Combination of analytical quality specifications based on biological within- and between-subject variation

Per Hyltoft Petersen[1,2], Callum G Fraser[3], Lone Jørgensen[4], Ivan Brandslund[4], Marta Stahl[4], Elizabeth MS Gowans[5], Jean-Claude Libeer[6] and Carmen Ricós[7]

Addresses
[1]Department of Clinical Biochemistry
Odense University Hospital
DK-5000 Odense C, Denmark
[2]NOKLUS (Norwegian Centre for External Quality Assurance of Primary Care Laboratories)
Division for General Practice
University of Bergen
Bergen, Norway
[3]Biochemical Medicine
Tayside University Hospitals NHS Trust
Ninewells Hospital and Medical School
Dundee, UK
[4]Department of Clinical Biochemistry
Vejle Hospital, Vejle, Denmark
[5]Department of Biochemistry
Princess Royal Hospital
Apley Castle, Telford, UK
[6]Scientific Institute of Public Health
Brussels, Belgium
[7]Laboratoris Clinnics Hospital, Vall d'Hebron
Barcelona, Spain

Correspondence
Per Hyltoft Petersen, MS
E-mail:
per.hyltoft.petersen@ouh.fyns-amt.dk

Abstract

At a conference on 'Strategies to Set Global Analytical Quality Specifications in Laboratory Medicine' in Stockholm 1999, a hierarchy of models to set analytical quality specifications was decided. The consensus agreement from the conference defined the highest level as 'evaluation of the effect of analytical performance on clinical outcomes in specific clinical settings' and the second level as 'data based on components of biological variation'. Here, the many proposals for analytical quality specifications based on biological variation are examined and the outcomes of the different models for maximum allowable combined analytical imprecision and bias are illustrated graphically. The following models were investigated. (1) The Cotlove *et al.* (1970) model defining analytical imprecision ($\%CV_A$) in relation to the within-subject biological variation ($\%CV_{W\text{-}S}$) as: $\%CV_A \leq 0{\cdot}5 \times \%CV_{W\text{-}S}$ (where %CV is percentage coefficient of variation). (2) The Gowans *et al.* (1988) concept, which defines a functional relationship between analytical imprecision and bias for the maximum allowable combination of errors for the purpose of sharing common reference intervals. (3) The European Group for the Evaluation of Reagents and Analytical Systems in Laboratory Medicine (EGE Lab) Working Group concept, which combines the Cotlove model with the Gowans concept using the maximal acceptable bias. (4) The External Quality Assessment (EQA) Organizers Working Group concept, which is close to the EGE Lab Working Group concept, but follows the Gowans *et al.* concept of imprecision up to the limit defined by the model of Cotlove *et al.* (5) The 'three-level' concept classifying analytical quality into three levels: optimum, desirable and minimum. The figures created clearly demonstrated that the results obtained were determined by the basic assumptions made. When $\%CV_{W\text{-}S}$ is small compared with the population-based coefficient of variation [$\%CV_P = (\%CV^2_{W\text{-}S} + \%CV^2_{B\text{-}S})^{1/2}$], the EGE Lab and EQA Organizers Working Group concepts become similar. Examples of analytical quality specifications based on biological variations are listed and an application on external quality control is illustrated for plasma creatinine.

Ann Clin Biochem 2002; **39**: 543–550

Introduction

A number of proposals for analytical quality specifications based on biological variation have been presented. One of those most widely cited and used, the model of Cotlove *et al.*,[1] is based on biological within-subject variation; others are based on total biological (population-based, i.e. within-subject plus between-subject, sometimes called 'group') variation. Furthermore, certain of the quality specifications relate to analytical imprecision only, whereas others combine analytical imprecision and bias according to different models. The concept of Gowans *et al.*[2] specifies the relationship between the maximum combination of analytical bias and imprecision for laboratories sharing common reference intervals within a defined geographical area.

Two European working groups under the auspices of European Group for the Evaluation of Reagents and Analytical Systems in Laboratory Medicine (EGE

Lab)[3] and EQA Organizers[4] have combined the concept of Cotlove *et al.*[1] for imprecision with the model of Gowans *et al.*[2] for maximum allowable bias in, for example, their recommendations for desirable analytical quality specifications to be used in external quality assessment (EQA).[5] Analytical quality specifications were generated from a combination of an imprecision goal of $\%CV_A \leqslant 0{\cdot}5 \times \%CV_{W\text{-}S}$ and a bias goal of $B \leqslant 0{\cdot}25 \times (\%CV^2_{W\text{-}S}+\%CV^2_{B\text{-}S})^{1/2}$ (where %CV is percentage coefficient of variation, 'A' denotes analytical, 'W-S' denotes within-subject, B is bias and 'B-S' denotes between-subject) by the EGE Lab Working Group[3] and a more complicated relationship between the same specifications in the EQA Organizers Working Group.[4] The relationships between imprecision and bias, however, were not clarified in detail. Further, Fraser *et al.*[6] added a three-level proposal to strategies for setting specifications for analytical quality with the same two formulae, as expressed with the values of a and b related to the factors in the general formulae $\%CV_A \leqslant a \times \%CV_{W\text{-}S}$ and $B \leqslant b \times (\%CV^2_{W\text{-}S}+\%CV^2_{B\text{-}S})^{1/2}$, where a and b are, respectively, 0·25 and 0·125 for optimum quality, 0·5 and 0·25 for desirable quality (as above) and 0·75 and 0·375 for minimum quality.

Since there is an ever-expanding demand for a combination of imprecision and bias as one single 'total error' (TE) specification, Fraser and Hyltoft Petersen[7] defined this total error (TE_A) by combining imprecision and bias according to the principle of Westgard *et al.*[8] $\{TE_A \leqslant |\pm[1{\cdot}65 \times (0{\cdot}5 \times \%CV_{W\text{-}S}) + 0{\cdot}25 \times (\%CV^2_{W\text{-}S}+\%CV^2_{B\text{-}S})^{1/2}]|\}$, where 1·65 is the z-value corresponding to a 95% confidence one-tailed probability. This formula was intended to make the combination easy to handle, but may have introduced more problems than solutions compared with the Westgard total error concept,[8] as the combination of the specifications for bias and imprecision to total error is fixed when the z-value is decided, whereas a total error can be divided into a great number of combinations of bias, imprecision and z-values. Consequently, the purpose of this overview is an attempt to clarify the relationships between the different assumptions and models that are possible for the generation of analytical quality specifications based on components of biological variation and to elucidate their relationships to the total error concept.

Basic considerations

Analytical quality specifications have been specified separately for imprecision[1] and bias[2] and for combinations of these,[3,4,6] as well as for total error.[7] The first proposals were developed solely for specifications of imprecision based on variance models simply by defining a fraction of the biological standard deviation, usually within-subject variation. Models followed for analytical quality specifications combining the separate terms of imprecision and bias in allowable combined errors according to more complicated models. Models for control of analytical quality were introduced some 30 years ago both for imprecision and bias separately and for total error, defined as $TE = B + z \times s$, where s is the stable (inherent) imprecision.[9] This latter model was then used for maximum allowable total error,[7] for which the derived analytical quality specifications for maximum allowable imprecision and bias were used in the formula. It is important to understand that the analytical quality specifications are based on the effect of analytical error on clinical decision-making, whereas total error is a combination of imprecision and bias for the purpose of, for example, EQA.[11]

Different theories and combinations of specifications

Quality specification for analytical precision

The model of Cotlove *et al.*[1] was based on the idea that the effect of analytical imprecision should not significantly affect monitoring of healthy individuals (and patients). This desirable low impact of imprecision can be obtained when $\%CV_A \leqslant 0{\cdot}5 \times \%CV_{W\text{-}S}$, as the combined (total) %CV would not increase by more than 12% compared with the pure biological $\%CV_{W\text{-}S}$; in other words, the clinical 'signal' would not be much confounded by the analytical 'noise'.

Quality specifications for analytical bias

Gowans *et al.*[2] based their approach on the guidelines promulgated by the Expert Panel on Theory of Reference Values of the International Federation for Clinical Chemistry (IFCC),[10] that a minimum sample size of 120 reference subjects should be used for estimation of a reference interval and that the 90% confidence interval around each reference limit is approximately $0{\cdot}25 \times (\%CV^2_{W\text{-}S}+\%CV^2_{B\text{-}S})^{1/2}$. The idea of Gowans *et al.*[2] was that it would be advantageous for the transfer of test result data between different geographical centres to have common reference intervals. This could be achieved by using much larger sample sizes than 120, as these have negligible confidence intervals, for establishing the common reference intervals, and then using the 90% confidence interval allowed by the guidelines of the IFCC (see above) as a quality specification for all laboratories using the common reference intervals. The formula for the inter-relation between imprecision and bias is then $B = (1{\cdot}96+0{\cdot}25) \times (\%CV^2_{W\text{-}S} + \%CV^2_{B\text{-}S})^{1/2} - 1{\cdot}96 \times (\%CV^2_{W\text{-}S}+\%CV^2_{B\text{-}S}+\%CV^2_A)^{1/2}$, the second component of this formula representing the dispersion of the population-based reference interval.

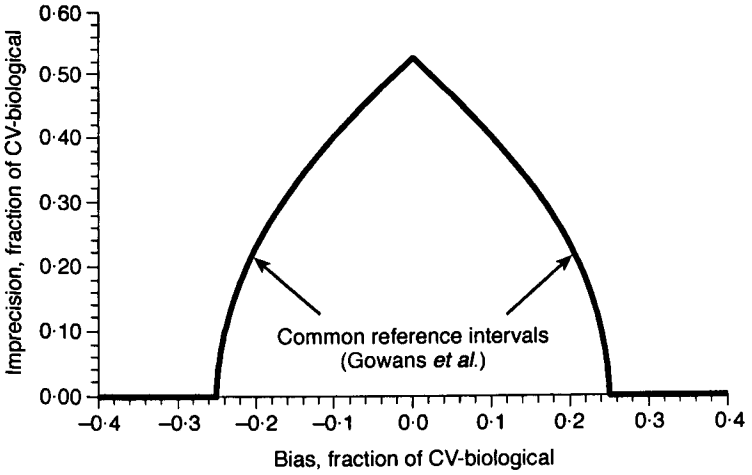

Figure 1. *Illustration of the analytical quality specifications for sharing common reference intervals according to Gowans* et al.[2] *The plot of imprecision as a function of bias shows the maximum allowable combination of the two errors.*

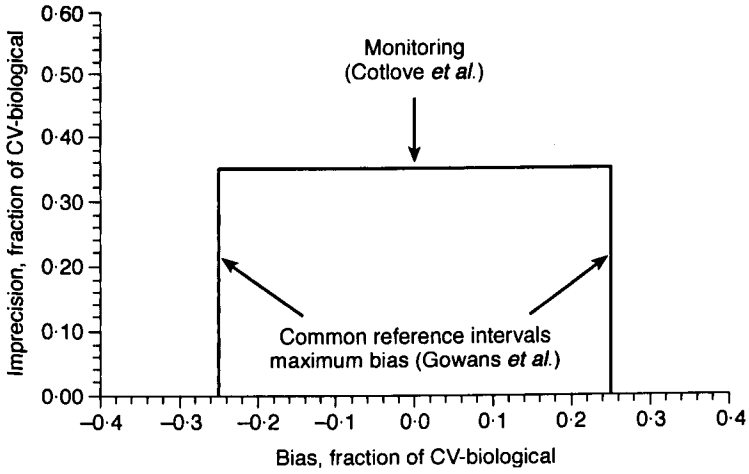

Figure 2. *The EGE Lab Working Group concept.*[3] *Allowable imprecision and bias are defined independently, resulting in a rectangular figure.*

This combination of B and $\%CV_A$ (both as variables for the purpose) is illustrated in Fig. 1, with bias (positive as well as negative) as abscissa and imprecision as ordinate. This figure illustrates the limits for maximum combination of imprecision and bias according to the maximum limit for a reference limit plus the limit of the 90% confidence interval. The plot is symmetrical since a gaussian curve is symmetrical. Combined errors inside the area bound by the lines are acceptable, whereas combined errors outside the area bound by the lines are unacceptable for the purpose of sharing common reference intervals. It should be recognized that these would be as good as if each laboratory established its own reference interval according to IFCC guidelines. The figure presented here is different to that originally published[2] because that showed only the numerical bias. Moreover, the axes were reversed. However, this type of figure has been used previously; for example, in the Nordic project on plasma proteins, such a figure was used for validation of laboratories in EQA.[12]

The EGE Lab Working Group quality specifications

In the EGE Lab Working Group recommendations,[3] maximum allowable imprecision was defined exactly as per the model of Cotlove *et al.*,[1] whereas the maximum allowable bias was defined rather simply (and possibly naively) as the allowable bias in the concept of Gowans *et al.*,[2] in which precision was zero. Thus, the two quality specifications are completely independent of each other. This can be illustrated as shown in Fig. 2. The ordinate is not shown as $\%CV_{W\text{-}S}$ but as a fraction of group biological variation $[(\%CV^2_{W\text{-}S}+\%CV^2_{B\text{-}S})^{1/2}]$ in order to make the figure comparable to the other figures presented here. If $\%CV_{W\text{-}S}$ is $0{\cdot}7\times(\%CV^2_{W\text{-}S}+\%CV^2_{B\text{-}S})^{1/2}$, then the goal of $\%CV_A < 0{\cdot}5\times\%CV_{W\text{-}S}$ corresponds to $(0{\cdot}5\times0{\cdot}7=)$ $0{\cdot}35\times(\%CV^2_{W\text{-}S}+\%CV^2_{B\text{-}S})^{1/2}$. No illustrative figures were presented in the EGE Lab Working Group recommendations.[3]

The European EQA Organizers Working Group concept

The European EQA Organizers Working Group concept[4] is rather more complex, since both within-subject and group biological variation are combined. In retrospect, we now know that the figures in the original paper were not completely correct; they should have been drawn combining Figs 1 and 2 here, as illustrated in Fig. 3. In this figure, imprecision and

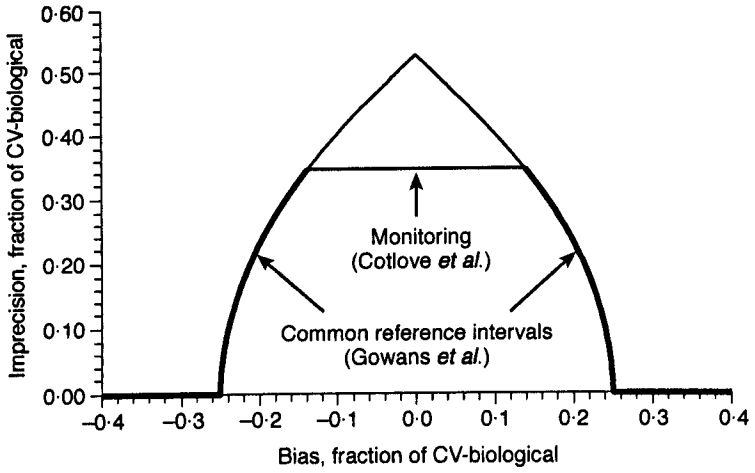

Figure 3. *The EQA Organizers Working Group concept.[4] The maximum allowable imprecision is defined according to Cotlove* et al.,*[1] whereas the maximum allowable combination of imprecision and bias is defined according to the Gowans* et al. *concept.[2] When these two models are combined, the top is cut from Fig. 1.*

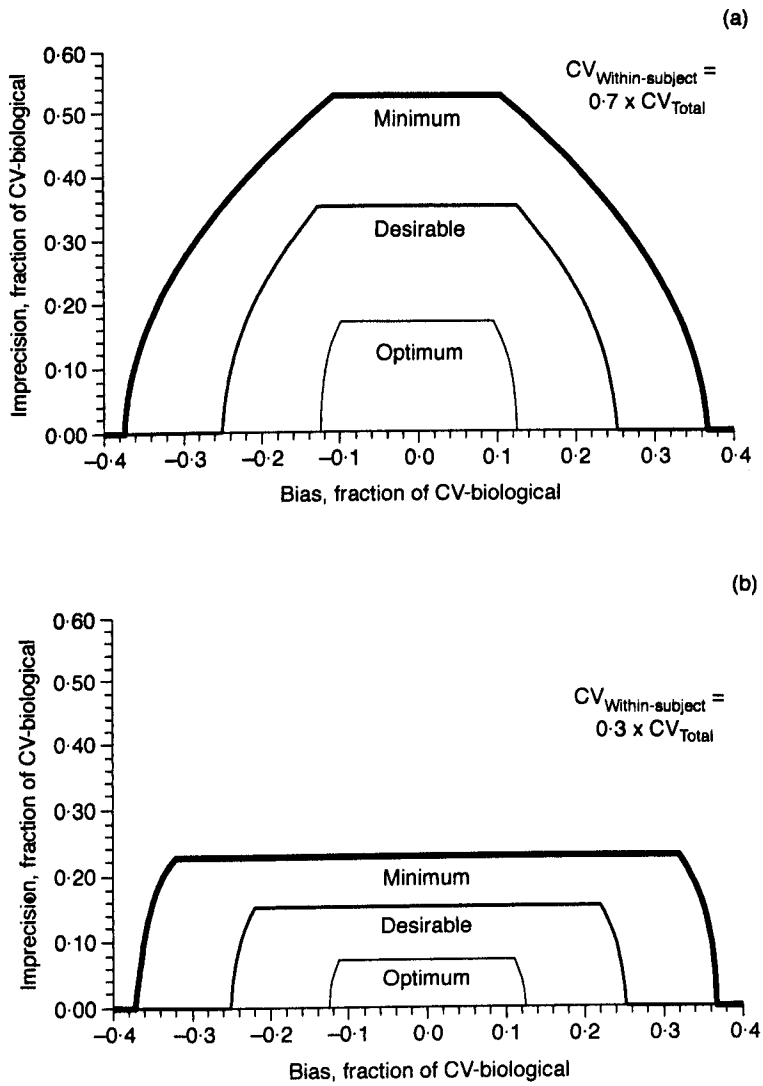

Figure 4. *The three-level concept.[5] Three levels are defined for the classification of analytical quality: optimum quality, which classifies a quality that has a negligible effect on the outcome; desirable quality; and minimum quality. See text for explanation. (a)* $\%CV_{W\text{-}S} = 0{\cdot}7 \times \%CV_{total}$. *(b)* $\%CV_{W\text{-}S} = 0{\cdot}3 \times \%CV_{total}$ *(where %CV = percentage coefficient of variation, W-S denotes within-subject).*

bias combined (from Fig. 1) are illustrated with the imprecision line from Fig. 2. The quality specification for imprecision is now delineated by the horizontal line. This concept is not as simple as the EGE Lab Working Group concept,[3] but it combines the basic models for setting quality specifications in a logical way according to the intentions of the EQA Organizers Working Group.

The three-level model

This model[5] defines three levels of quality specification for precision ($\%CV_A \leqslant a \times \%CV_{W\text{-}S}$) and for bias ($B \leqslant b \times (\%CV^2_{W\text{-}S} + \%CV^2_{B\text{-}S})$, where a and b are, respectively, 0·25 and 0·125 for optimum quality, 0·5 and 0·25 for desirable quality and 0·75 and 0·375 for minimum quality. This is illustrated in Fig. 4, in which each of the three quality specifications have the same shape as that drawn in Fig. 3, but are now represented as three different curves for the concept of sharing common reference intervals, cut by the corresponding limits for imprecision according to the monitoring concept ($a \times \%CV_{W\text{-}S}$). In Fig. 4a, $\%CV_{W\text{-}S}$ is arbitrarily set equal to $0{\cdot}7 \times (\%CV^2_{W\text{-}S} + \%CV^2_{B\text{-}S})^{1/2}$, as an example, corresponding to $0{\cdot}75 \times \%CV_{W\text{-}S}$, $0{\cdot}5 \times \%CV_{W\text{-}S}$ and $0{\cdot}25 \times \%CV_{W\text{-}S}$, which correspond to 0·525, 0·35 and 0·175 times $(\%CV^2_{W\text{-}S} + \%CV^2_{B\text{-}S})^{1/2}$, respectively. In Fig. 4b the factor is set to 0·3.

Relations to the total error concept

In the formula $TE \leqslant |\pm[1{\cdot}65 \times (0{\cdot}5 \times \%CV_{W\text{-}S}) + 0{\cdot}25 \times (\%CV^2_{W\text{-}S} + \%CV^2_{B\text{-}S})^{1/2}]|$, a concept has been used[5] with $\%CV_{W\text{-}S}$ and $(\%CV^2_{W\text{-}S} + \%CV^2_{B\text{-}S})^{1/2}$ as constants that can be readily derived from published data on biological variation.[13] Thus, TE is constant for a certain quantity, except for the z-score used (here 1·65) defining the probability (here 5%), for which errors will exceed the value of TE when imprecision and bias are both at their maximum, as described for difference plots by Hyltoft Petersen *et al.*[14] When this TE is defined as a constant and used in some EQA schemes,[6] it may be used for defining other combinations of allowable imprecision and bias within this value.

When TE is broken down into its components using the formula used by Westgard *et al.*,[8] the two variables can be combined according to the formula $TE \times a = B + z \times RE \times s$, where s is the stable (inherent) imprecision and RE is the error increasing the real imprecision; for example, if RE = 2, then the actual imprecision is twice the stable imprecision. Thus, when TE and s are constant, the formula $TE = B + z \times RE \times s$ describes the linear relationship between TE and RE. In formulae used by Westgard *et al.*,[8] B is often substituted by the term systematic error.

Table 1. *Some data on biological variation*, $\%CV_{W\text{-}S}$, $\%CV_{B\text{-}S}$ and analytical quality specifications based on biology***

Component	$\%CV_{W\text{-}S}$	$\%CV_{B\text{-}S}$	Allowable $\%CV_A$	Allowable bias (%)
Metabolites				
Plasma creatinine	4	13	2	4
Plasma uric acid	9	17	5	5
Plasma urea	12	18	6	5
Electrolytes				
Plasma calcium	2	3	1	1
Plasma sodium	1	1	1	1
Plasma magnesium	4	6	2	2
Plasma potassium	5	6	2	2
Enzymes				
Plasma alkaline phosphatase	6	25	3	6
Plasma lactate dehydrogenase	7	15	4	4
Plasma creatine kinase	23	40	12	12
Haematology				
Blood haemoglobin	3	7	4	2
Blood leucocytes	11	20	6	6
Blood thrombocytes	9	22	5	6
Urine components				
Urinary albumin (morning)	36	55	18	16
Urinary creatinine	24	25	12	9

*From http://www.westgard.com/biodatabase/htm, http://www.westgard.com/guest21.htm and ref. 12.

**According to EGE Lab recommendations.[3] $\%CV_{W\text{-}S}$ = within-subject percentage coefficient of variation; $\%CV_{B\text{-}S}$ = between-subject percentage coefficient of variation; $\%CV_A$ = analytical percentage coefficient of variation.

Examples of analytical quality specifications

Table 1 lists a number of biological %CVs, together with the calculated allowable $\%CV_A$ and allowable bias (%). It is clear from the table that the electrolytes are physiologically well regulated with accordingly low %CV values and, thereby, analytical quality specifications, whereas urinary components are widely variable, with corresponding high %CV values. It can also be seen that $\%CV_{W\text{-}S}$ is generally about half the value of $\%CV_{B\text{-}S}$ or less.

Plasma creatinine is used as an example to translate the theoretical figures into control data to validate the control result. For the control material, the target value for creatinine is 125 μmol/L. Thus the calculations from %CV values (*see* Table 1) are 4% of 125 μmol/L = 5 μmol/L and $(4^2+13^2)^{1/2}$ μmol/L = 17 μmol/L. The quality specifications according to the three-level model are shown in Fig. 5, together with the result of ten replicates of the control with a mean of 129·7 μmol/L (bias = +47 μmol/L) and a standard deviation of 3·5 μmol/L. The combined bias and imprecision are shown in Fig. 5 with 95% confidence intervals.

Discussion

It is well known that there are different models for the generation of analytical quality specifications, as demonstrated by the report from a conference in Stockholm in 1999 on 'Strategies to Set Global Analytical Quality Specifications in Laboratory Medicine'.[15] Here, the highest level in a hierarchical structure is the evaluation of analytical quality specifications based on clinical strategies/outcome, followed by analytical quality specifications based on biology. The specifications based on clinical situations are different according to the purpose, but, even when these specifications are based on biology, different models are available and may be combined according to varying assumptions, as demonstrated in this article. The main recommendations, however, agree on the combination of the concepts of Cotlove *et al.*[1] for imprecision and of Gowans *et al.*[2] for bias, but with small differences in application, as is clearly seen from Figs 2 and 3. The results of combining these concepts are easy to follow from Fig. 1, in which only the idea of sharing common reference intervals is investigated and no specific assumptions in relation to monitoring are applied. In the EGE Lab Working Group concept,[3] the combination of monitoring with the common reference interval concept is clearly simplified, resulting in a rectangular shape (Fig. 2), where the bias specifications are identical in Figs 1 and 2 for imprecision equal to zero. In Fig. 3 (the European EQA Organizers Working Group concept[4]) the area from Fig. 1 is cut by the horizontal 'monitoring' line from Fig. 2, reducing the imprecision specification of the Gowans *et al.*[2] concept to the Cotlove *et al.*[1] specification. As shown for the three-level quality specification[5] concept in Fig. 4, the specifications are highly dependent on the magnitude of the within-subject biological variation as a fraction of the group biological variation. In Fig. 4a, the within- and between-variations are of the same magnitude, as can be seen, for example, for the within- and between-coefficients for plasma sodium and potassium.[13]

In Fig. 4b the within-subject variation is about one-third that of the total, as exemplified for serum creatinine.[13] It is evident that, when within-subject biological variation is small compared to the total (and between as well), then the plot (Fig. 4b) gets closer to the EGE Lab Working Group concept[3] (Fig. 2). Likewise, when the within-subject biological variation approaches the population-based variation, the EQA Organizers Working Group concept (Fig. 4a) gets closer to the Gowans *et al.* concept[2] (*see* Fig. 1).

One of the advantages of the biological approach to defining analytical quality specifications is that data

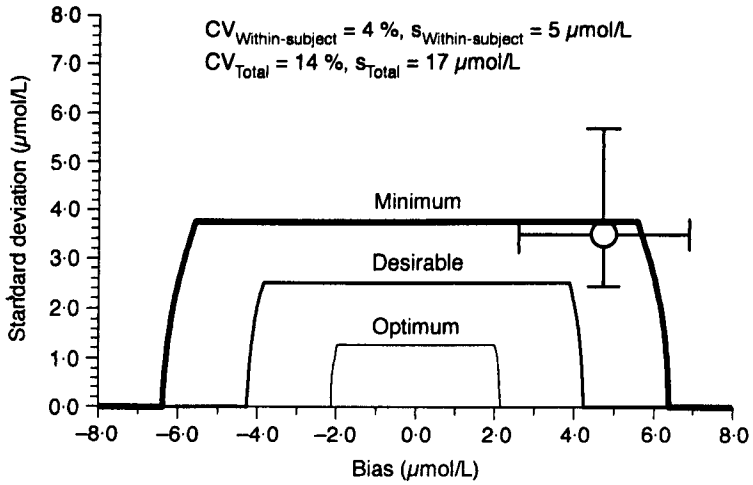

Figure 5. *Illustration of the concept in external control of plasma creatinine with* $\%CV_{W\text{-}S} = 4$ *and* $\%CV_{B\text{-}S} = 13$ *(*$\%CV_{total} = 14$*) (where %CV = percen-percentage coefficient of variation, W-S denotes within-subject, B-S denotes between-subject). For a control sample with a target value of 125 μmol/L, these variables correspond to standard deviations of 5 μmol/L and 16 μmol/L (*$s_{total} = 17$ *μmol/L), respectively. The three levels of quality are shown, together with the result of ten replicates of the control with a mean of 129·7 μmol/L (bias+4·7 μmol/L) and standard deviation of 3·5 μmol/L. The combined bias and imprecision is shown with 95% confidence intervals.*

are available from the literature[13] and easily applicable to the model, as seen in Table 1 and illustrated for external control of creatinine in Fig. 5 with a direct validation according to the analytical quality specifications.

The effects of analytical bias and imprecision are fundamentally different, as bias relates to the calculated mean (of a considerable number of replicates) minus a true value, whereas imprecision is a measure of random error, expressed as a standard deviation or coefficient of variation. In principle, a bias can be corrected when a reference method or reference preparation is available, whereas the random error cannot be corrected; although, by measuring in replicate and calculating the mean value, the variation of this mean can be reduced. Some external control strategies include replicate measurements and thus are able to give estimates of both bias and imprecision; the model for validation of control data in an external survey illustrated in Fig. 5 has been used for plasma proteins.[16] In control schemes (e.g. proficiency testing) in which only single measurements are performed, the concept of total error is applied. Use of total error is often used in quality management systems describing a maximum error for single determinations of quality control materials based on known and previously defined acceptance limits for imprecision and bias. The formula $TE < \pm [1{\cdot}65 \times (0{\cdot}5 \times \%CV_{W\text{-}S}) + 0{\cdot}25 \times (\%CV^2_{W\text{-}S} + \%CV^2_{B\text{-}S})^{1/2}]$ for total error, based on the biological variation data and the EGE Lab Working Group concept[3] with $\%CV_{W\text{-}S}$ and $(\%CV^2_{W\text{-}S} + \%CV^2_{B\text{-}S})^{1/2}$, as constants is an attempt to define total error based on biology.[7] The splitting of TE into bias (or SE) and RE (increased imprecision) is another use of the linear function of $TE = B + z \times RE \times s$. The total error approach, however, honours the use of the test results in various medical situations, whereas the biological variation approach provides an independent framework of well-defined analytical quality regardless of the medical situation in which the test result is applied.

Conclusion

This overview illustrates the well-documented but often not appreciated fact that varying assumptions and different models will yield different results. What the figures clearly show is that it is possible to obtain a rough impression of how serious a specific assumption or model will influence the results compared with other assumptions and models.

It is hoped that this overview will be helpful in the interpretation of our own as well as published data within the field of analytical quality specifications based on biology.

Acknowledgements

We thank Linda Thienpont and Dietmar Stöckl (Laboratorium voor Analytische Chemie, Faculteit van de Farmaceutische Wetenschappen, Universiteit Gent, Gent, Belgium) and Ole Blaabjerg (Department of Clinical Biochemistry, Odense University Hospital, Odense, Denmark) for fruitful discussions.

References

1 Cotlove E, Harris EK, Williams GZ. Biological and analytical components of variation in long term studies of serum constituents in normal subjects. III. Physiological and medical implications. *Clin Chem* 1970; **16**: 1028–32

2 Gowans EMS, Hyltoft Petersen P, Blaabjerg O, Hørder M. Analytical goals for acceptance of reference intervals for laboratories throughout a geographical area. *Scand J Clin Lab Invest* 1988; **48**: 757–64

3 Fraser CG, Hyltoft Petersen P, Ricós C, Haeckel R. Proposed quality specifications for the imprecision and inaccuracy of analytical systems for clinical chemistry. *Eur Clin Chem Clin Biochem* 1992; **30**: 311–7

4 Stöckl D, Baadenhuijsen H, Fraser CG, Libeer J-C, Hyltoft Petersen P, Ricós C. Desirable routine analytical goals for quantities assayed in serum. Discussion paper from the members of the external quality assessment (EQA) working group A on analytical goals in laboratory medicine. *Eur J Clin Chem Clin Biochem* 1995; **33**: 157–69

5 Libeer J-C, Baadenhuijsen H, Fraser CG, Hyltoft Petersen P, Ricós C, Stöckl D, *et al.* Characterization and classification of external quality assessment schemes (EQA) according to objectives such as evaluation of method and participant bias and standard deviation. Discussion paper from the members of the external quality assessment (EQA) working group A on analytical goals in laboratory medicine. *Eur J Clin Chem Clin Biochem* 1996; **34**: 665–78

6 Fraser CG, Hyltoft Petersen P, Libeer J-C, Ricós C. Proposals for setting generally applicable quality goals solely on biology. *Ann Clin Biochem* 1997; **34**: 8–12

7 Fraser CG, Hyltoft Petersen P. Quality goals in external quality assessment are best based on biology. *Scand J Clin Lab Invest* 1993; **53 (Suppl 212)**: 8–9

8 Westgard JO, Carey RN, Wold S. Criteria for judging precision and accuracy in method development and evaluation. *Clin Chem* 1974; **20**: 825–33

9 Westgard JO, Groth T, de Verdier C-H. Principles for developing improved quality control procedures. *Scand J Clin Lab Invest* 1984; **44 (Suppl 172)**: 19–41

10 Solberg HE. Approved recommendations (1987) on the theory of reference values. Part 5. Statistical treatment of collected reference values: determination of reference limits. *Eur J Clin Chem Clin Biochem* 1987; **25**: 645–56

11 Ricós C, Baadenhuijsen H, Libeer J-C, Hyltoft Petersen P, Stöckl D, Thienpont L, *et al.* External quality assessment: currently used criteria for evaluating performance in European countries, and criteria for future harmonization. *Eur J Clin Chem Clin Biochem* 1996; **34**: 159–65

12 Hyltoft Petersen P, Blaabjerg O, Irjala K, eds. Assessing quality in measurements of plasma proteins. In: *The Nordic Protein Project and Related Projects.* Helsinki: NORDKEM (Nordic Clinical Chemistry Project), 1994

13 Ricós C, Alvarez V, Cava F, Garcia-Lario JV, Hernandéz A, Jiménez CV, *et al.* Current databases on biological variation: pro, conc and progress. *Scand J Clin Lab Invest* 1999; **59**: 491–500

14 Hyltoft Petersen P, Stöckl D, Blaabjerg O, Pedersen B, Birkemose E, Thienpont L, *et al.* Graphical interpretation of analytical data from comparison of a field method with a reference method by use of difference plots. *Clin Chem* 1997; **43**: 2039–46

15 Hyltoft Petersen P, Fraser CG, Kallner A, Kenny D, eds. Strategies to set global analytical quality specifications in laboratory medicine. *Scand J Clin Lab Invest* 1999; **59**: 475–585

16 Hyltoft Petersen P, Blaabjerg O, Irjala K, eds. Assessing quality in measurements of plasma proteins. The Nordic Protein Project and Related Projects. *Upsala J Med Sci* 1994; **99**: 195–389

Accepted for publication 26 July 2002