Upsala J Med Sci 98: 259–274, 1993

# 5.2 Transferability of Clinical Laboratory Data

Torgny Groth[1] and Carl-Henric de Verdier[2]

[1]*Unit of Biomedical Systems Analysis and* [2]*Department of Clinical Chemistry, University of Uppsala, Uppsala, Sweden*

## INTRODUCTION

Transferability of clinical laboratory measurement results is an important problem, both within and between laboratories and hospitals (2,4,5,6). For instance, it cannot be taken for granted that measurement results can be transferred over time in longitudinal statistical studies, in the use of reference values, and in the monitoring of patients. Well-designed statistical control procedures are required which can assure the specified analytical quality. Careful documentation is required of changes performed in measurement procedures and which affect the analytical results.

A very common problem of transferability within a laboratory is related to the reporting of measurement results from a group of instruments measuring the same components. Even with instruments of the same type and from the same manufacturer, and with identical reference materials for calibration it may well be inter-instrument differences that jeopardize the analytical quality goals.

However, if the instruments show stable performance within defined limits it would be possible to "tune" the instruments on the basis of *correction functions* estimated from simultaneous measurements on reference material (*e.g.* selected patient samples) covering the whole measurement interval of interest ('analytical bias assessment programme').

Transferability problems are more commonly realised in connection with communication of laboratory results between laboratories, hospitals and health care centres. In the first place in the care of individual patients referred from another hospital, but also in connection with multi-centre collection of reference values, and in using reference limits and decision limits originating from outside the own laboratory. More recently, the transferability problem has also been recognised in the application of decision rules and computer-based interpretative programmes in places not directly involved in the development of the knowledge-bases.

In order to decrease the analytical inter-laboratory variation, improvement is generally required in various ways (5), *e.g.* with regard to

- *Analytical methodology* , to get rid of unspecific methods;
- *Calibration procedures*, to reduce analytical between-run variation and drift;
- *Internal analytical quality control procedures*, to achieve stability in analytical performance within defined limits of allowable analytical errors;
- *External quality assessment (EQA) programmes*, to obtain a reliable characterization of the individual laboratory with regard to analytical bias ("trueness problems") (for terminology see Chapter 11);
- *Numerical correction* of known systematic deviations from conventional true values (analytical bias).

All these activities also presuppose further development and use of *definitive or reference method technology,* and well designed *reference materials* for calibration, analytical quality assessment and control. With other words, the traceability must increase.

Correction of measurement results has been "taboo" in the clinical laboratory field for a long time, but during recent years there has been an increasing awareness of some basic principles of metrology (5), *e.g.* that *a measured value should be corrected for known systematic errors before it is stated.*
Since metrology is (1) the field of knowledge, which "includes all aspects both theoretical and practical with reference to measurements, whatever their uncertainty, and in whatever fields of science and technology they occur", clinical chemistry and other clinical laboratory disciplines should be no exceptions.

As mentioned in Chapter 2 (7), one of the main subprojects of the present NORDKEM project on 'Medical need for quality specifications in laboratory medicine' is focused on the pragmatic application of metrological principles in order to improve the transferability of clinical laboratory results. The project has been performed in two steps, an introductory phase (11), followed by a main phase.
The work in the main phase has been supported by the EU-AIM Openlabs project (A) which is gratefully acknowledged.

## HOW TO EXPRESS LABORATORY QUALITY SPECIFICATIONS

The Total Error (TE) of an analytical procedure *cp.*(14) can be expressed as a sum of different types of errors, which - in order to be meaningful - ought to be possible to estimate.

$$TE = SE_{corr} + SE_{noncorr} + \{\Delta SE \cdot s_{A0}\} + z\,\{\Delta RE \cdot s_{A0}\}$$

where

$SE_{corr}$ = a well-determined - and thus correctable - part of the total systematic error. Usually referred to as 'analytical bias. Can have a positive or negative value.

$SE_{noncorr}$ = a non-determined, and thus non-correctable part of the total systematic error. May be due to low specificity of the measurement procedure, *e. g.* matrix effects. Should be estimated as a maximum error.

$\Delta SE$ = a temporary increase of the systematic error, not detected/eliminated by the quality control system of the laboratory. This increase in error is expressed as a multiple of the inherent random error of the measurement procedure ($s_{A0}$). Can have a positive or negative value and be of the order $(0;\ \Delta SE_{detect})$.

$\Delta SE_{detect}$ = lower detection limit of QC procedure.

$\Delta RE$ = a temporary increase in random error not detected/eliminated by the quality control system of the laboratory. The change in random error is expressed as multiples of the inherent random error ($s_{A0}$), and is in the order $(1;\Delta RE_{detect})$.

$\Delta RE_{detect}$ = lower detection limit of QC procedure.

$z$ = a multiplier related to the portion of the distribution exceeding the quality requirement, often set as 1.65 to fix the maximum defect rate to 5 %.

A laboratory 'Analytical Quality Specification' (AQSpec) should be expressed as a 'Total Allowable Error' ($TE_a$) at a defined concentration. The total error TE should always be kept smaller than $TE_a$.

In the process of formulating AQSpecs (7) it is essential to compare the clinical goals with the characteristics of the analytical measurement procedures and the performance of the quality control procedures, and judge what is realistic for the laboratory to set up as

its AQSpecs. The quality specifications ought to have probability limits - *e. g.* that 95 % of the produced results should fall within these limits. In a 'Quality policy' document (8), it is also insistent to declare the 'traceability' of the values of the laboratory (for definition see Chapter 10).

## GENERAL CONSIDERATIONS FROM THE INTRODUCTORY PHASE OF THE STUDY

In the introductory phase of the project linear regression techniques were used for assessment of analytical stability and bias over a wider concentration range for selected analytical methods. The results of analyses of S--Creatinine and S--Urate methods in 17 laboratories in the Swedish Uppsala-Örebro health care region have been published (11). The analytical procedures of the various laboratories were described by their *'transformation functions'* (4)

$$c = k\,c^* + l\,;\ c_{min} \leq c^* \leq c_{max}$$

as estimated by linear regression analysis of results, $c$, from measurements on 'quality assessment material' distributed in a specially designed external programme. The analyses of unknown samples were performed weekly during three 14-week-periods.

Contrary to conventional external quality assessment programmes, with one or two concentration levels, the assigned values ($c^*$) were spread over the whole measurement interval of interest ($c_{min}$; $c_{max}$) (Fig 1A).

For illustration (Fig. 1B), the total allowable error ($TE_a$ )was here taken from the Medicare/CLIA criteria for minimum performance, *viz.* for S--Creatinine ± 15% or ± 26.5 μmol/L, whichever is larger. The wider dotted lines represent an area that should contain 95 % of the observations. The more narrow – · – lines represent the limits for the *'critically sized systematic error'*. Given a value of $TE_a$ the critically sized systematic error $\Delta SE_{crit}$ can be calculated with use of the simple formula:

$$\Delta SE_{crit} = TE_a / s_{A0} - 1.65\,;$$

and expressed in units of the standard deviation of the measurement procedure ($s_{A0}$). The acceptable region of the regression line $c = k\,c^* + l$ would be

$$c = c^* \pm \Delta SE_{crit} \cdot s_{A0}$$

where the inherent random error ($s_{A0}$) could be a more or less complex function of $c^*$, *e. g.* $s_{A0} = CV \cdot c^*$ (relative error) or $s_{A0}$ = constant (absolute error). The coefficient of analytical variation should be the actual value from the laboratory. An organizer of an external quality assessment programme may instead choose an average within-laboratory

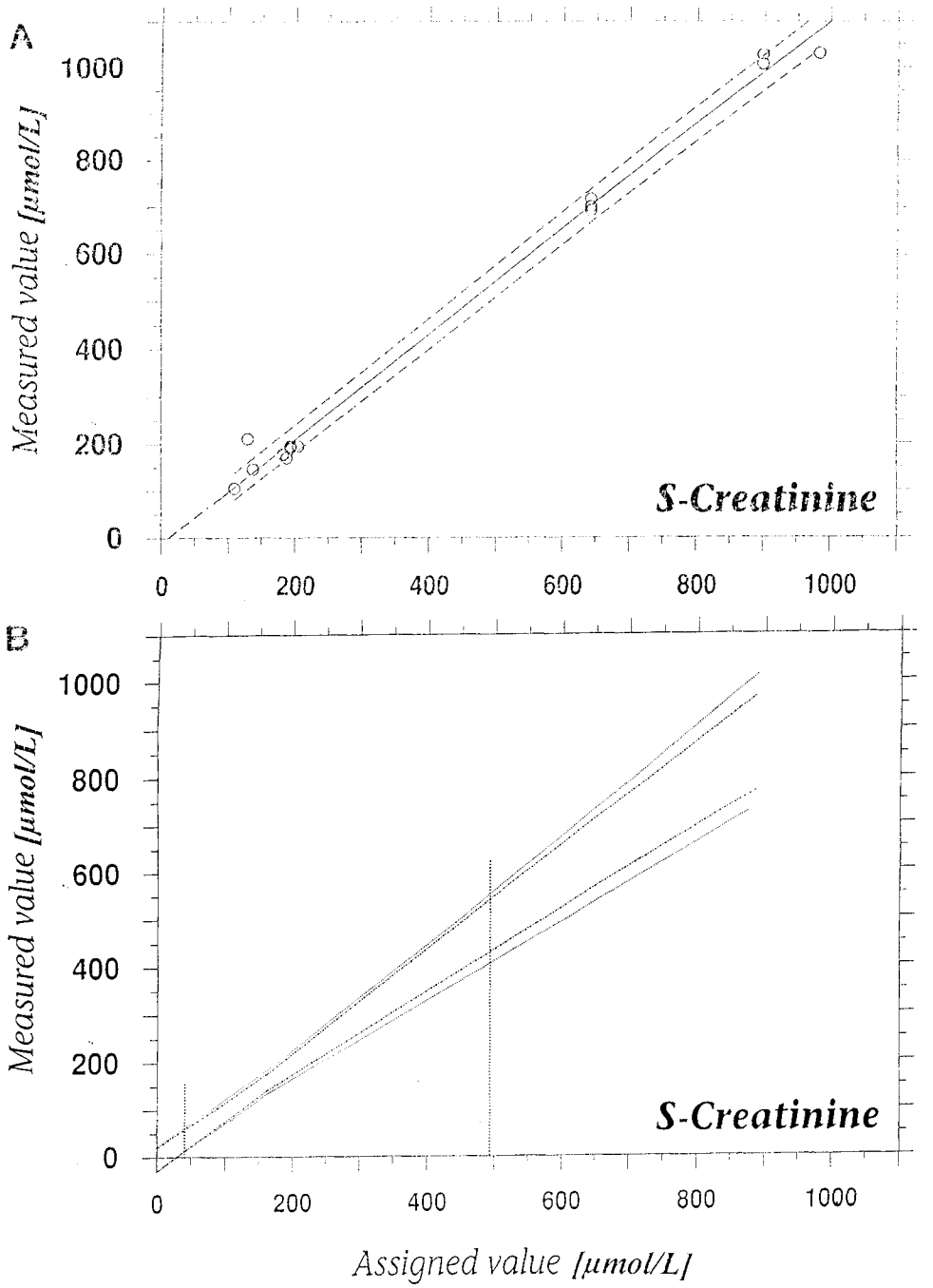


**Fig. 1A-B. A.** Shows the 'transformation function for a S--Creatinine method of a laboratory. A linear least squares fit of measured vs. assigned values from samples with "quality asessment material" from one laboratory during a 14-week study period. The broken thin lines indicate the 95% confidence limits for the regression line.
**B.** The dotted lines represent the limits for $TE_a$ taken from the Medicare/CLIA criteria effective in the USA. The more central –··– lines represent the limits for 'the critically sized systematic errors' ($\Delta SE_{crit}$ ) calculated from the 'total allowable error' ,$TE_a$ (see text). $c_{min}$ = 45 µmol/L; $c_{max}$ = 495 µmol/L.

variation ($CV_{AwL}$ ).

The regression lines are calculated from a number of measurement values over a defined time period. The slope (k) and the intercept (l) of the regression line can be used to characterize each laboratory in a strict error assessment procedure.

The results can be illustrated in a two-dimensional plot as shown for S--Creatinine in Fig. 2. Here S--Creatinine results from four of the seventeen laboratories in the introductory study were chosen (11). Each laboratory was assessed at three time periods (each of 14 weeks duration), and the performance with regard to analytical stability and bias is reflected in the clustering of consecutive points in the k-l plane. Ideally the slope should be 1.0 and the intercept equal to 0, but a comparison with the hexagonal *allowable k-l area* (specially marked in Fig. 2 and corresponding to the allowable region in Fig. 1B), shows that two of the four laboratories (marked with triangles and circles) fulfil the requirements with regard to both analytical 'trueness' and 'stability' (*cp.*11), and that another laboratory (squares) fulfils the requirements with regard to analytical 'stability', while the fourth laboratory (hexagons) fails with regard to both these characteristics of measurement, as they are defined by the Medicare/CLIA criteria for minimum performance within the measurement interval 45 - 495 μmol/L.

## THE MAIN PHASE STUDIES

In a continued study seven Nordic Clinical Chemistry Laboratories from five Nordic countries participated[1]. Three analyses were included in the study: S--Creatinine, S--Cholesterol, and S--Calcium. We gratefully acknowledge the production of a 'reference sample package' of 'analytical bias assessment material' by SERO A/S, Stasjonsveien. 44, N-1362 Billingstad, Norway. The 'package' consisted of five different samples with increasing concentrations of the three components according to Table 1. The 'package' was weekly analyzed one week-day during a four-week period and later during a period of at least two weeks together with 40 patients' samples.

[1]A full report of this project supported by NORDKEM will be published. The following contact persons participated and provided the measurement data: Kristoffer Hellsing Uppsala, Peter McNair Hvidovre, Lennart Nordström Karlstad, Elin Olafsdottir Reykjavik, Olli Peltola Turku, Gunnar Ronquist Uppsala, Helge Erik Solberg Oslo, and Kalle Willman Jyväskylä.

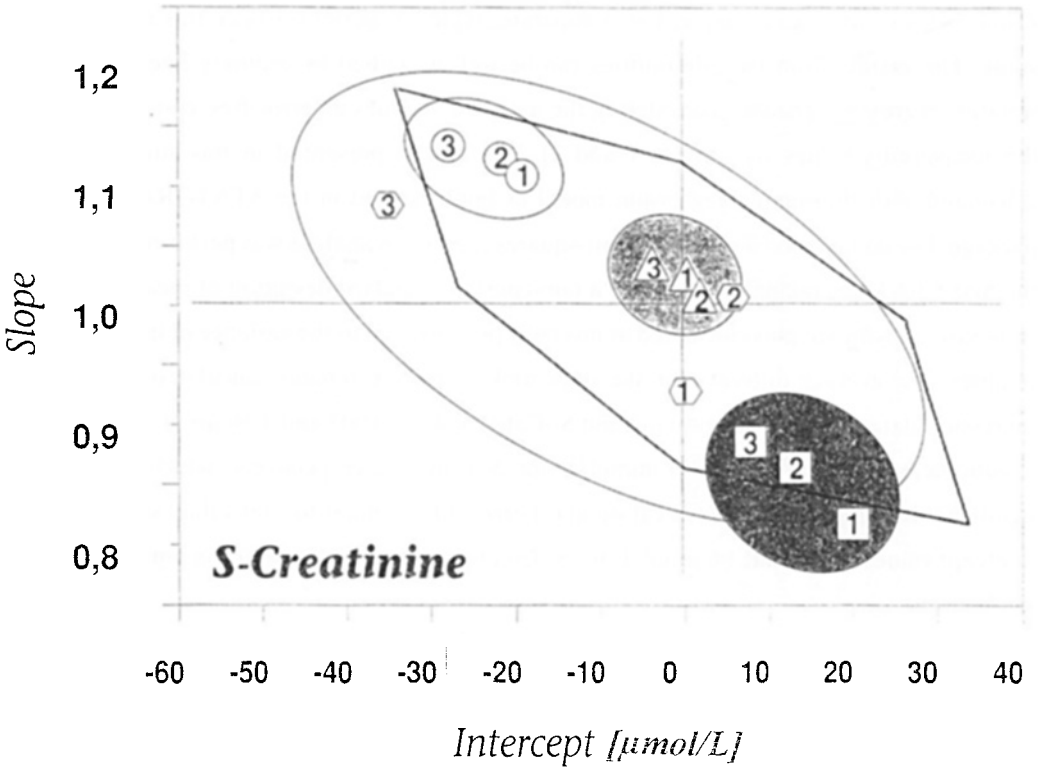


**Fig. 2.** S--Creatinine. Slope and intercept of regression lines of measured *vs.* assigned values for four selected laboratories plotted over three 14-week periods. The more or less shaded ellipses represent the 95% confidence area around an average value of slope and intercept for the three periods. The hexagonal area in the centre of the figure represents the 'allowable k-l region'. Intentional measurement interval: 45-495 µmol/L. Correction procedures will help a laboratory to hit such an area. Printed from (11) with permission.

**Table 1.** The Reference Sample Package contained five freeze-dried, serumlike reference samples. The ampoules were diluted with 5.00 mL distilled water. The assigned concentration values (mean ± 1 SEM; n = 10) were calculated using transferred values from an internal masterlot.

| | *S--Creatinine* | *S--Calcium* | *S--Cholesterol* |
|---|---|---|---|
| | *µmol/L* | *mmol/L* | *mmol/L* |
| Sample 1 | 45 ± 0.35 | 1.90 ± 0.003 | 2.14 ± 0.003 |
| Sample 2 | 101± 0.35 | 2.39 ± 0.003 | 4.52 ± 0.003 |
| Sample 3 | 140 ± 0.40 | 2.63 ± 0.003 | 5.32 ± 0.006 |
| Sample 4 | 270 ± 0.50 | 2.92 ± 0.003 | 5.96 ± 0.012 |
| Sample 5 | 495 ± 0.62 | 3.38 ± 0.003 | 6.83 ± 0.022 |

Using S--Calcium as an example, Fig. 3 illustrates typical experiences from this extended study. The results from the laboratories can be well described by ordinary linear least-squares regression analysis, considering the assigned x-values as error-free compared to the measured y-values (*cf.* Tables 1 and 3). The results presented in this study were calculated with this simple regression model as implemented in the STATGRAPHIC® package. For comparison a weighted least-squares regression analysis was performed using the SAS® package, taking into account a non-constant standard deviation of measured y-values by applying weights calculated as inversely proportional to the variance of individual y-values. The average difference in the slope and intercept estimates, calculated from 27 regression data sets for S--Creatinine and S--Calcium were 0.003 and 1.36 μmol/L for S--Creatinine, and 0.006 and 0.017 mmol/L for S--Calcium, respectively; which is non-significant in comparison with typical standard errors of estimates for individual slope and intercept values: 0.006 and 1.6 μmol/L for S--Creatinine, and 0.025 and 0.068 mmol/L for S--Calcium.

The 95 % confidence limits of estimated slopes and intercepts are represented in Fig. 3 by vertical and horizontal bars. The elliptic circumscriptions represent 95 % confidence regions in the k-l-plane. In our view, this is a comprehensive and objective way of describing the analytical performance of a laboratory. In Figs. 2 - 4 the target areas are indicated by rhomboids or hexagons. Defining the measurement (concentration) interval and the critical systematic errors ($\Delta SE_{crit}$) will provide a region in a plot similar to that in Fig. 1B. It is only lines that fully pass through this region that are considered to have acceptable k- and l-values. The measurement intervals and the $TE_a$ used for the calculation of the target areas in Fig. 2, 3, and 4A-C are listed in Table 2. The experiences are similar in this 'main phase study': some laboratories demonstrate 'good trueness' and 'analytical stability', some have bias problems and a few have both bias problems and less good analytical stability.

**Table 2.** The analytical measurement intervals and the total allowable errors (obtained from Medicare/CLIA, USA) used for calculation of the target areas in Figs. 2 - 4.

| *Analyte* | *Measurement interval* | | *Total allowable error, $TE_a$* |
|---|---|---|---|
| | $c_{min}$ | $c_{max}$ | |
| S--Creatinine | 45 μmol/L | 495 μmol/L | ± 15 % or 26.7 μmol/L |
| S--Calcium | 1.90 mmol/L | 3.38 mmol/L | ± 0.25 mmol/L |
| S--Cholesterol | 2.14 mmol/L | 6.83 mmol/L | ± 10 % |

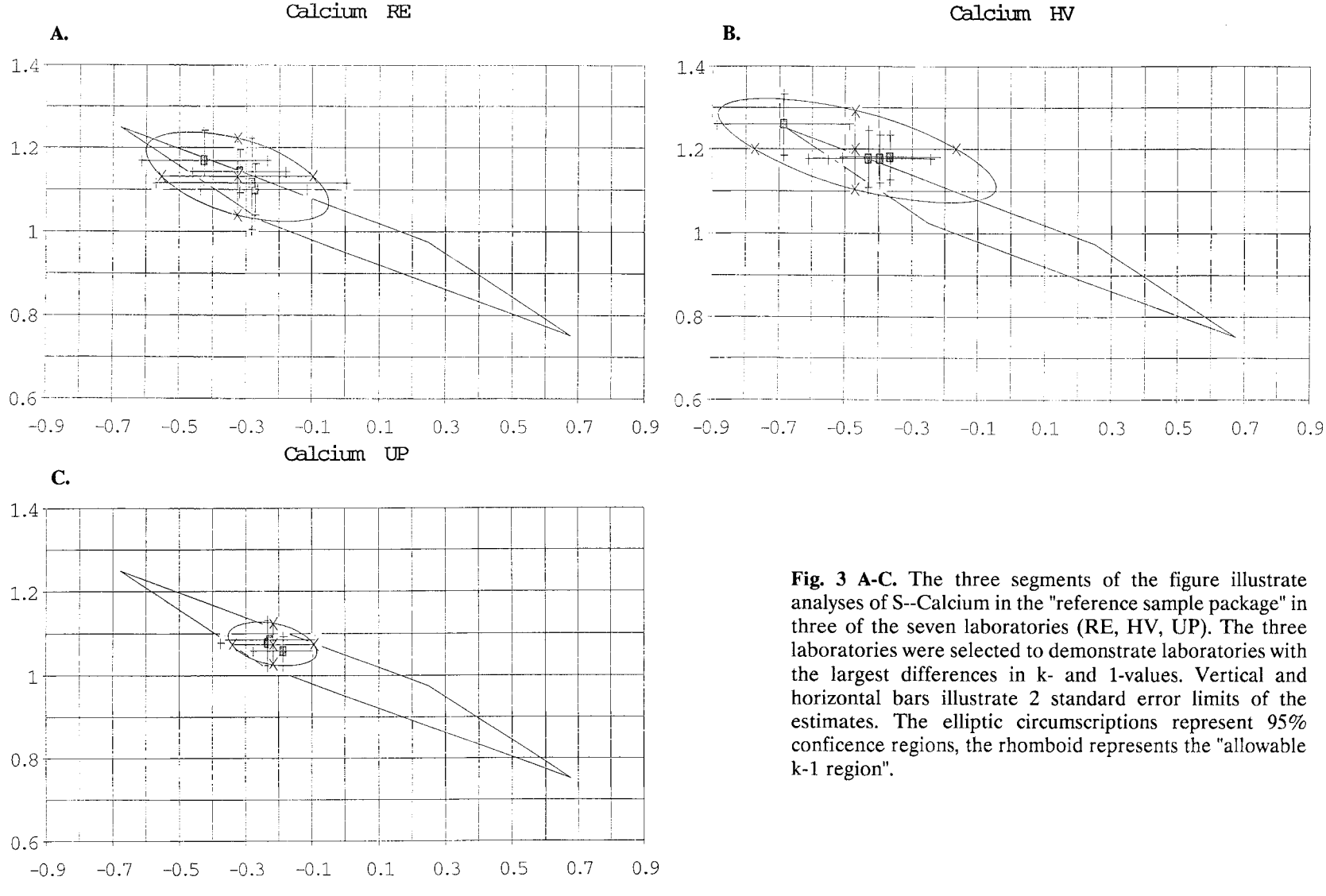


**Fig. 3 A-C.** The three segments of the figure illustrate analyses of S--Calcium in the "reference sample package" in three of the seven laboratories (RE, HV, UP). The three laboratories were selected to demonstrate laboratories with the largest differences in k- and 1-values. Vertical and horizontal bars illustrate 2 standard error limits of the estimates. The elliptic circumscriptions represent 95% conficence regions, the rhomboid represents the "allowable k-1 region".

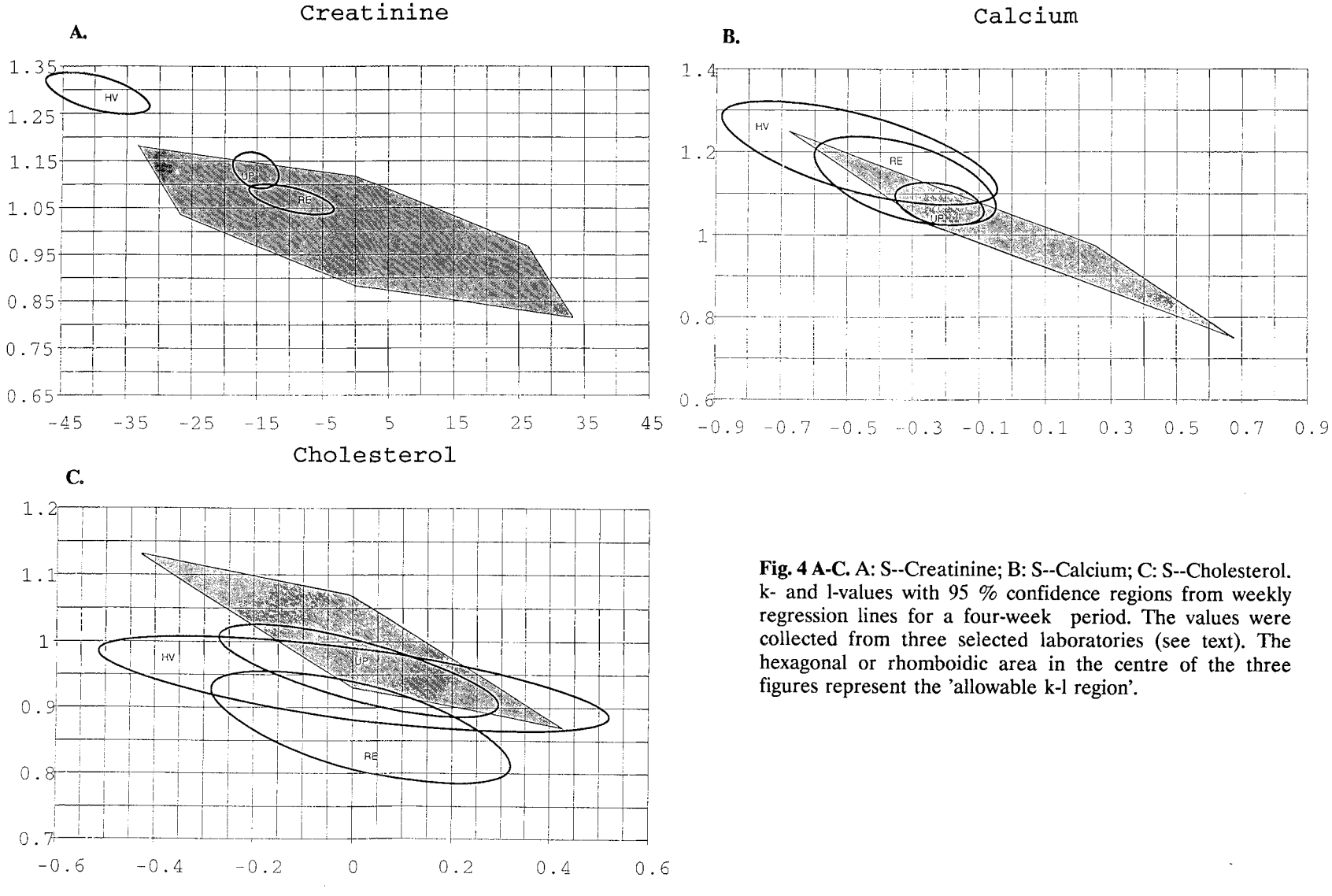


**Fig. 4 A-C.** A: S--Creatinine; B: S--Calcium; C: S--Cholesterol. k- and l-values with 95 % confidence regions from weekly regression lines for a four-week period. The values were collected from three selected laboratories (see text). The hexagonal or rhomboidic area in the centre of the three figures represent the 'allowable k-l region'.

## ATTEMPTS TO CORRECT MEASUREMENT VALUES FOR ANALYTICAL BIAS

One important objective of the 'main phase study' was to investigate if the application of a correction function would reduce the between-laboratory standard deviation ($s_{Ab}$). For that purpose 40 different, partly pooled, patient samples were collected and distributed to the seven laboratories, and analyzed there surrounded by the 'reference sample package'. In order to obtain high variation of the creatinine concentrations we tried to attain pooling of individual samples of similar creatinine concentration .

Under certain conditions, *e.g.* for analytical procedures with high specificity and documented stable analytical performance, numerical correction of the 'correctable systematic error' can be performed to decrease interlaboratory variation to a level specified by medical needs. Corrected values could be calculated with use of the 'correction function'

$$c_{corr} = ( c_{meas} - I ) / k ,$$

where $c_{meas}$ is the measured value. Figs. 5A-C illustrate the results of numerical correction of the measurement results. Along the y-axes the ratios between 'analytical between-laboratory variation' ($s_{AbL}$ ) for corrected and for directly measured values $[s_{AbL[corr.]}/s_{AbL[meas.]}]$ are presented in relation to the measured concentration values for S--Creatinine (A), S--Cholesterol (C), and S--Calcium (B). Fig 5A indicates that correction for all samples reduced the between-laboratory variation and that the mean ratio was as low as 0.45. It is also evident from Fig. 5A that the improvement by correction is concentration dependent. Fig. 5B demonstrates decrease in between-laboratory variation for most of the samples with a mean ratio of 0.81. There was no clear relation to measured concentration. Finally Fig. 5C does not show any improvement in between-laboratory variation by correction. The mean ratio increased to 1.40 and the change was concentration independent.

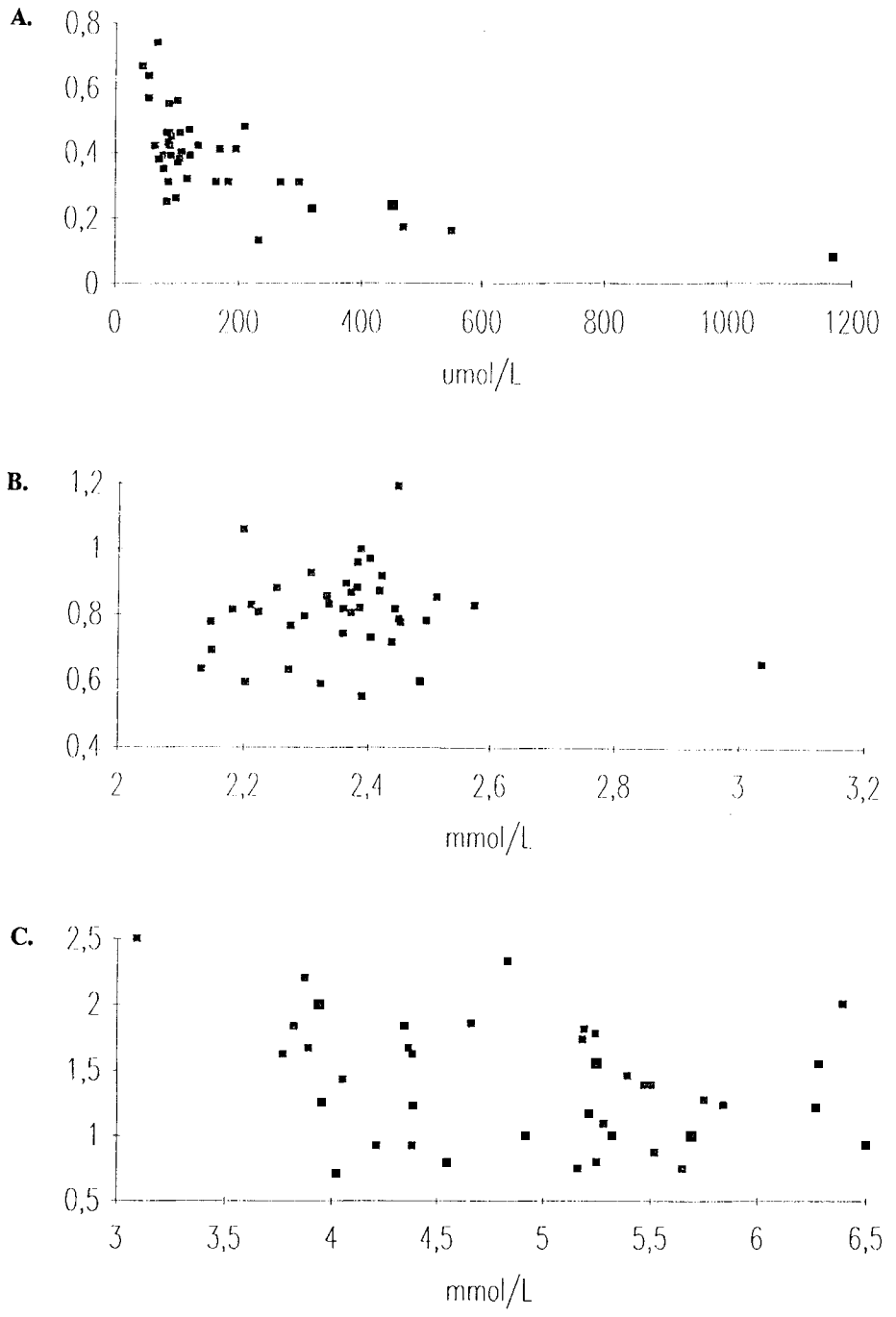


**Fig. 5 A-C.** A: S--Creatinine; B: S--Calcium; C: S--Cholesterol. The abscissa shows the measured substance concentration, the ordinate the ratio between variation (expressed as SD) of corrected concentrations and measured concentrations from 40 patient sera. Ratios below 1 indicate that correction procedures reduce between laboratory variation.

## DISCUSSION and CONCLUSIONS

In the present study it is shown that it is possible to reduce analytical between-laboratory variation to a remarkable degree by correction procedures for specially selected analyses. Logically, analytical methods with high stability, low analytical within-laboratory variation ($s_{AwL}$) and relatively high between-laboratory variation ($s_{AbL}$) ought to be good candidates for correction programmes. In Table 3 we have compiled CV-data and the mean values of the ratios of analytical 'between-laboratory variation' ($s_{AbL}$ ) for corrected and for directly measured values.

**Table 3.** Comparison between the means for the ratios of the standard deviations for the corrected and the directly measured values (bottom of table) and the ratios between the coefficients of variations for within and between laboratory variation. The top line presents estimates from other laboratories of the average CV for within-laboratory variation. These values are used together with the average total CV obtained in this study for the calculation of the average between-laboratory CV.

| | *S-Creatinine* | *S-Calcium* | *S-Cholesterol* |
|---|---|---|---|
| $CV_{AwL}$ % | 2.0 | 1.5 | 1.9 (pathonorm) |
| $CV_{AL}$ % | 7.8 | 2.0 | 2.3 |
| $CV_{AbL}$ % | 6.2 | 1.3 | 1.2 |
| $CV_{AwL}/CV_{AbL}$ | 0.32 | 1.15 | 1.58 |
| Mean ratio | 0.45 | 0.81 | 1.40 |

$s_{AbL[corr.]} / s_{AbL[meas.]}$

The $CV_{AwL}$-data were obtained from Ref (9), $CV_A$ data from the present study and the $CV_{AbL}$-values were calculated using analysis of variance. Comparison of the ratios between $CV_{AwL} / CV_{AbL}$ and the ratios indicating effect of correction, demonstrate that both are lowest for the S--Creatinine method and highest for S--Cholesterol.

It is evident that clinical laboratories of today are not able any longer to live an isolated life of their own. Patients are changing hospitals and doctors, health care units are changing clinical laboratories, and decision levels generally assessed in single or groups of

clinics with access to clinical laboratories with different calibration procedures, equipment, and reagents. Traceability is today stressed in connection with application for certification or accreditation of clinical laboratories. It is one way to try to deal with the problem. The present study based on results from seven large Nordic clinical chemical laboratories shows that further improvement can be obtained. In the ongoing European discussions of "quality development programmes" in clinical laboratories, see *e.g.* articles (10,13) in a recent issue of EQAnews, there seems to be a general understanding that "the aims and procedures of the external quality assessment (EQA) schemes should be reviewed to better accommodate aspects of transferability and descriptions of laboratories' performance". We propose that "multi-level analytical bias assessment programmes" of the design described in the present paper should complement/replace the conventional type of passive EQA activities. The proposed statistical approach and the graphical presentation provides a means to assess "analytical stability" (definition see Chapter 11) and "analytical bias" of individual laboratories in relation to specified analytical quality requirements (AQSpecs). In addition we are also advocating for a more common use of numerical correction procedures in clinical chemistry in order to increase the transferability of measurement results.

An often heard argument against such corrections is that matrix effects may lead to wrong conclusions and that mathematical corrections may introduce ambiguities (cf. 3,12). As emphasized in the beginning of this paper, this type of correction procedure should only be applied to correct for well-known and well-determined analytical bias ("stable bias"); this presupposes that matrix effects on the measurements on "bias assessment material" can be disregarded. The non-determined part of the total systematic error of patient results, *e.g.* due to matrix effects should not be corrected for, but estimated and reported as a maximum error limit. With a relatively large maximum error estimate of the latter type of systematic error component, it may seem less important to correct for "stable bias". It goes without saying that non-specific analytical methods should be replaced by more specific methods, if available and affordable. But even with highly specific methods there may be a need for estimation (and correction) for analytical bias due to other factors; this is according to basic metrological principles.

It is also evident that a correction may be more effective for methods with a small ratio of analytical within-laboratory to between-laboratory variation.

An "Analytical Quality Assurance" (AQA) programme (see Chapter 11), combining analytical bias assessment and active correction, can be implemented in different ways, *e.g.*

1) as an external programme providing the applicable correction functions when there is

a need to transfer laboratory results "over time and space"; or

2) as an external assessment followed by internal correction of measurement result before reporting.

It should be noted that "analytical bias assessment and correction programmes" should be run strictly separate from EQA and Proficiency Testing programmes organized by various regulatory bodies.

Supervised correction procedures could - as indicated earlier in this paper - be of great value especially in the following situations:

- when used within a laboratory organization to compare results over time, and from different equipments
- when importing or exporting reference values from/to another laboratory
- when frequently exchanging patients' data between two hospitals or health care organizations
- when health care organizations are issuing new general recommendations for clinical actions based on laboratory data.

We are convinced that with the present development of analytical stability in many analytical measurement systems and of new technology for interlaboratory data communication the time now has come to start implementing systems for correction and transfer of clinical laboratory data.

## REFERENCES


1. BIPM, IEC, IFCC, ISO, IUPAC,IUPAP,OIML. International Vocabulary of Basic and General Terms in Metrology, 2nd ed. ISO, 1993:1-59.
2. de Verdier C-H. Improved transferability of clinical laboratory data. Short survey of advantages and problems. Scand J Clin Lab Invest 1984;44 Suppl 171:305.
3. de Verdier C-H, Groth T. Desirable Analytical Quality Specifications for Clinical Chemistry Laboratories. Letter to the Editor. Clin Chem 1994; 40(1):168-169.
4. Groth T. Strategies for decreasing interlaboratory variation. Studies by computer simulation. Scand J Clin Lab Invest 1984;44 Suppl 171:331-347.
5. Groth T. Internal and external quality assurance. Scand J Clin Lab Invest 1989;49 Suppl 193:87-94.

6. Groth, T. and de Verdier, C.-H. Analytical quality goals and assessment to ensure transferability of laboratory results. Clin Chim Acta; 1993; 222:129-139.
7. Groth T, de Verdier C-H, Hyltoft Petersen P. Background. Upsala J Med Sciences 1993;98(3):191-193.
8. ISO . Quality - Vocabulary, 1986:1-16.
9. Jansen RTP, Bullock DG, Vassault A, et al. Between-country comparability of clinical chemistry results: an international quality assessment survey of 17 analytes in six European countries through existing National Schemes. Ann Clin Biochem 1993;30:304-314.
10. Kallner A. Quality Improvement in Laboratories by Accreditation according to EN 45001. EQAnews 1993;4(3):2-3.
11. Olafsdottir E, Aronsson T, Groth T, de Verdier C-H. Transferability of Clinical Laboratory Data within a Health Care Region. Scand J Clin Lab Invest 1992; 52:679-687.
12. Stöckl D. Desirable Performance Criteria for Quantitative Measurements in medical laboratories Based on Biological Analytical Variation-Hindrances to Reaching Some and Reasons to Surpass Some. Letter to the Editor. Clin Chem 1993 (5):913-914.
13. Uldall A. Quality Development Programmes Should Be the Main Scope for EQA Work. EQAnews 1993;4(3):4-5.
14. Westgard JO. Charts of Operational Process Specifications ("OPSpecs Charts") for Assessing the Precision, Accuracy, and Quality Control Needed to Satisfy Proficiency Testing Performance Criteria. Clin Chem 1992;38:1226-1233.

*Correspondence:*
Carl-Henric de Verdier, M.D., Professor
Department of Clinical Chemistry
University Hospital
S-751 85 Uppsala, Sweden