The control of analytical accuracy and day to day precision helps for the follow-up of patients and is essential when using biochemical data from the literature.

C. Bachmann
Laboratoire Central de Chimie Clinique
Centre Hospitalier Universitaire Vaudois
CH 1011 Lausanne, Switzerland

This article was published as
ERNDIM Biomed2 Newsletter
Number 3 in May 2000

The theoretical background of Quality Control (QC) has been presented in an earlier BIOMEDII ERNDIM Newsletter (February 1999).
For stimulating participants of the ERNDIM QAP we want to remind you that the internal QC results allow you to assess rationally (no guess-work!) whether consecutive changes of laboratory results are most likely reflecting the biochemical course of the disease and of its treatment or if such changes can merely be explained by the existing day to day imprecision of the analytical process. In the appendix we describe our recipe for preparing QC-material for amino acid analysis to allow you to do this yourself. Finally we discuss the need for external quality control and awareness of the accuracy of your results in making recommendations or comparisons with data from the literature for managing your patients.

Precision

Let us assume that you are monitoring the treatment of a 2 month old patient in the out-patient clinic. PKU has been detected in this boy by newborn screening; you have confirmed the increased plasma phenylalanine in plasma by chromatography and a defect of tetrahydrobiopterin metabolism has been excluded. At the last control, 2 weeks ago, his phenylalanine was 150 µmol/L with a prescribed nutritional intake of 45 mg Phe /kg per day.
You assay the phenylalanine now and find 64 µmol/L.

Can this decrease between the 2 consultations merely be explained by the expected analytical imprecision or is it due to overanxious parents who restrict the intake of natural protein even more than you prescribed and thereby excessive restriction of phenylalanine? How can you find out ?

Your analytical imprecision should be known to you from the interseries imprecision results of the internal quality control of your lab. This is the day to day variance of the analytical process tested in a sample material which is identical to the patients analytical samples (i.e. heparinized plasma in the example) at a concentration which is reasonably close to the critical levels you face (e.g. intervention or diagnostic limits) and for which you have determined in your setting the acceptable mean (and confidence limits calculated from SEM = SD/ n) and multiples of standard deviation and calculated the coefficient of variation (CV= SD/mean).

Population variance (SD^2) of a analyte (e.g. in reference values) is the sum of the inter-individual, the intra-individual and the analytical variance. In the situation presented here i.e. in the follow-up with 2 sequential values, we deal with a single subject, so that inter-individual variance does not apply. We are interested in determining which part of the fluctuations of the two results can be attributed to the analytical variance and which is due to intra-individual variance. The latter is dependent on the impact of treatment and evolution of the disorder. Also it should not be forgotten that preanalytical factors also play an important role. These are e.g. the time of sampling (fasting, time after last meal), the delay and ambient temperature during any delay (cystine drops!) before centrifugation of the blood, the patients muscular activity (increase of glutamine after seizures) or the type of sample (capillary vs. venous) .

If it is ensured that the sampling conditions are well controlled and the analytical imprecision is known, we can assess whether the critical difference for analytical results of two consecutive samples can be attributed, with a 95 probability, to analytical variance alone:

To perform the calculation we need:
1. The two results (x_1 et x_2) of the patient and compute the actual difference
2. The coefficient of variation of the day to day quality control for which the target value is of the order of magnitude of the higher of the two patient samples.

One compares the actual (absolute) difference with the critical difference.

For those who like statistics : as in the t-test , p = 2.5% to p=97.5% the integral of which with Gaussian distribution = 1.96 =Z ;
Test quotient compared to Z : Difference between the 2 values/ (SD of difference). SD = mean × Coefficient of variation (as %)/100.

Actual difference = $\left|x_1 - x_2\right|$

Critical difference = $1.96 \times \sqrt{2} \times x_{\max} \times CV(\%)/100 = 2.82 \times x_{\max} \times CV(\%)/100$

$$Critical\,Difference \approx 3 \times x_{\max} \times CV(\%)/100$$

where x_{max} is the higher of the 2 results. If the actual difference is larger than the critical difference, the decrease (or increase) observed cannot be merely "explained" by the analytical imprecision.

Taking the example above :
Actual difference of phenylalanine (µmol/l) = 150 – 64 = 86 µmol/L

If your day to day imprecision at the time is 7% (CV at a level of e.g. 96 µmol/L (target value))
the critical difference will be : 2.82 * 150 µmol/L * 0.07 = 30 µmol/L

The actual difference is greater than this critical difference, thus the drop cannot be explained solely by analytical imprecision.
If however a method with a higher CV, of let us say 18%, is used e.g. the Guthrie Test, the critical difference will be 76 µmol/l (2.82 * 150 µmol/L * 0.18), and the change from one control visit to the next can then be fully explained by the analytical variance of this screening method (which is not designed for the follow-up of patients). With such an imprecision the conclusion that the patient tends to be dangerously deprived of phenylalanine by excessive treatment cannot be made since the drop of phenylalanine can fully be attributed to the analytical imprecision. You will have to use another method and rapidly control the patient again.

From the responses to the questionnaire for the ERNDIM Laboratory Directory it is evident that some laboratories do not perform internal quality control routinely; thus a practical procedure for preparing the QC-material yourself is presented in the appendix. The results of three critical amino acids in the QC of one batch used during 17 months is shown in Figure 1. It demonstrates that such a material stored at –80°C in aliquots is stable for a year. The cause of the increase during the last few months could be the water loss by "lyophilisation" when sample tubes are not hermetically closed.
Since amino acids such as glutamine are only stable at an acid pH the QC material we use is not untreated plasma but includes the sample treatment before chromatography. Thus in addition to the amino acid concentrations in the QC material we also evaluate the areas of the internal standards in the patients samples and compare them to those of the QC material to make sure that there was no error made in the sample preparation (pipetting imprecision when adding 25 µl of the deproteinisation-internal standard solution to 100 µl patient plasma !).

Accuracy

Due to the rarity of patients with IEMs it is virtually impossible for any single centre to follow enough patients and practice independently of observations or recommendations of other groups. We all use data from the literature, preferably from prospective multicenter studies which pool a significant number of affected patients. Lessons can be learnt from the exemplary work which has been carried out in the domain of assessing cardiovascular risk factors. Method standardisation and external quality control programs (WHO) allow us to minimise the bias of cholesterol determination (below 3%). We propose that in our field of rare inherited errors of metabolism we should also aim at the goal reached with cholesterol determination. We have chosen that example to illustrate the impact of analytical inaccuracy on preventive therapeutic decisions.

What would be the consequences if a laboratory had a bias for cholesterol determination of plus / minus 10% at a decision level of 6.5 mmol/L ?

Figure 2 shows the histogram of serum cholesterol levels in our normal male population of 45-54 years of age. 30% of the population has a level above 6.5 mmol/L and comes into consideration for chronic preventive treatment. If a laboratory had a bias of cholesterol measurement of +10%, which means that it would give a result of 6.5 mmol/L when the true cholesterol is actually 5.85 mmol/L, an additional 21 % of the patients would be alarmed and possibly subjected to an excessive preventive treatment. Worse, if the bias were – 10% the laboratory would miss 17% of the patients who would actually need a treatment.

As already mentioned we are often forced to use data from the literature for amino acids, organic acids, carnitine, lactate, pyruvate and other special assays. For example in adapting treatment in PKU we feel secure if the phenylalanine values lie between 40 and 250 µmol/l until puberty. We might decide to increase the supply of essential amino acids in urea cycle disorders or supplement isoleucine in propionic or methylmalonic acidaemia when fasting isoleucine in plasma drops below 25 µmol/L in order to avoid appetite loss and chronic catabolism. Lactate/pyruvate ratios above 30 lead us to focus investigations on respiratory chain disorders or pyruvate carboxylase deficiency and dismiss other genetic causes of lactic acidemia. In those cases where we have not established our own age dependent reference values but use those of the literature are we really sure that our own analytical methods (and sampling conditions) correspond to those of the authors whose magic numbers we use ? The results of the QAP programs lead us think otherwise and that major efforts need to be made to correct the analytical bias in order to avoid false decisions. Especially for the follow-up of patients and the optimisation of treatment we have definitely evolved from qualitative or semi-quantitative assay results to quantitative results with their inherent imprecision and inaccuracy which we must take into account. However waiting for clinical signs and symptoms cannot be the criteria for assuring the plausibility of laboratory data since irreversible damage might have already occurred.

It is obvious that multicenter studies with several laboratories involved need to have a reliable QAP for avoiding systematic errors.

Summary

The results of internal quality control are not only helpful for early detection of random or systematic analytical errors by using the Westgard rules (http:\www.westgard.com), but also for a rational use of consecutive results in the follow-up of patients. For collaborative studies and the use of published reference values or recommendations based on biochemical data, information about the accuracy of own measurements is a necessity for applying extraneous data correctly. When using data from other sources one should not only compare the methodology but also the performance including accuracy and precision. Such data, even if they are not published, can be requested from the authors of published literature.

We encourage all ERNDIM participants to review their latest QAP results in conjunction with their decision algorithms for prevention, diagnosis and adaptation of treatment as illustrated above for phenylalanine and to draw the necessary conclusions.

References:

Fraser CG, Petersen PH (1991) The importance of imprecision. Ann Clin Biochem 28:207-211

Jones RG, Payne RB (Eds) Clinical Investigation and Statistics in Laboratory Medicine, ACB Venture Publications, Association of Clinical Biochemists, London 1997 (ISBN 0 902429 21 3)

Appendix

Preparation of QC-material for internal quality control.
First calculate approximately how much material you will need for a year. We perform an internal QC analysis after every tenth analysis (and at least once per week) and after every calibration (and change of reagents). To this sum you add 30 samples (for establishing the initial mean and standard deviation). Multiply the number of samples by the sample volume and add 10 % for pipetting loss. Multiply the whole by the number of levels at which you will do a control (decision levels), e.g. 2 levels, one rather low or normal and one high.
Estimate from this how much fresh blood you need assuming a hematocrit of 50%.

Obtain from the transfusion centre enough fresh blood collected from a fasting subject with the identical anticoagulant (e.g. lithium heparin) you use for your patient sampling (not ACD-blood!). Have this tested for HIV and hepatitis. It might be advisable to check also that the subject has no cryoglobulins. Centrifuge rapidly and collect the plasma. Assay it for establishing the amino acid concentration and freeze all the plasma.

> Remark : Commercial preparations of serum (e.g. calf serum) are not suitable because they usually have been heated which leads to losses of amino acids and gross increase of glutamate (with relatively low glutamine).

Deproteinize the plasma as you do for patients samples but on a higher scale: Per 100 ml of plasma (e.g. in 50 ml Falcon tubes) add 25 ml of the deproteinisation solution which also contains the internal standard(s).You might choose to add solid amino acids for spiking. The deproteinisation solution contains 16.0g of sulfosalicylic acid / 100 ml of nanopure water.
Vortex for 2 minutes, distribute it into tubes which withstand centrifugation at 15'000 x g
Centrifuge : 30 min at 4°C at 15'000 x g
Recover supernatant and repeat the centrifugation of the supernatants twice.
Pool the final supernatants and dilute with buffer, as you would do for the sample preparation before chromatography mixing carefully in the cold. We use a pH 3.3 chromatography buffer (1:1 v/v) which results in a final pH of 2.2. An acid pH is needed for avoiding losses of glutamine and asparagine even if the samples are deep frozen.
Verify the pH, aliquot the whole pool (e.g. 180µl) into Eppendorf tubes which should be tightly closed.
The tubes are labelled and stored at -80°C.
Assay twenty samples of QC with recalibrations of your method between each of the assays in order to calculate the starting target values of mean and standard deviation of this QC material.

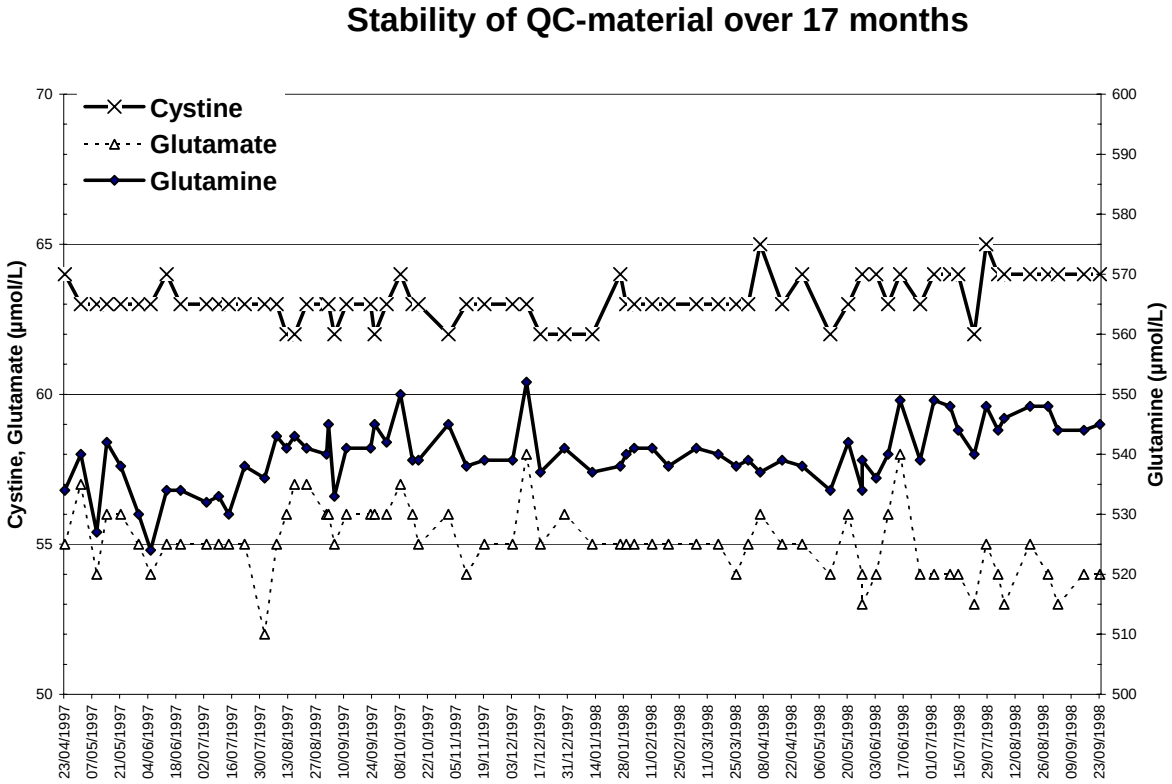

Figure 1

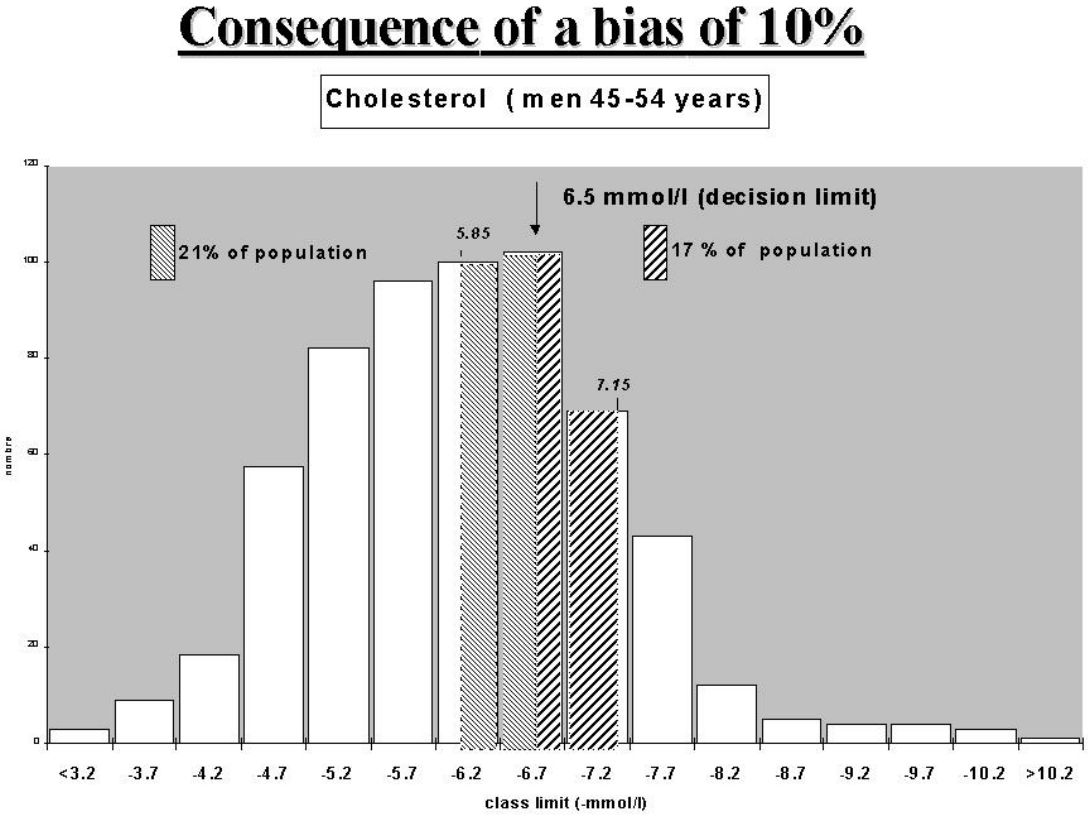

Figure 2