

Opinion Paper

Wytze P. Oosterhuis* and Sverre Sandberg

Proposal for the modification of the conventional model for establishing performance specifications

DOI 10.1515/cclm-2014-1146

Received November 21, 2014; accepted March 24, 2015; previously published online April 22, 2015

Abstract: Appropriate quality of test results is fundamental to the work of the medical laboratory. How to define the level of quality needed is a question that has been subject to much debate. Quality specifications have been defined based on criteria derived from the clinical applicability, validity of reference limits and reference change values, state-of-the-art performance, and other criteria, depending on the clinical application or technical characteristics of the measurement. Quality specifications are often expressed as the total error allowable (TE_A) – the total amount of error that is medically, administratively, or legally acceptable. Following the TE_A concept, bias and imprecision are combined into one number representing the “maximum allowable” error in the result. The commonly accepted method for calculation of the allowable error based on biological variation might, however, have room for improvement. In the present paper, we discuss common theories on the determination of quality specifications. A model is presented that combines the state-of-the-art with biological variation for the calculation of performance specifications. The validity of reference limits and reference change values are central to this model. The model applies to almost any test if biological variation can be defined. A pragmatic method for the design of internal quality control is presented.

Keywords: allowable error; analytical variation; biological variation; quality specifications; total error.

Introduction

A good quality of test results is fundamental to the work of the medical laboratory. How to define the level of quality needed is a question that has been subject to much debate, and more than one consensus agreement has been reached to define quality specifications [1–4]. Quality specifications have been defined on the basis of criteria derived from the clinical applicability, validity of reference limits and reference change values, state-of-the-art performance, and other criteria, depending on the application and the characteristics of the test.

Quality specifications are often expressed as the total error allowable (TE_A) – the total amount of error that is medically, administratively, or legally acceptable. Following the TE_A concept, bias and imprecision are combined into one number representing the “allowable” error in the result. Internal quality control (IQC) procedures, as well as external quality assessment (EQA) can be shaped according to the TE_A of the analyte [5]. The Six Sigma concept is also linked to TE_A , as the sigma value is derived from this entity [5, 6]. “Sigma-metrics” are valuable to “normalize” quality to a common scale.

The commonly accepted method for calculation of the allowable error based on biological variation might, however, have room for improvement. The addition of the bias and imprecision terms according to this method has been shown to overestimate TE_A [7]. In the present paper, we discuss the common theories on the determination of quality specifications. A modified model for the calculation of quality specifications is presented. The validity of reference limits and reference change values are central to this model that is a modification of existing models (further called the “modified model”) [8–10].

Analytical quality specifications

According to the new consensus [1], quality specifications for clinical applicability should preferably be based on

*Corresponding author: Wytze P. Oosterhuis, Atrium-Orbis, Department of Clinical Chemistry and Haematology, Henri Dunant straat 5 Heerlen 6419PC, The Netherlands, E-mail: woosterhuis@atriummc.nl

Sverre Sandberg: Norwegian Quality Improvement of Primary Care Laboratories (Noklus), Institute of Global Health and Primary Health Care, University of Bergen and Laboratory of Clinical Biochemistry Haukeland University Hospital, Bergen, Norway

clinical outcome, which for most situations is the same as specifications based on decision levels and evaluation of tolerable false-positive (and false-negative) results. It was, however, acknowledged that, in many cases, this is difficult or even impossible to achieve as for most analytes decision levels cannot be unambiguously defined. According to the consensus document, the preferred alternative approach for these measurands is to derive specifications from biological variation. It should be noted that there is a fundamental difference between these methods, as the specifications based on biological variation are not related to clinical needs but tries to minimize the “signal-to-noise” ratio between analytical variation and the (natural) biological variation.

Analytical performance specifications based on biological variation are now broadly used in clinical chemistry, whereas in clinical guidelines, performance based on clinical criteria is, in most cases, preferred [11, 12].

Cotlove et al. [13] proposed that the tolerable analytical variation (analytical standard deviation) should be less than half of the total biological variation:

$$CV_A < 0.5CV_B \quad (CV_B = \text{coefficient of variation, total biological}). \quad (1)$$

Harris [14] then proposed that quality specifications for individual monitoring should be calculated using the formula

$$CV_A < 0.5CV_I \quad (CV_I = \text{coefficient of variation, within subject}). \quad (2)$$

For medical diagnosis, the total biological variation was used. In the case of monitoring to detect trends in the results from an individual over a period of time, the within-subject CV_I is used instead of the total biological variation. This strategy was adopted by the College of American Pathologists at the 1976 Aspen Conference [3] and by the Subcommittee on Analytical Goals in Clinical Chemistry of the World Association of Societies of Pathology in London in 1978 [4, 8].

Definition of bias and imprecision

Mathematically, analytical bias is clearly different from imprecision. Bias or systematic measurement error is defined as an error that in replicate measurements remains constant or varies in a predictable manner [15]. However, it is also stated that, “Systematic measurement error, and its causes, can be known or unknown. A correction can be applied to compensate for a known systematic measurement error.” In practice, the distinction between bias and

imprecision is, however, less clear. “Systematic” implies a certain time period. As Klee [16] pointed out, bias tends to be dependent on the time interval considered. In this paper, bias is used for the net shift in test values, relative to the set point of the assay when the reference data on patients were collected. Although bias should be removed when possible, in some circumstances, bias is inevitably encountered, such as systematic differences between analyzers measuring the same analyte.

Total error concept

The total error (TE) is an expression of the total deviation of the test result from the true value. Westgard et al. [17] presented this TE concept using the argument that physicians think rather in terms of the total analytical error, which includes both random and systematic components.

The TE limits are defined by a maximum percentage of test results, generally taken as 5%, that exceeds this limit (one-sided). For example, assume the true value of a plasma glucose measurement is 9 mmol/L and assume the TE, calculated from actual bias and imprecision, is 10%. In that case, there is up to a 5% chance that this actual result will exceed the TE limit. This means that the probability that the true result will be <8.1 mmol/L or will be >9.9 mmol/L will each be 5%. Whether this result meets the quality criteria depends on the specification of the quality limits (see below).

The basic expression most generally used for calculation of the TE is [17]

$$TE = \text{bias} + Z \times CV_A. \quad (3)$$

The Z-value is generally taken as 1.65 (95% one-sided) (Note 1).

Total error allowable (TE_A)

As shown above, TE can be calculated from actual bias and imprecision when these are known. However, when a limit for TE is predefined (the total error allowable, TE_A), the maximum allowable bias and imprecision can be derived for the acceptable analytical performance. Medical decision levels should be specified, at which concentration the performance of a method is critical. Just as an example, one decision level for glucose could be at 2.8 mmol/L with a TE_A of 20% (0.56 mmol/L). The maximum allowable bias and imprecision can then be calculated using

$$TE_A = \text{bias} + Z \times CV_A. \quad (4)$$

Many combinations of bias and imprecision can meet the limit set by the value selected for TE_A . In our example, the extreme values are $\text{bias}=20\%$ (with $CV_A=0$) and $CV_A=(20/1.65)=12.1\%$. Bias and imprecision have a linear (inverse) relationship: a higher bias requires a low imprecision, a high imprecision a low bias. As will be shown below, this model is valid only when biological variation does not play a role.

Quality specifications based on biological variation

Quality specifications should preferably be based on clinical outcome. The specification of TE_A is not always as straightforward as mentioned before. An alternative is to derive these specifications from biological variation [1].

How can this be achieved? Allowable imprecision and bias had been defined as follows:

- Allowable imprecision $CV_A < 0.5CV_I$ [14].
- Allowable bias $< 0.25(CV_I^2 + CV_G^2)^{1/2}$ [8].

In the case of EQA, these specifications should be fulfilled separately and EQA schemes could be designed accordingly. Fraser and Hytloft Petersen [18] proposed that in case only a single determination of each survey material is used or allowed, the 95% acceptance range for each laboratory from the target value was proposed to be the sum of both values:

$$\begin{aligned} 95\% \text{ acceptance range} = \\ \text{target value} \pm [1.65(0.5CV_I) + 0.25(CV_I^2 + CV_G^2)^{1/2}]. \end{aligned}$$

In terms of the TE_A :

$$TE_A = 1.65(0.5CV_I) + 0.25(CV_I^2 + CV_G^2)^{1/2}. \quad (5)$$

Two flaws in the conventional model

Although the purpose of expression (5) was the application in EQA, it is commonly accepted and used for other purposes outside EQA, such as the identification of appropriate limits for IQC [5, 19].

As shown above, the quality specification for imprecision is, in general, $CV_A < 0.5CV_B$. In the case of diagnosis, this can be written as $CV_A < 0.5(CV_I^2 + CV_G^2)^{1/2}$. In case of

monitoring, only the within-subject variation is included: $CV_A < 0.5CV_I$.

The maximum allowable bias was derived as $0.25CV_B$ or $0.25(CV_I^2 + CV_G^2)^{1/2}$ [8]. It should be noticed, however, that in the conventional model, this bias term is applied in the case of *monitoring* although this expression had been derived from a reference value model and only applies to *diagnosis*. For that reason, in the case of monitoring, we applied in the present study the reference change value model that is only based on CV_I and not on CV_G [9, 10, 20].

Secondly, it has been a pragmatic solution proposed for the use in EQA to add both maxima of allowable bias and imprecision to obtain TE_A as in Eq. (5). The theoretical basis for this is, however, lacking, as two “maximum” errors are added, each allowing 5% of the test results exceeding the limit, and only valid under the mutual exclusive assumptions of zero bias and zero imprecision, respectively. The sum will allow an increase of the percentage of test results exceeding the predefined limits [7].

What could be a rational and correct alternative to combine the effects of bias and imprecision on patient test results?

Theoretical models for quality specifications

Several models have been developed to derive maximal bias and imprecision based on reference values and the maximum number of false positives [8, 16]. The model presented here is the model according to Gowans et al. [8].

Model of Gowans (Appendix 1)

The model of Gowans et al. [8] (here referred to as the model of Gowans) is based on the influence of bias and imprecision on the proportion of results outside reference limits. Performance specifications were derived from the maximum number of results outside the reference limits. In the model of Gowans, bias and imprecision are combined into one model. The influence on the false-positive rate is calculated based on a Gaussian distribution.

Owing to the effects of bias, imprecision, or a combination of both, more cases will be outside the reference limits. Instead of the usual 2.5% outside a reference limit at 1.96 SD, a maximum of 4.6% (based on the IFCC guideline on reference values [21]) outside the same limits was

assumed to be acceptable. Thus, Gowans' model allows a maximum increase of 84% (Note 2) in false-positive results.

Following Gowans' model, a curve can be calculated that defines the maximum bias and imprecision with 4.6% of results outside the reference limit. The maximum bias and imprecision at the extreme ends of this curve are (see Appendix 1) discussed below.

The maximum allowable error was calculated as follows:

- Maximum bias (when imprecision=0)= $0.275CV_B$.
- Maximum imprecision (when bias=0)= $0.597CV_B$.

where CV_B is the total biological variation (in this model, not further specified with respect to CV_I and CV_G).

Between these two extremes, a curve describes the combination of bias and imprecision such that the condition is fulfilled: 4.6% of the results outside the reference limits (one-sided) (Figure 1) (Note 3).

According to Gowans' model, the maximum allowable bias ($0.275CV_B$) only applies when CV_A is minimal (the hypothetical situation with $CV_A=0$). On the other hand, when bias is minimal (bias=0), the allowed CV_A

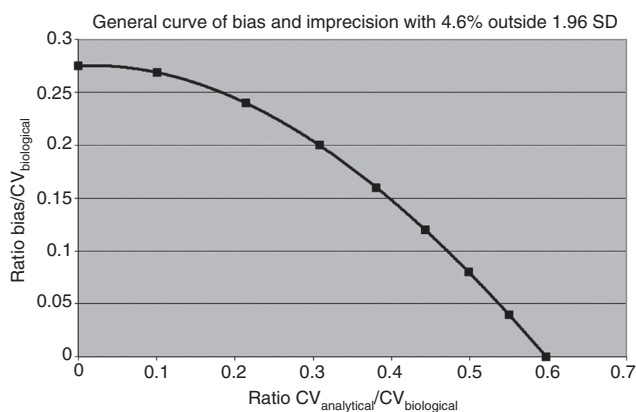


Figure 1: Model according to Gowans et al. [8]. Curved relation between bias and imprecision that describes the combinations of bias and imprecision with 4.6% of the results outside the upper reference limit.

is at a maximum ($CV_A=0.597CV_B$). This concept is clearly not in accordance with the model proposed by Fraser and Hyltoft Petersen, mentioned above, where maximum bias and maximum imprecision are summed in the expression of TE_A .

It is important to note that TE_A in Gowans' model is not a constant but varies from $0.275CV_B$ (at $CV_A=0$) to 1.65 ($0.597CV_B$) (at bias=0, with Z-factor=1.65). See also Appendix 1: Gowans for calculation of maximum bias and imprecision of creatine kinase (CK) and sodium, and Table 1.

Why was analytical variation not included in the definition of performance specifications?

The starting point of the models of Gowans and others like the model of Klee [16] is the validity of reference limits. The analytical performance specifications of the tests are derived from this concept. Gowans' model has the same assumption as the model of Klee: both define the reference limits without taking the analytical variation into account. This is clearly not the situation in common practice, as reference limits include analytical variation.

Why was this definition of reference limits used without inclusion of the analytical variation? The reason can be understood from the paper itself [8]: by including the analytical variation in the reference interval, the performance specification for analytical variation will, in part, be determined by the analytical variation itself.

Different applications of quality standards

How to solve the problem of defining performance specifications without this circular argument involved in defining, applying, and controlling analytical quality?

Table 1: Quality specifications with maximum bias (at $CV_A=0$) and maximum imprecision (at bias=0) based on different models for CK and Na.

Model	CK		Na	
	Bias, %	Imprecision, %	Bias, %	Imprecision, %
Conventional (monitoring) [22]	30.3	18.4	0.73	0.44
Gowans (diagnosis)	12.7	27.2	0.25	0.54
Modified (diagnosis)	12.7	27.1	0.39	1.34
Modified (monitoring)	9.0	13.7	0.28	1.30

Three different quality objectives should be separated at this point. The *first* is the achievement of the minimum analytical performance needed for clinical use of a test – answering the question: is the quality of the assay acceptable for routine use? This is a concept of the clinical utility of a test. For this decision, different criteria can be applied here, as covered by consensus-based quality specifications, e.g., $CV_A < 0.5 CV_I$.

The *second* quality objective is the achievement of the minimum analytical performance to maintain the validity of the reference limits (or, in the case of monitoring, reference change values), corresponding to the situation at the time the test was taken into use.

Third is the inclusion of IQC into the concept. To maintain the minimum analytical performance, some extra quality margin is needed due to the limited sensitivity for bias and imprecision of IQC procedures (see below).

In the case of sodium, the analytical variation is generally higher than the biological variation. The reference limits are, for the greater part, determined by the analytical variation, not by the biological variation. With quality criteria based on the presented models, the test would fail these quality requirements (in theory, this could be overcome by replicate measurements). The quality could also be related to the state-of-the-art performance; however, this would mean that the required quality is not based anymore on any theoretical model, and the validity of reference intervals – according to the presented models – is not maintained.

We propose another approach, by which the quality specifications are based on the same principles but with an accurate calculation of reference limits or reference change values. For analytes with a high analytical variation relative to the biological variation, this would result in more realistic quality goals.

This approach is made independent of the criterion by which the test was approved for clinical use. In theory, even a low-quality test with a very high imprecision could be introduced as a routine test by a laboratory. This would result in reference limits and reference change values that are determined predominantly by the analytical variation. In that case, the modified model does still apply.

In the text below, quality specifications based on reference values are presented that apply in the case of diagnosis. It is acknowledged that most tests will be used for monitoring, and the model based on reference change values should be used. Mathematically, this model is very similar to the reference value model. For that reason and for reasons of readability, we refer to

Appendix 3: performance specifications for the reference change model.

The modified model

This model is an adaptation of Gowans' model (Appendix 1), and the reference change values' model [9, 10] (Appendix 3), based on the following principles:

1. The model describes the maximum bias and imprecision allowable that still maintains the validity of reference values (or reference change values in case of monitoring).
2. The reference limits are defined by both biological and analytical variation.

As in other models, a distinction is made between quality criteria for diagnosis and monitoring. For diagnosis, the CV of the reference value (CV_{ref}) is used as starting point:

$$CV_{ref} = (CV_G^2 + CV_I^2 + CV_{A0}^2)^{1/2},$$

where CV_{A0} is the CV analytical of the test at $t=0$, when the reference limits were determined or confirmed.

$$CV_G = CV(\text{group}) \text{ and } CV_I = CV(\text{within person}).$$

For monitoring, the reference change model is applied (see Appendix 3). This is an adaptation of the model as described before [9, 10]. In both cases, diagnosis and monitoring, the underlying mathematical principles are the same. The only difference is the description of the total variation.

The actual (total) variation of test results in a reference population is based on biological variation and CV_A (CV actual analytical):

$$CV_T(\text{total}) = (CV_I^2 + CV_G^2 + CV_A^2)^{1/2}.$$

As in the model of Gowans, a maximum of 4.6% of the test results outside a reference limit is considered acceptable (any other percentage will not change the principle of the model).

The consequence of including CV_{A0} in the expression is that an increase in CV_A with respect to CV_{A0} determines the quality, not the absolute value of CV_A . In this model, CV_{A0} can be within the quality specification $CV_A < 0.5 CV_I$ but does not need to be, e.g., when the state of the art does not meet this specification. In the model of Gowans, a Gaussian distribution is assumed with $CV = CV_B$ with reference limits at the point where 2.5% of the results are outside the limits. Analytical variation (or analytical variation in combination with bias) is then added to the model with a limit of 4.6% test results outside the reference

limits (in other words, 4.6% misclassification instead of 2.5%). In contrast to this, the modified model starts with a Gaussian distribution with $CV=(CV_B^2+CV_{A0}^2)^{1/2}$ to which additional analytical variation (or analytical variation in combination with bias) is then added.

With $CV_A=0$, the maximum bias is

$$\text{Bias}_{\max}=0.275(CV_B^2+CV_{A0}^2)^{1/2}.$$

With bias=0, the maximum imprecision is

$$CV_{A,\max}=((1.96/1.68)^2 \times (CV_B^2+CV_{A0}^2)-CV_B^2)^{1/2}.$$

The model can be illustrated by the examples in Figures 2 and 3 for CK and sodium.

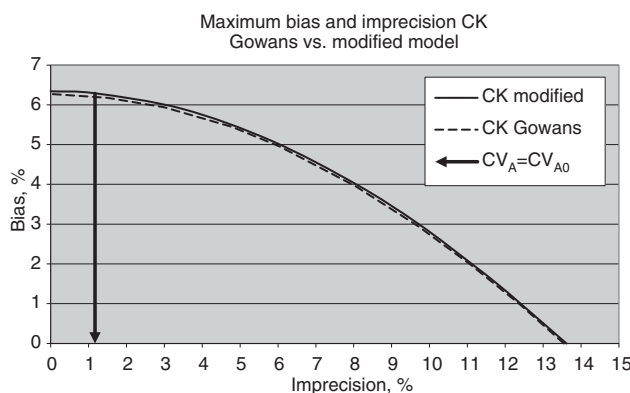


Figure 2: Curves describing the combinations of bias and imprecision for CK according to Gowans' and the modified models. The arrow indicates the position of CV_{A0} (see Appendix 2). Owing to the low value of the analytical compared to the biological variation, both models are almost identical.

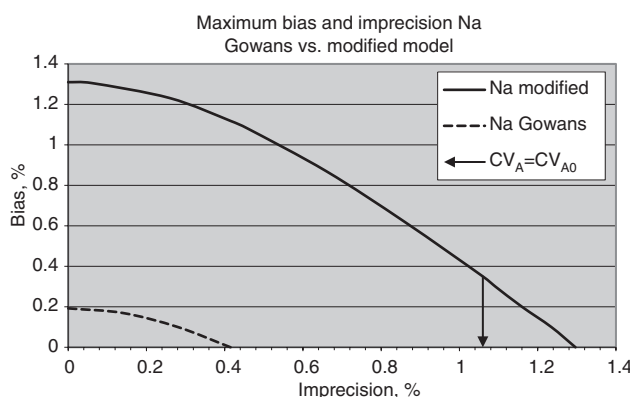


Figure 3: Curves describing the combinations of bias and imprecision for sodium according to Gowans' and the modified models. Owing to the high value of the analytical compared to the biological variation, both models differ considerably. The arrow indicates the position of CV_{A0} (see Appendix 2). The analytical variation is clearly outside the specifications of Gowans' model but inside the specifications of the modified model.

Quality control, TE_A , and the sigma concept

One of the most important uses of quality specifications, in terms of maximum allowable bias and imprecision, is the development of an IQC program. An important concept here is the TE_A . TE_A combines bias and imprecision to one fixed number. This concept is only valid when bias and imprecision show a linear relation (see Appendix 4: linear). The fixed number for TE_A also means that one does not need to know whether the deviation of a control sample result should be attributed to an increase of bias or imprecision. Only one limit TE_A for the combination of bias and imprecision suffices.

The TE_A concept is valid when IQC results are considered. The distribution of quality control results is described by the analytical variation only, as biological variation plays no role here. For these results, a TE_A limit can be defined for the combination of bias and imprecision; bias and imprecision will show a linear relation as described in Appendix 4.

A problem arises, however, when this concept is translated to patient data. As we have shown, both in the concept of reference values and of reference change values, there is no linear relation between bias and imprecision. When biological variation is taken into account, the linearity changes to a curved relation. This curved relation implies that the tolerance for additional imprecision will increase compared to the tolerance for additional bias, which will remain unchanged. (What might also be taken into consideration is the fact that the sensitivity of multirules [17] are not the same for bias and imprecision.)

Now the problem can be described more precisely: when we have an ICQ result with a certain deviation from the target value, we cannot ascribe this to an effect of bias or imprecision. In a linear model, we have shown that this is not of importance, as long as the deviation is within the TE_A limits. As the TE_A concept fails with patient results, with no linear relation between bias and imprecision, how can we decide whether this deviation is acceptable or not?

There is no accepted solution for this problem. We could, however, assume that the imprecision of the test system will remain constant. With that assumption, deviations of IQC results will be ascribed solely to the effect of bias. This pragmatic solution would result in a definition of TE_A for quality control results as

$$TE_A = \text{bias} + Z \times CV_A.$$

With (for diagnosis):

$$Z=1.65.$$

with $CV_A = CV_{A0}$ (see Appendix 2).

$$CV_{T0} = (CV_I^2 + CV_G^2 + CV_{A0}^2)^{1/2}.$$

$$\text{Bias}_{\max} = -1.68CV_{T0} + 1.96CV_{T0} = 0.275CV_{T0}.$$

$$TE_A = 0.275CV_{T0} + 1.65CV_{A0}.$$

For monitoring, see Appendix 3: modified RCV.

A test can have a certain analytical quality that is close to the limit of the quality specifications. In that case, there will be a problem with the maintenance of this level of quality. For instance, when $CV_A = 0.5CV_B$ (see Figure 1), the analytical variation is almost equal to the limit of the desired quality specification according to the criteria based on biological variation [23]. In Gowans' model, no additional bias or imprecision is allowed. This makes it almost impossible to apply quality control specifications to the test, even with the quality specification fulfilled. The modified model, however, has a considerable margin for additional bias and imprecision in the case of CK (Figure 2, vertical line), and some margin for sodium (Figure 3).

In conclusion, some extra margin of quality is needed because of the limited sensitivity of ICQ procedures. Note that within the Six Sigma theory, a margin of 1.5 SD is assumed necessary to maintain the results within the performance specifications.

Discussion

In this study, we compared several models for analytical quality specifications, including a modified model based on the calculation of reference values and reference change values, taking the analytical variation into account. Central to this model is the assumption that the validity of reference limits (for diagnosis) or reference change values (for monitoring) determines the minimum analytical quality. This concept can, of course, in itself be discussed.

A distinction is made with, on the one hand, the analytical performance needed for routine clinical use of a test, answering the question of whether a test should be taken into routine use by the laboratory. On the other hand is the analytical performance required to maintain the validity of the reference limits and reference change values, once the choice has been made to take a test into use. The second follows the first: when a test is accepted for routine use based on clinical or other criteria, reference limits are subsequently determined (or existing reference limits validated). Quality criteria can subsequently

be derived from the model that maintains the validity of the reference limits or reference change values.

Both models of Klee and Gowans [8, 16] are based on the reference value concept and do include the biological variation, but do not include analytical variation in their definition of reference limits. This can lead to unrealistic values and quality criteria for tests like sodium, where the analytical variation is the dominant component of the reference interval.

The calculation of TE_A is often based on biological variation. It has been shown that in the calculation of TE_A , the summation of both maximum allowable bias and imprecision term [expression (2)] does lack a theoretical basis [7]. In the conventional model, the total allowable error is assumed to be a constant, with an inherent linear relation between bias and imprecision. The biological variation is, however, not correctly included in this model. In the presented model, TE_A – with analytical variation included – is not anymore a constant but depends on the ratio of bias and analytical variation.

The proposed modified model can be seen as an adaptation of existing models based on reference values [8] and reference change values [9, 10]. In these models, reference limits were calculated based on biological variation alone. In the present model, reference limits are based on both biological and analytical variation. In tests with CV_B considerably larger than CV_A , CV_A can, however, be neglected in the calculation, and the model equals to the existing models. When this condition is not fulfilled, the quality goals according to the existing models will tend to be too strict. For example, for sodium, it would be almost impossible to meet the quality specifications (Figure 2). In contrast, in the modified model, both bias and imprecision do meet the quality specifications and have still some margin for increase. In the case of a less-than-perfect test, application of the modified model will lead to more realistic quality goals.

In the model of Gowans, the quality specification for misclassification was 4.6%. In the presented model, we applied this specification of 4.6% both for the case of diagnosis and for monitoring, although in the model of Gowans this specification has been derived for the validity of reference values (and thus for diagnosis) only. Another specification could, however, be applied, without changing the principles of the presented model.

On the one hand, we have quality specifications based on the validity of reference values and reference change values. On the other hand, we have the problem of maintaining this quality with quality control procedures. These procedures have a limited sensitivity for errors, and an extra quality margin is needed to be able to guarantee that results are within quality limits. This margin will depend

on the quality control procedures applied, and is not part of the models presented here.

Conclusions

We propose a modified model that offers an alternative method for the calculation of performance specifications. It is based on maintaining the validity of reference limits and reference change values. The model applies to almost any test if biological variation can be estimated. A pragmatic method for the design of IQC is presented.

Acknowledgments: We thank Dietmar Stöckl for valuable comments, as well as Henne Kleinveld, and Niels de Jonge (chair) and other members of the Six Sigma guideline working group of the Netherlands Society for Clinical Chemistry and Laboratory Medicine.

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Financial support: None declared.

Employment or leadership: None declared.

Honorarium: None declared.

Competing interests: The funding organization(s) played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

Notes

Note 1

Z-value 1.65 vs. 1.96.

When the bias is equal or larger than the analytical imprecision, the Z-value of 1.65 applies, and 95% of the test results will fall within the limits defined by the TE_A . The test result can be reported with 95% certainty to be within the limit defined by this TE_A .

The Z-factor has been shown to be dependent on the ratio of bias and imprecision and could be 1.96 (two-sided 0.95 or 95% confidence limit) when bias can be neglected compared to the imprecision. For ratios of bias and imprecision, the Z-value was calculated as 1.96 (bias/imprecision=0), 1.77 (ratio=0.25), 1.68 (ratio=0.5), 1.65 (ratio=0.75), and 1.645 (ratio $\geq$ 1) [24].

For example, when the true value of glucose is 9 mmol/L with a bias of 6% and an imprecision of 2% (ratio $>$ 1), then the bias is 9 mmol/L $\times$ 6/100 or 0.54 mmol/L

and the imprecision 9 mmol/L $\times$ 2/100 or 0.18 mmol/L. This makes TE 0.54+1.645 $\times$ 0.18 or 0.84 mmol/L greater than the actual concentration of 9.0 mmol/L, so that the TE limit is 9.84 mmol/L.

In this case, the bias is positive and substantially greater than the imprecision. The Z-value is 1.645 in this case, as only the upper limit is relevant: knowing with 95% certainty that the true glucose concentration will not exceed the 9.84 mmol/L limit. With this bias, the measured value will never be lower than the true value.

If, on the other hand, the bias is low, the Z-value of 1.96 should be applied. For example, when the true value of glucose is 9 mmol/L with a bias of 0% and an imprecision of 2%, then Z is 1.96. The imprecision component is 9 mmol/L $\times$ 1.96 $\times$ 2/100 or 0.35 mmol/L. This makes the TE 0.0+0.35 or 0.35 mmol/L higher or lower than the actual concentration of 9.0 mmol/L so that the TE limits are 8.65 and 9.35 mmol/L.

As a known bias can be corrected, the presence of a bias that is equal or larger than the imprecision of a test should, in many cases, be corrected. As a result, use of the Z-value of 1.65 in all cases could be questioned.

Note 2

Gowans' model allows a maximum increase of 84%: (4.6–2.5)/2.5 $\times$ 100%=84%.

Note 3

Note that these results are in close agreement with the accepted performance criteria [25]: for bias, compare 0.275 CV_B (CV biological) with the maximum (desirable) bias of 0.25 CV_B .

For imprecision, compare 0.597 CV_B with 0.5 CV_I . Note that Gowans' model only mentioned biological variation and did not take into account the difference between diagnosis (based on within group and within person variation) and monitoring (based only on within-person variation).

Appendix

Appendix 1. Model of Gowans et al. (Gowans)

When the recommendations of the IFCC are applied, reference values will be calculated on the basis of test results

in a group of at least 120 persons [21]. There is always an inherent uncertainty in the determination of the reference limits for every analyte. This uncertainty is such that with $n=120$, a maximum of 4.6% (one-sided) of the results could be outside the “inner” confidence limit of the reference limits when using 1.96 SD.

Following this model, a curve can be calculated that defines the maximum bias and imprecision with 4.6% of results outside the reference limit. Maximum bias and imprecision are at the extreme ends of this curve are (see Figure 1).

In general, the relation between bias and imprecision can, in this case, be described as (see Appendix 4)

$$\text{Bias} = -1.68\text{CV}_T + 1.96\text{CV}_B,$$

with $\text{CV}_T = \text{total variation} = (\text{CV}_B^2 + \text{CV}_A^2)^{1/2}$; $\text{CV}_B = \text{total biological variation}$; $\text{CV}_A = \text{analytical variation}$; 1.96 represents the Z-value with 2.5% and 1.68 the Z-value with 4.6% outside the limit.

In this expression, bias and imprecision (analytical variation) have a linear relationship with a slope of -1.68 (see Appendix 4). The intersection with the y-axis represents the bias at $\text{CV}_A = 0$:

$$\begin{aligned}\text{Bias}(\text{CV}_A = 0) &= -1.68(\text{CV}_B^2 + \text{CV}_A^2)^{1/2} + 1.96\text{CV}_B \\ &= -1.68(\text{CV}_B^2)^{1/2} + 1.96\text{CV}_B \\ &= (1.96 - 1.68)\text{CV}_B.\end{aligned}$$

$$\text{Maximum bias (when imprecision}=0)=0.275\text{CV}_B.$$

Note that this number differs from 0.25 in the study of Gowans et al. [8].

The intersection with the x-axis represents CV_A at bias=0

$$\begin{aligned}\text{CV}_A(\text{bias}=0): 0 &= -1.68(\text{CV}_B^2 + \text{CV}_A^2)^{1/2} + 1.96\text{CV}_B \\ 1.96\text{CV}_B &= 1.68(\text{CV}_B^2 + \text{CV}_A^2)^{1/2}.\end{aligned}$$

$$\text{Maximum imprecision (when bias}=0)=0.597\text{CV}_B.$$

Between these two extremes, a curve describes the combination of bias and imprecision such that the condition is fulfilled: 4.6% of the results outside the reference limits (one-sided) (Figure 1) (Note 3).

Calculation of maximum bias and imprecision of CK and sodium:

Example 1: CK

Gowans' model, with maximum imprecision and bias, respectively:

$$\text{CV}_{A,\text{max}} = 0.597\text{CV}_B.$$

$$\text{Bias}_{\text{max}} = 0.275\text{CV}_B.$$

$$\text{CV}_I = 22.8\%; \text{CV}_G = 40.0\% [22].$$

$$\text{CV}_B = (\text{CV}_I^2 + \text{CV}_G^2)^{1/2} = 46.0\%.$$

$$\text{CV}_{A,\text{max}} = 0.597\text{CV}_B = 27.2\%.$$

$$\text{Bias}_{\text{max}} = 0.275\text{CV}_B = 12.65\%.$$

Example 2: sodium

$$\text{CV}_I = 0.6\%; \text{CV}_G = 0.7\% [22].$$

$$\text{CV}_B = (\text{CV}_I^2 + \text{CV}_G^2)^{1/2} = 0.92\%.$$

$$\text{CV}_{A,\text{max}} = 0.597\text{CV}_B = 0.54\%.$$

$$\text{Bias}_{\text{max}} = 0.275\text{CV}_B = 0.25\%.$$

Note that these quality limits apply for diagnosis, not for monitoring situations.

Appendix 2. Modified model, adapted Gowans' model

Below is the modification of Gowans' model, based on the following principles:

1. The model describes the maximum bias and imprecision allowable to maintain the validity of reference values (or reference change values in case of monitoring).
2. The reference limits are defined by both biological and analytical variation.

As in other models, a distinction is made between quality criteria for diagnosis and monitoring. For diagnosis, the CV of the reference value (CV_{ref}) is used as starting point:

$$\text{CV}_{\text{ref}} = (\text{CV}_G^2 + \text{CV}_I^2 + \text{CV}_{A0}^2)^{1/2},$$

with $\text{CV}_{A0} = \text{CV analytical of the test at } t=0$, when the reference limits were determined or confirmed; $\text{CV}_A = \text{CV actual analytical}$.

For monitoring, the reference change model is applied (see Appendix 3). This is an adaptation of the model as described before [9, 10]. In both cases, diagnosis and monitoring, the general principles presented here are the same.

The actual (total) variation CV_T of test results in a reference population is based on biological variation and actual analytical variation CV_A :

$$\text{CV}_T = (\text{CV}_I^2 + \text{CV}_G^2 + \text{CV}_A^2)^{1/2}.$$

As in the model of Gowans, a maximum of 4.6% of the test results outside a reference limit is considered acceptable (any other percentage is, however, possible and will not change the principle of the model). The curve in this modified model is identical to the curve according to Gowans, with the following substitutions.

The relation between bias and imprecision can, in this case, be described as (see Appendix 4)

$$\begin{aligned}\text{Bias} &= -1.68(CV_B^2 + CV_A^2)^{1/2} + 1.96(CV_B^2 + CV_{A0}^2)^{1/2} \\ &= -1.68(CV_I^2 + CV_G^2 + CV_A^2)^{1/2} + 1.96(CV_G^2 + CV_I^2 + CV_{A0}^2)^{1/2} \\ CV_B &= (CV_I^2 + CV_G^2)^{1/2},\end{aligned}$$

where 1.96 represents the Z-value with 2.5% and 1.68 the Z-value with 4.6% outside the limit, respectively.

The consequence of including CV_{A0} in the expression is that an *increase* in CV_A with respect to CV_{A0} determines the quality, not the absolute value of CV_A .

In the model of Gowans, a Gaussian distribution is assumed with $CV = CV_B$ with reference limits at the point where 2.5% of the results are outside the limits. Analytical variation (or analytical variation in combination with bias) is then added to the model with a maximum of 4.6% test results outside the reference limits. In contrast to this, the modified model starts with a Gaussian distribution with $CV = (CV_B^2 + CV_{A0}^2)^{1/2}$ to which analytical variation (or analytical variation in combination with bias) is then added.

With bias=0, the maximum allowable imprecision should fulfill the condition

$$\text{Bias}=0 = -1.68(CV_B^2 + CV_{A,\max}^2)^{1/2} + 1.96(CV_B^2 + CV_{A0}^2)^{1/2}.$$

The maximum CV_A can be calculated from this expression to be

$$CV_{A,\max} = ((1.96/1.68)^2 \times (CV_B^2 + CV_{A0}^2) - CV_B^2)^{1/2}.$$

With $CV_A=0$, the maximum bias is

$$\text{Bias}_{\max} = 0.275(CV_B^2 + CV_{A0}^2)^{1/2}.$$

The model can be illustrated by the examples in Figures 1 and 2.

Example 1: CK

In the case that $CV_{A0} \ll CV_B$, the term CV_{A0} vanishes from the expression and the modified model equals Gowans' model (Figure 1), with maximum imprecision and bias, respectively

$$CV_{A,\max} = 0.597 CV_B.$$

$$\text{Bias}_{\max} = 0.275 CV_B.$$

Note that, in this case, the modified model will become equal to Gowans' model.

CK approximates this condition with $CV_{A0} \ll CV_B$. Numbers from the authors' laboratory:

$$CV_{A0} = 1.17\%.$$

$$CV_I = 22.8\%; CV_G = 40.0\% [22].$$

The contribution to the reference interval by the analytical variation is only 0.03%.

Note:

$$CV_I = 22.8\%; CV_G = 40.0\%.$$

$$(CV_I^2 + CV_G^2 + CV_{A0}^2)^{1/2} = (22.8^2 + 40.0^2 + 1.17^2)^{1/2} = 46.01\%.$$

$$(46.01 - 46.0) / 46 \times 100\% = 0.03\%.$$

$CV_{A,\max}$ according to the modified model:

$$(CV_{\text{biol}}^2 + CV_{A,\max}^2)^{1/2} = 1.16(CV_{\text{biol}}^2 + CV_{A0}^2)^{1/2} \rightarrow CV_{A,\max} = 27.1\%.$$

$$\text{Bias}_{\max} = 0.275(CV_{\text{biol}}^2 + CV_{A0}^2)^{1/2} = 12.65\%.$$

This illustrates that for $CV_{A0} \ll CV_B$, both models become equal.

Example 2

On the other end of the spectrum, we have $CV_{A0} \gg CV_B$, and now the term CV_B vanishes from the expression. Gowans' model does not apply to this situation, as CV_{A0} lies outside the area of the curve (outside the minimum quality limit).

Sodium approximates this condition with $CV_{A0} > CV_B$. Numbers from the authors' laboratory (see Figure 2):

$$CV_{A0} = 1.06\%.$$

$$CV_I = 0.6\%; CV_G = 0.7\% [22]; CV_B = 0.92\%.$$

The contribution to the reference interval by the biological variation is 17%.

$CV_{A,\max}$, modified model:

$$(CV_B^2 + CV_{A,\max}^2)^{1/2} = 1.16(CV_B^2 + CV_{A0}^2)^{1/2} \rightarrow CV_{A,\max} = 1.34\%.$$

$$\text{Bias}_{\max} = 0.275(CV_B^2 + CV_{A0}^2)^{1/2} = 0.39\%.$$

Appendix 3. Performance specification based on reference change values

Modified model

Below is the adaptation of the model based on reference change value [9, 10], according to the following principles:

1. The model describes the maximum bias and imprecision necessary to maintain the validity of reference change values.

2. The reference change values are defined by both biological and analytical variation.
3. By definition, reference change values only apply for monitoring, not for diagnosis.

The reference change value (RCV) is defined as

$$RCV = \sqrt{2} \times Z \times (CV_I^2 + CV_A^2)^{1/2},$$

$$CV_A = CV \text{ analytical},$$

$$CV_I = CV \text{ within person},$$

where Z is the number of standard deviations appropriate to the desired probability.

The differences between two consecutive values within one patient are described by a Gaussian curve with CV :

$$CV_{RCV} = \sqrt{2} (CV_I^2 + CV_A^2)^{1/2}.$$

With a Z -value of 1.96, 2.5% of the test results will respectively be outside the upper and lower limits. In analogy with the quality limits described before, we chose to set the quality standard at a maximum of 4.6% outside upper and lower limits, instead of 2.5%. In other words, bias and imprecision – or the combination of these – are allowed to increase until 4.6% of the differences (change values) of a reference population are outside a reference limit, resulting in a 4.6% misclassification instead of 2.5%.

Again, we substitute CV_{A0} (CV_A at the time the test was introduced or validated) for CV_A . The consequence of this is that it is the *increase* in CV_A with respect to CV_{A0} that determines the quality, not the absolute value of CV_A .

$$RCV = \sqrt{2} \times Z \times (CV_I^2 + CV_{A0}^2)^{1/2}.$$

The combinations of maximum bias and imprecision are again described with a curve.

With imprecision = CV_{A0} , there is no additional imprecision compared to CV_{T0} , and the maximum bias can be calculated (a decrease of imprecision relative to CV_{T0} is possible but not considered here):

$$\text{Bias} = 1.96 \times \sqrt{2} CV_{T0} - 1.68 \times \sqrt{2} CV_T.$$

Note: 1.96 is Z -value corresponding with 2.5%; 1.68, with 4.6% outside the quality limit.

$$\text{With } CV_{T0} = (CV_I^2 + CV_{A0}^2)^{1/2},$$

$$CV_T = (CV_I^2 + CV_A^2)^{1/2}.$$

Maximum bias is allowed when $CV_A = CV_{A0}$, and $CV_T = CV_{T0}$

$$\text{Bias}_{\max} = (1.96 - 1.68) \times \sqrt{2} CV_{T0},$$

or

$$\text{Bias}_{\max} = 0.275 \times \sqrt{2} (CV_I^2 + CV_{A0}^2)^{1/2}.$$

With bias=0, the maximum allowable imprecision should fulfill the condition:

$$\text{Bias}=0 = 1.96 \times \sqrt{2} CV_{T0} - 1.68 \times \sqrt{2} CV_T,$$

$$\text{With } CV_{T0} = (CV_I^2 + CV_{A0}^2)^{1/2},$$

$$\text{And } CV_T = \sqrt{2} (CV_I^2 + CV_A^2)^{1/2},$$

$$\text{Bias}=0 = 1.96 \times \sqrt{2} (CV_I^2 + CV_{A0}^2)^{1/2} - 1.68 \times \sqrt{2} (CV_I^2 + CV_A^2)^{1/2}.$$

The maximum CV_A can be calculated from this expression to be

$$CV_{A,\max} = ((1.96 / 1.68)^2 \times (CV_I^2 + CV_{A0}^2) - CV_I^2)^{1/2}.$$

Example 1

CK approximates this condition with $CV_{A0} < CV_B$. Numbers from the authors' laboratory:

$$CV_{A0} = 1.17\%.$$

$$CV_I = 22.8\% [22].$$

$$CV_{A,\max} \text{ according to the reference change model:}$$

$$CV_{A,\max} = ((1.96 / 1.68)^2 \times (CV_I^2 + CV_{A0}^2) - CV_I^2)^{1/2} = 13.7\%.$$

Maximum bias:

$$\text{Bias}_{\max} = 0.275 \times \sqrt{2} (CV_I^2 + CV_{A0}^2)^{1/2} = 8.97\%.$$

Example 2

Sodium approximates the condition with $CV_{A0} > CV_B$. Numbers from the authors' laboratory (see Figure 2):

$$CV_{A0} = 1.06\%.$$

$$CV_I = 0.7\% [22].$$

$$CV_{A,\max} \text{ according to the reference change model:}$$

$$CV_{A,\max} = ((1.96 / 1.68)^2 \times (CV_I^2 + CV_{A0}^2) - CV_I^2)^{1/2} = 1.3\%.$$

Maximum bias:

$$\text{Bias}_{\max} = 0.275 \times \sqrt{2} (CV_I^2 + CV_{A0}^2)^{1/2} = 0.28\%.$$

Estimated TE_A for monitoring

$$TE_A = \text{bias} + Z \times CV_A,$$

with (for monitoring)

$$Z = 1.65.$$

$$CV_A = CV_{A0}.$$

$$CV_{T0} = (CV_I^2 + CV_G^2 + CV_{A0}^2)^{1/2}.$$

$$\text{Bias} = -1.68 \times \sqrt{2} CV_{T0} + 1.96 \times \sqrt{2} CV_{T0} = 0.275 \times \sqrt{2} CV_{T0} = 0.39 CV_{T0}.$$

$$TE_A = 0.39 CV_{T0} + 1.65 CV_{A0}.$$

Appendix 4. Linear relation between bias and imprecision

Assume that test results show a normal (Gaussian) distribution with coefficient of variation CV_{T0} (total coefficient of variation at $t=0$, biological and analytical combined). We can define a limit (e.g., a quality limit) with a fixed number (percentage) of test results outside this predefined limit. The position of this limit is expressed as Z-factor (number of CV). When the percentage of test results outside the limit is 2.5% (one-sided), Z equals 1.96.

If the bias (shift of the curve) increases from zero to a positive number, the percentage of test results outside the limit $+1.96 \times CV_{T0}$ will increase, above the predefined number of 2.5%. To fulfill the condition of 2.5%, CV_T (actual total coefficient of variation) should decrease. The relation between bias and CV_T and CV_{T0} can be expressed as

$$\text{Bias} = -Z \times CV_T + Z \times CV_{T0}.$$

In the case of zero bias, the (maximum) CV will be equal to CV_B , and the condition of 2.5% outside the predefined limit is fulfilled.

In the case of zero CV_T , the (maximum) bias will be equal to $Z \times CV_{T0}$, a shift equal to the position of the predefined limit. (Strictly speaking, $CV_T=0$ outside the model, as this CV_T does not define a normal distribution. However, as CV approaches zero, the bias approaches the limit $Z \times CV_B$.)

In the case of CV_T half of CV_{T0} , the (maximum) bias will be equal to $0.5 \times Z \times CV_{T0}$, or a shift of $0.98 CV_{T0}$.

There is another situation if the distribution of results is defined by CV_{T0} , the limit at $Z=1.96$ with 2.5% of the results outside the predefined limit. A new maximum percentage can be set at 4.6%, at the same limit of $1.96 CV_{T0}$; 4.6% corresponds to a limit at $1.68 CV_{T0}$, so the corresponding Z-value is 1.68. The relation between bias and CV_T can now be expressed as

$$\text{Bias} = -Z' CV_T + Z \times CV_{T0} = -1.68 CV_T + 1.96 CV_{T0}.$$

In the case of zero bias, the (maximum) CV_T will be equal to $Z/Z' \times CV_{T0}$, or $(1.96/1.68) \times CV_{T0}$, and the condition of 4.6% outside the predefined limit is fulfilled.

In the case of zero CV_T , the (maximum) bias will be equal to $Z \times CV_{T0}$ ($1.96 CV_{T0}$), a shift equal to the position of the predefined limit. The linear relation between bias and CV only exists as the distribution is described by CV. The total variation can be composed of components of biological and analytical variation. The relation between a component (e.g., the analytical variation) and bias will not be linear, but is described by a curve (see Figure 1).

Appendix 5. Definition of pragmatic quality control limits

Here, it is assumed that the deviation of IQC results is mainly due to bias, and that the increase of the analytical variation can be neglected. For diagnosis:

$$TE_A = \text{bias} + Z \times CV_A, \quad (1)$$

with

$$Z = 1.65. \quad (2)$$

We assume that the actual CV_A is equal to the CV_A at time=0 (stable CV_A):

$$CV_A = CV_{A0}. \quad (3)$$

For diagnosis, the total variation at $t=0$:

$$CV_{T0} = (CV_I^2 + CV_G^2 + CV_{A0}^2)^{1/2}. \quad (4)$$

For bias (see Appendix 4: linear):

$$\begin{aligned} \text{Bias} = & -1.68 (CV_I^2 + CV_G^2 + CV_A^2)^{1/2} \\ & + 1.96 (CV_G^2 + CV_I^2 + CV_{A0}^2)^{1/2}. \end{aligned} \quad (5)$$

With $CV_A = CV_{A0}$ [5], this becomes

$$\text{Bias}_{\max} = -1.68 CV_{T0} + 1.96 CV_{T0} = 0.275 CV_{T0}. \quad (6)$$

Expression (1) combined with (2), (3), and (6) becomes

$$TE_A = 0.275 \times CV_{T0} + 1.65 CV_{A0}. \quad (7)$$

Compare this with the expression for monitoring (Appendix 3: RCV):

$$\text{Bias} = 1.96 \times \sqrt{2} (CV_I^2 + CV_{A0}^2)^{1/2} - 1.68 \times \sqrt{2} (CV_I^2 + CV_A^2)^{1/2}. \quad (8)$$

With $CV_A = CV_{A0}$ [5], this becomes

$$\text{Bias} = 0.275 \times \sqrt{2} (CV_I^2 + CV_{A0}^2)^{1/2}. \quad (9)$$

Expression (1) combined with (2), (3), and (6) becomes

$$TE_A = 0.275 \times \sqrt{2} (CV_I^2 + CV_{A0}^2)^{1/2} + 1.65 CV_{A0}.$$

$$TE_A = 0.39 (CV_I^2 + CV_{A0}^2)^{1/2} + 1.65 CV_{A0}. \quad (10)$$

For CK, this will give a sigma score (numbers as used before):

With $CV_{AO}=1.17\%$; $CV_I=22.8\%$.

$TE_A=10.8\%$.

Sigma score (with $CV_A=CV_{AO}$): $TE_A/CV_{AO}=9.3$.

For sodium:

With $CV_{AO}=1.06\%$; $CV_I=0.6\%$.

$TE_A=2.2\%$.

Sigma score (with $CV_A=CV_{AO}$): $TE_A/CV_{AO}=2.1$.

It can be calculated that for a sigma score of 3.0, CV_A (with $CV_A=CV_{AO}$) should be 0.18%.

References

- Sandberg S, Fraser CG, Horvath AR, Jansen R, Jones G, Oosterhuis WP, et al. Defining analytical performance specifications: consensus statement from the 1st Strategic Conference of the European Federation of Clinical Chemistry and Laboratory Medicine. *Clin Chem Lab Med* 2015;53: 833–35.
- Kenny D, Fraser CG, Hyltoft Petersen P, Kallner A. Consensus agreement. *Scand J Clin Lab Invest* 1999;59:585–6.
- Elevitch FR, editor. Analytical goals in clinical chemistry. CAP Aspen Conference, 1976. Skokie, IL: College of American Pathologists, 1977.
- World Association of Societies of Pathology. Proceedings of the Subcommittee on Analytical Goals in Clinical Chemistry. Analytical goals in clinical chemistry: their relationship to medical care. *Am J Clin Pathol* 1979;72:624–30.
- Burnett D, Ceriotti F, Cooper G, Parvin C, Plebani M, Westgard J. Collective opinion paper on findings of the 2009 convocation of experts on quality control. *Clin Chem Lab Med* 2010;48:41–52.
- Gras JM, Philippe M. Application of the six sigma concept in clinical laboratories: a review. *Clin Chem Lab Med* 2007;45: 789–96.
- Oosterhuis WP. Gross overestimation of total allowable error based on biological variation. *Clin Chem* 2011;57:1334–6.
- Gowans EM, Hyltoft Petersen P, Blaabjerg O, Horder M. Analytical goals for the acceptance of common reference intervals for laboratories throughout a geographical area. *Scand J Clin Lab Invest* 1988;48:757–64.
- Hyltoft-Petersen P, Fraser CG, Westgard JO, Larsen ML. Analytical goal-setting for monitoring patients when two analytical methods are used. *Clin Chem* 1992;38:2258–60.
- Larsen ML, Fraser CG, Hyltoft Petersen P. A comparison of analytical goals for haemoglobin HbA1c assays derived using different models. *Ann Clin Biochem* 1991;28:272–8.
- Bruns DE. Laboratory related outcomes in healthcare. *Clin Chem* 2001;47:1547–52.
- Boyd JC, Bruns DE. Performance requirements for glucose assays in intensive care units. *Clin Chem* 2014;60:1463–5.
- Cotlove E, Harris EK, Williams GZ. Biological and analytic components of variation in long-term studies of serum constituents in normal subjects. 3. Physiological and medical implications. *Clin Chem* 1970;16:1028–32.
- Harris EK. Statistical principles underlying analytic goal-setting in clinical chemistry. *Am J Clin Pathol* 1979;72:374–82.
- International vocabulary of metrology – basic and general concepts and associated terms (VIM). ISO/IEC Guide 99, 2007.
- Klee GG. Tolerance limits for short-term analytical bias and analytical imprecision derived from clinical assay specificity. *Clin Chem* 1993;39:1514–8.
- Westgard JO, Carey RN, Wold S. Criteria for judging precision and accuracy in method development and evaluation. *Clin Chem* 1974;20:825–33.
- Fraser CG, Hyltoft Petersen P. Quality goals in external quality assessment are best based on biology. *Scand J Clin Lab Invest* 1993;53(Suppl):8–9.
- Schoenmakers CH, Naus AJ, Vermeer HJ, van Loon D, Steen G. Practical application of Sigma Metrics QC procedures in clinical chemistry. *Clin Chem Lab Med* 2011;49:1837–43.
- Stöckl D, Baadenhuijsen H, Callum G, Fraser CG, Jean-Claude Libeer JC, Hyltoft Petersen P, et al. Desirable routine analytical goals for quantities assayed in serum. *Eur J Clin Chem Clin Biochem* 1995;33:157–69.
- Solberg HE. Approved recommendation (1987) on the theory of reference values. Part 5. Statistical treatment of collected reference values: determination of reference limits. *J Clin Chem Clin Biochem* 1987;25:645–56.
- Ricos tables on biological variation. Available from: www.westgard.com/biodatabase1.htm. Accessed 10 April, 2015.
- Fraser CG, Hyltoft Petersen P. Analytical performance characteristics should be judged against objective quality specifications. *Clin Chem* 1999;45:321–23.
- Stöckl D, Thienpont LM. About the z-multiplier in total error calculations. *Clin Chem Lab Med* 2008;46:1648–9.
- Fraser CG, Hyltoft Petersen P, Libeer JC, Ricos C. Proposals for setting generally applicable quality goals solely based on biology. *Ann Clin Biochem* 1997;34:8–12.