levels and used in the evaluation. The use of $RMSE^{1/2}$ as control parameter in interlaboratory quality control has two major advantages: a) $RMSE^{1/2}$ is well defined; b) $RMSE^{1/2}$ includes the laboratory's systematic (δ) and random error (σ). Hence, $RMSE^{1/2}$ evaluation provides sufficient information of the method.

Conclusions

In 1994 19 laboratories participated in the DEQAS. In the design of the DEQAS weight has been attached to voluntary participation and the educational role of the scheme. The scheme is designed to be in accordance with international standards and guidelines in the field. Reports are designed to graphically present the laboratory performance, e.g. random, systematic and total error illustrated by the MEF (Fig. 3) and $RMSE^{1/2}$ plots (the

$RMSE^{1/2}$ plot is not shown), and to enable comparison with the performance of other laboratories, e.g. illustrated by the ratio (Fig. 1), z-score (Fig. 2) and Youden plots (Fig. 4).

Submitted on invitation.
Accepted on 5 September 1995.

REFERENCES

1. 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.

2. ANGLOV, J.T.B., HOLST, E., DYG, S. & CHRISTENSEN, J.M. 1993. Iron, manganese, copper and titanium in welding fume dust on filters for internal and external quality assurance. *Fresenius' J. Anal. Chem.* **345**: 335-339.

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

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

5. 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.

6. CHRISTENSEN, J.M., POULSEN, O.M. & ANGLOV, T. 1992. Protocol for the design and interpretation of method evaluation in atomic absorption spectrometric analysis. Application to the determination of lead and manganese in blood. *J. Anal. At. Spectrom.* **7**: 329-334.

Appendix. - Summary of the scheme

Country	Denmark.
Name of scheme	Danish external quality assessment scheme (DEQAS).
Status of scheme	International. Voluntary. Run by the National Institute of Occupational Health, Denmark. *Aims:* to provide external quality control for laboratories in the occupational health sector in Denmark. *Participants:* varying number (19 in 1994). Analytical chemical laboratories.
Scheme description	*Control materials:* usually prepared "in-house". Samples enriched by "spiking". Blood samples are freeze-dried. Target values usually determined as consensus mean values or by one or more reference laboratories if only a few laboratories participate. These two methods of establishing the target values are optional in the PC-programme that makes the statistical elaboration. *Organization of EQA exercises*: frequency of exercises: 2 per year; number of samples: 5 at different concentration levels; time schedule for returning the results: within 2 weeks. *Elaboration of results:* by means of a PC-programme. Reports and control sheets are generated automatically, and returned to the laboratories within 2 weeks. *Criteria for evaluation of laboratory performance:* Z-score plot, Youden plot, ratio plot, MEP plot and $RMSE^{1/2}$ plot. *Measures taken against poor performers:* advice is given in the report. No other measures.
Organization	National Institute of Occupational Health att. Kirsten Byrialsen Lersø Parkallé 105 DK-2100 Copenhagen, Denmark Tel + 45 39 29 97 11. Fax + 45 39 27 01 07 e-mail: amilpost@uts.uni.c.dk
Analytes and matrices covered	Lead in human blood. Iron, manganese and titanium in welding fume. Organic solvents on charcoal tubes.