Accred Qual Assur (2006) 11: 319–328
DOI 10.1007/s00769-006-0113-0

GENERAL PAPER

Andrea Konnert
Sabine Arends
Stefan Schubert
Christoph Berding
Cas Weykamp
Carla Siebelder

Uncertainty calculation for calibrators of the IFCC HbA1c standardization network

Received: 30 December 2005
Accepted: 15 February 2006
Published online: 18 March 2006

A. Konnert (✉) · S. Arends · S. Schubert · C. Berding
Roche Diagnostics GmbH, Department of Biostatistics,
Nonnenwald 2,
82377 Penzberg, Germany
e-mail: andrea.konnert@roche.com
Tel.: +49-8856-604647
Fax: +49-8856-604152

A. Konnert
University of Dortmund, Department of Statistics,
Vogelpothsweg 87,
44221 Dortmund, Germany

S. Schubert
University of Ulm, Department of Mathematics and Economics,
Helmholtzstraße 18,
89069 Ulm, Germany

C. Weykamp · C. Siebelder
Queen Beatrix Hospital,
Beatrix Park 1, 7101 BN, Winterswijk,
The Netherlands

Abstract Within the last decade, the IFCC HbA1c standardization network has established the metrologically highest reference measurement procedure for HbA1c testing. Based on this procedure, reference calibrators are produced which in turn provide the starting point for the standardization of the manufactures routine HbA1c assays. According to the IVD directive, the uncertainty of the reference calibrators must be calculated and reported together with their assigned values. Within this article, we elaborate the uncertainty calculation according to GUM (guide to the expression of uncertainty in measurement) in detail. Finally, the results are validated by a simulation study.

Keywords HbA1c reference method · Uncertainty · Standardization

Introduction

The measurement of hemoglobin A1c (beta-N-terminal glycated hemoglobin A, denoted further as HbA1c) in percent of total hemoglobin (HbA1c and HbA0 – non-glycated hemoglobin A) in human blood is the most important biomedical marker for long-term assessment of the glycemic status in patients with diabetes mellitus, and goals for therapy are set at specific HbA1c target values [1]. The International Federation of Clinical Chemistry (IFCC) recognized the need for a reliable anchor of this major biomedical analyte and installed the IFCC Working Group on HbA1c standardization [2]. This group succeeded to develop a reference system of highest metrological order, which has been approved by all member national societies of the IFCC [3]. In the IVD Directive 98/79/EC [4] as well as in international standards [5, 6] essential requirements for in vitro medical devices are described in general terms, where a major specified requirement, directly related to standardization, is metrological traceability: when available, the calibration of biomedical assays must be traceable to reference measurement procedures and/or reference materials of a higher metrological order.

The components of the HbA1c reference system cover the upper part of the traceability chain: HbA1c is defined on basis of its molecular structure, based on pure HbA1c and HbA0 primary calibrators as mixtures of these analytes are obtained. The reference measurement procedure is the approved reference method (enzymatic digestion followed by HPLC [3]) and the secondary calibrators are whole blood

panels to which values have been assigned with the reference method.

Traceability, defined as "property of the result of a measurement or a value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties" [7], requires the evaluation of the uncertainty of calibrators on every metrological level.

In this article, we present the uncertainty evaluation of the primary and secondary calibrator. In addition to the usual evaluation approach according to GUM [8], we calculated the uncertainties based on a simulation study in order to validate the GUM results, as recommended by Cox and Harris [9].

The article is organized as follows: First, we give a short summary of uncertainty calculation. In the second section, the production process of the primary calibrators is explained as well as the application of the GUM approach to calculate the uncertainty of the primary calibrators. Afterwards, the uncertainty calculation of the secondary calibrators is derived. Finally, we demonstrate by a simulation study the appropriateness of the results obtained via the GUM approach.

Uncertainty calculation

According to GUM, the uncertainty calculation is based on a well-defined metrological model, where error-prone input quantities are related to the measurement quantity under consideration by a model function. The estimates as well as the standard uncertainties of the input quantities need to be known. Based on the model function and by means of the law of error propagation, the value of the output quantity and its uncertainty can be derived.

Denote with $x_1, \ldots, x_r$ measurements of the input quantities, with $u(x_1), \ldots, u(x_r)$ their corresponding standard uncertainties and with $\rho_{ij}, 1 \leq i, j \leq r$ the correlation between the measurements x_i and x_j. The output quantity or measurement result is then given by $y = f(x_1, \ldots, x_r)$.

The variance of the measurement result can be approximated by

$$u^2(y) = \sum_{i=1}^{r} \left(\frac{\partial f}{\partial x_i} \right)^2 \cdot u^2(x_i) + \sum_{\substack{i,j \\ i \neq j}} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} \cdot \rho_{ij} \cdot u(x_i) \cdot u(x_j). \quad (1.1)$$

The partial derivatives of the function f with respect to x_i, are evaluated for the actual values of the estimates. These are called sensitivity coefficients and describe the change of the measurement result y in response to changes, in the input quantity x_i. The second sum of Eq. (1.1) vanishes under the assumption of mutual independence of the input quantities. We will see, however, that in some cases it is important to take the correlation structure of the estimates of the input quantities into account.

The uncertainty of y is then given by the square root of the variance.

Based on the resulting uncertainty for y a coverage region can be determined which covers the true value of y with a given confidence, e.g., 95%. Under the assumption that the output quantity is normally distributed and unbiased, the 95% coverage region is approximated by $[y - 2 \cdot u(y), y + 2 \cdot u(y)]$.

However, if the measurement contains uncorrected bias, the coverage region, which covers the true value with a given confidence, may no longer be centered around the measurement result. In this case, the above coverage region must be expanded in the direction of the bias. Denote with $\delta_{\min}(y)$ the minimal bias and with $\delta_{\max}(y)$ the maximum bias. According to Phillips and Eberhardt [10], a coverage region for the true value under the assumption that the measurements carry bias is given by

$$[y - 2 \cdot u(y) - \delta_{\min}(y), y + 2 \cdot u(y) + \delta_{\max}(y)]. \quad (1.2)$$

However, Eq. (1.2) still involves assumptions that may not be fulfilled in our present situation. In particular, the first-order approximation of Eq. (1.1) may not be sufficient. Moreover, although the assumption of independence is often not fulfilled for the sake of simplicity, the necessary but cumbersome evaluation of the correlation structure is often not done. Hence, a validation of the derived confidence intervals (Eq. (1.2)) by a simulation study is quite appropriate.

Primary calibrators

A primary standard containing as the major hemoglobin derivative HbA1c and a primary standard containing as the major hemoglobin derivative HbA0 are mixed to give the primary calibrator (reference material) for the HbA1c reference method. The preparation of the primary calibrators, as well as the involved materials and methods are explained in detail in [11]. Knowing the hemoglobin concentration of these primary standards, a calibrator panel comprising six levels (0, 3, 6, 9, 12, and 15 [HbA1c/(HbA1c + HbA0) %]) is produced.

The reference method is calibrated with the above primary calibrator panel by linear calibration. This designated primary reference system conduces to assign target values to other samples (e.g., whole blood samples), which then in turn serve as secondary calibrators.

Traces of HbA0 in the HbA1c primary standard are estimated based on a standard addition procedure and the calculated ratio of HbA1c to total hemoglobin is corrected for these traces. The amount of HbA0 in the HbA0 primary standard is stated by the manufacturer to be higher than 98%, however we could demonstrate that it is in fact higher than 99.98%, where the resulting 0.02% are mostly traces of HbA1c. A more exact determination of the traces of HbA1c in the HbA0 primary standard is difficult and

laborious, therefore the assigned values of the calibrators are not corrected for this amount, but are included as maximum bias within the uncertainty calculation.

Thus, the uncertainty of the primary calibrators has the following contributions: (i) Errors in the estimation process of the amount of hemoglobin in the HbA0, HbA1c primary standards, respectively. (ii) Errors in the weighting and pipetting process when mixing the HbA0 and HbA1c primary standards to generate the six concentration levels. (iii) Errors in the estimation of HbA0 traces in the HbA1c primary standard. (iv) Bias due to the HbA1c traces in the HbA0 primary standard.

Total hemoglobin estimation

In a first step, the concentration of total hemoglobin in the HbA0 and HbA1c primary standards is estimated by photospectrometrical measurements according to the ICSH reference method [12] in two independent laboratories. Each laboratory performs measurements in triplicate. For each measurement, multiple absorption readings are made and the respective concentration is derived from the mean of the read absorptions. The concentration of the solutions is measured in [mg Hb/g solution].

To minimize an instrument bias (photospectrometer measuring too high or too low) or a concentration depending bias (different absorption at low or high concentrations) both standards are measured in the same run with approximately the same concentration. To each standard, the mean of the six reaction mixes is assigned as total hemoglobin concentration: $\bar{c}_{hb(A0)}$, $\bar{c}_{hb(A1c)}$.

Mixing procedure

Once the concentrations of the primary standards are known, the required amounts of the primary standards are calculated to obtain mixtures close to the specified calibrator levels. The exact amounts of the respective primary standards are determined by weighing rather than by pipetted volumes in order to reduce pipetting errors. Based on the weighing procedure, the assigned values of the primary calibrators are calculated and replace the originally specified calibrator levels. Hence, the weighed masses of the primary standards are influenced by the error of the weighing process exclusively.

The precision of the balance is stated by the manufacturer. We will denote the weighed masses of the primary standards for calibrator level j by $w^j_{hb(A0)}$, $w^j_{hb(A1c)}$.

Amount of HbA0 in the HbA1c primary standard

The HbA1c primary standard carries traces of HbA0, which need to be determined. The amount of these traces is estimated via a standard addition procedure [13]. Denote with $c_{hb} = c_{HbA1c} + c_{HbA0}$ the amount of total hemoglobin in an unspiked sample based on the HbA1c primary standard and with c^s_{HbA0} the amount of HbA0 spiked in. By the spiking method, samples with 0 to 12.5% of c^s_{HbA0}/c_{hb} are produced. These in turn are measured twice by the reference method with two repetitions for each measurement. Based on the integrated peak areas of HbA1c to HbA0, the response ratio of $(c^s_{HbA0} + c_{HbA0})/c_{HbA1c}$ is obtained. As the relationship between the response ratio and the concentration ratio is linear (see [3]), a linear regression of the response ratios on the concentration ratios can be performed.

The amount of HbA0 in the HbA1c primary standard, denoted by $i_{A0/A1c}$, is obtained by interpolating the regression line to the intercept with the x-axis [14].

Some elements of this procedure may suffer from errors: inaccuracies in dilution and pipetting heterogeneities in the digestion reaction, measurement and instrument effects, as well as impurities of the HbA0 primary standard. The estimation and combination of all these contributions into an appropriate mathematical model might be quite complicated. Thus, the uncertainty of the estimated amount of HbA0 is determined experimentally, using a design that includes the most relevant of the aforementioned variances. The whole procedure–standard dilution, digestion, measurement–is done twice with duplicates. As the uncertainty contribution of the traces of HbA1c within the HbA0 standard, used for spiking, is very small, it has been neglected.

The uncertainty of the resulting estimated amount, $\bar{i}_{A0/A1c}$ has been estimated by variance component analysis [15, 16] (see also Appendix A) with components "digest" and "repetition", where digest abbreviates the complete procedure including standard addition.

Amount of HbA1c in the HbA0 primary standard

The amount of HbA1c in the HbA0 primary standard, denoted by $i_{A1c/A0}$, is checked by injecting a ten times concentrated HbA0 standard as well as a ten times concentrated HbA0 standard spiked with 0.02% HbA1c on capillary electrophoreses. With the 0.02% HbA1c specimen, a peak is seen at the position of HbA1c, while this is missing for the

Table 1 Uncertainty components and sensitivity coefficients of primary calibrators

Uncertainty component	Sensitivity coefficient $\partial f/\partial x_i$
$\bar{c}_{hb(A0)}$	$\dfrac{w^j_{hb(A1c)} \cdot \bar{c}_{hb(A1c)} \cdot (i_{A0/A1c} - 100) \cdot w^j_{hb(A0)}}{\bar{c}^2_{hb}}$
$\bar{c}_{hb(A1c)}$	$-\dfrac{w^j_{hb(A0)} \cdot \bar{c}_{hb(A0)} (i_{A0/A1c} - 100) \cdot w^j_{hb(A1c)}}{\bar{c}^2_{hb}}$
$w^j_{hb(A1c)}$	$\dfrac{w^j_{hb(A1c)} \cdot \bar{c}_{hb(A1c)} \cdot (i_{A0/A1c} - 100) \cdot \bar{c}_{hb(A0)}}{\bar{c}^2_{hb}}$
$w^j_{hb(A0)}$	$-\dfrac{w^j_{hb(A0)} \cdot \bar{c}_{hb(A0)} \cdot (i_{A0/A1c} - 100) \cdot \bar{c}_{hb(A1c)}}{\bar{c}^2_{hb}}$
$i_{A0/A1c}$	$-\dfrac{w^j_{hb(A1c)} \cdot \bar{c}_{hb(A1c)}}{\bar{c}_{hb}}$

HbA0 standard. From this, it is concluded that the amount in question is less than 0.02%. This deviation from a theoretical 100% pure solution is not used for an adjustment of the calibrator assigned value and thus enters the calculation of the coverage region as bias. The bias term is added on the right-hand side of the confidence interval, which results in a non-symmetric interval around the assigned value of the calibrator.

Combined uncertainty

The assigned values of the primary calibrators in percent of HbA1c/(HbA1c + HbA0) are obtained by

$$c^j_{pcal} = f\left(\bar{c}_{hb(A0)}, \bar{c}_{hb(A1c)}, w^j_{hb(A0)}, w^j_{hb(A1c)}, i_{A0/A1c}\right)$$
$$= \frac{c^j_{HbA1c}}{\bar{c}^j_{hb}} \cdot 100 = \frac{w^j_{hb(A1c)} \cdot \bar{c}_{hb(A1c)} \cdot \left(1 - \frac{i_{A0/A1c}}{100}\right)}{w^j_{hb(A0)} \cdot \bar{c}_{hb(A0)} + w^j_{hb(A1c)} \cdot \bar{c}_{hb(A1c)}} \cdot 100.$$

According to GUM, the uncertainty of the assigned values of the primary calibrators can be obtained by applying Eq. (1.1). The sensitivity coefficients of the different components are given in Table 1.

However, as the amount of HbA1c in the HbA0 primary standard is not taken into account in value assignment, its contribution expands the coverage region by the bias term.

The maximum bias due to the impurity of HbA1c in the HbA0 primary standard for each calibrator level is calculated as:

$$\delta_{\max}\left(c^j_{\text{pcal}}\right) = 100 \cdot \frac{w^j_{hb(A0)} \cdot \bar{c}_{hb(A0)} \cdot i_{A1c/A0}}{\bar{c}^j_{hb}}.$$

The minimum bias is zero, which leads to a non-symmetric confidence interval around the assigned value:

$$\left[c^j_{\text{pcal}} - k \cdot u\left(c^j_{\text{pcal}}\right), c^j_{\text{pcal}} + k \cdot u\left(c^j_{\text{pcal}}\right) + \delta_{\max}\left(c^j_{\text{pcal}}\right)\right].$$

As has already been mentioned, this coverage region is approximate. Another approach to derive a reasonable coverage region utilizes simulation. Simulation studies allow for an easy modeling of correlation structures and distributions. Indeed, Cox and Harris [9] propose to use simulation studies for the validation of the model assumptions and the respective results (see the Simulation section).

Quantifying uncertainties

Exemplarily we present the uncertainty calculations for the primary calibrators pcal 2004. For the production of these calibrators, the concentrations of the HbA0 primary standard (PRM Lot nr. 90479699) and HbA1c primary standard (PRM Lot nr. 90847099) are assigned in two independent laboratories and three reaction mixes per laboratory by multiple photospectrometrical measurements. Standard variance component analysis allows for the estimation of the between-laboratory variation as well as the within-laboratory variation of the photospectrometrical measurements. Based on both variances, the uncertainty of the mean of the spectrometric measurements is given by

$$u(\bar{c}_{hb(A0)}) = \sqrt{\frac{1}{2}u^2_{\text{betw.}} + \frac{1}{2 \cdot 3}u^2_{\text{with.}}}.$$

For the estimation of both variances, see Appendix A.

For the pcal 2004 calibrators, this results in $\bar{c}_{hb(A0)} = 118.487$ [mg/g], $u(\bar{c}_{hb(A0)}) = 0.185$ [mg/g], $\bar{c}_{hb(A1c)} = 18.70$ [mg/g], $u(\bar{c}_{hb(A1c)}) = 0.026$ [mg/g].

The weighing of the primary standards takes place on a Sartorius BP 211 D, with stated uncertainty of the manufacturer 0.00005 [g]. The weights for the different calibrator levels are given in Table 2.

The estimated amount of HbA0 in the HbA1c primary standard is given by the mean of the four repeated calculations and its uncertainty is also calculated on the basis of the nested design of the data, by estimating the between-digest variation as well as the within-digest variation, resulting in $i_{A0/A1c} = 6.59$ [%], $u(i_{A0/A1c}) = 0.224$ [%].

In Table 2, the standard uncertainty, the maximum bias, and the 95% coverage region for the primary calibrators, according to GUM, are given. The main contribution to the combined uncertainty of the primary calibrators comes from the estimation process of the amount of HbA0 in the HbA1c primary standard (about 60%), followed by the estimation of the concentration of the HbA0 primary standard (20%) and HbA1c (17%) primary standard. The uncertainty contribution of the mixing procedure is marginal.

As for each level of the primary calibrators, the estimated concentrations of the primary standards, as well as the estimated amount of HbA0 in the HbA1c standard enter the calculation of the assigned value, these values are correlated. However, there is no closed form for the calculation of this correlation. Results of the simulation study revealed

Table 2 Uncertainty and coverage regions of the primary calibrators pcal 2004 according to GUM

Calibrator level (target values)	$w_{hb(A0)}$ (g)	$w_{hb(A1c)}$ (g)	Lower limit ($k = 2$) (%)	Assigned value (%)	Upper limit ($k = 2$ + bias) (%)	Bias (%)	Combined uncertainty (%)
A (0%)	1.86108	0.00000	0.00	0.00	0.02	0.020	0.000
B (3%)	1.78244	0.36463	2.90	2.92	2.96	0.019	0.009
C (6%)	2.15490	0.90167	5.75	5.79	5.84	0.019	0.018
D (9%)	1.68478	1.10010	8.67	8.73	8.80	0.018	0.027
E (12%)	2.01571	1.79165	11.42	11.49	11.58	0.018	0.035
F (15%)	1.56248	1.81598	14.39	14.48	14.58	0.017	0.043

that the correlation of the assigned values of two calibrator levels, is between 0.99 and 0.90, becoming smaller with increasing distance between the assigned values. This correlation should be kept in mind when considering the uncertainty of the assigned values of the secondary calibrators.

Secondary calibrators

Secondary calibrators of the HbA1c IFCC reference system are human blood samples collected under controlled conditions. They are measured by the reference measurement procedure in laboratories in the U.S., Europe and Asia, all being members of the HbA1c IFCC standardization network. The reference measurement procedure is calibrated with the primary calibrators, thus a proportional relationship is established between the ratio of integrated areas of the HbA1c peak and the HbA0 peak. Each secondary calibrator is measured in two digests and two repetitions per digest in each laboratory. "Digest" in this context refers to the enzymatical preprocessing of the analyte before the chromatographic analysis. A detailed description of the measurement process (HPLC – ESI/MS) is found in [11]; the data analysis is described in [17].

The following uncertainty contributions must be regarded for the assigned value of the secondary calibrators: (i) uncertainty component due to the measurement error of the reference method, (ii) uncertainty component due to the uncertainty of the primary calibrators and a bias component due to the bias of the primary calibrators.

Uncertainty due to measurement

After appropriate data analysis of the measured data points, the means of the measured results become the assigned values of the secondary calibrators. The uncertainty component due to the measurement error of the reference method is calculated considering the nested measurement design (laboratory, digestion, repetition) for each sample. The between-laboratory, between-digest, and within-digest variances enter the calculation of the uncertainty component of the secondary calibrators in the following way:

$$u(c_{\mathrm{scal}})_{\mathrm{meas}} = \sqrt{\frac{1}{L}u^2_{b.-lab} + \frac{1}{L\cdot D}u^2_{b.-dig} + \frac{1}{L\cdot D\cdot R}u^2_{w.-dig}},$$

where L denotes the number of laboratories, D the number of digests and R the number of repetitions (see also Appendix A).

Uncertainty due to calibration

As the reference measurement method is calibrated by means of the primary calibrators, which are already subject to uncertainty, this uncertainty must be transmitted to the secondary calibrators.

The uncertainty component of the secondary calibrators due to calibration is approximated by weighing the uncertainties of the next lower $(c^i_{\mathrm{pcal}}, u(c^i_{\mathrm{pcal}}))$ and higher primary calibrator $(c^{i+1}_{\mathrm{pcal}}, u(c^{i+1}_{\mathrm{pcal}}))$, compared to the assigned values of the secondary calibrators. The correlation between these two calibrator values, denoted by $\rho_{i,i+1}$ is also considered, resulting in

$$u(c_{\mathrm{scal}})_{\mathrm{cal}} = s^2\cdot u^2\left(c^{i+1}_{\mathrm{pcal}}\right) + (1-s)^2\cdot u^2\left(c^{i}_{\mathrm{pcal}}\right) + \rho_{i,i+1}\cdot s\cdot(1-s)\cdot u\left(c^{i}_{\mathrm{pcal}}\right)\cdot u\left(c^{i+1}_{\mathrm{pcal}}\right), \quad (1.3)$$

where

$$s = \frac{c_{\mathrm{scal}} - c^i_{\mathrm{pcal}}}{c^{i+1}_{\mathrm{pcal}} - c^i_{\mathrm{pcal}}}.$$

The derivation of this formula is presented in Appendix B.

For the actual data the correlation between two calibrator values was approximated with $\rho_{i,i+1} = 0.99$.

Combined uncertainty and bias

The combined uncertainty of the assigned values of the secondary calibrators then takes the form:

$$u^2(c_{\mathrm{scal}}) = u^2(c_{\mathrm{scal}})_{\mathrm{meas}} + u^2(c_{\mathrm{scal}})_{\mathrm{cal}}.$$

Finally, the bias of the primary calibrators must also be taken into account, eventually leading to a bias of the secondary calibrators, too. The same ideas that led to Eq. (1.3) can be applied to calculate the maximum bias for each secondary calibrator. In Appendix B we show that this bias is given by

$$\delta_{\max}(c_{\mathrm{scal}}) = s\cdot\delta_{\max}\left(c^{i+1}_{\mathrm{pcal}}\right) + (1-s)\cdot\delta_{\max}\left(c^{i}_{\mathrm{pcal}}\right).$$

Thus, a 95% confidence interval $(k=2)$ of the assigned values of the secondary calibrators is obtained:

$$[c_{\mathrm{scal}} - 2\cdot u(c_{\mathrm{scal}}),\ c_{\mathrm{scal}} + 2\cdot u(c_{\mathrm{scal}}) + \delta_{\max}(c_{\mathrm{scal}})].$$

Quantifying uncertainties

The primary calibrators pcal 2004 are used as calibrators within the recent studies. Therefore, the uncertainty calculation for the secondary calibrators, to which values are assigned within the Orlando 2 study, is presented. These secondary calibrators are named Orlando 6 – Orlando 10.

Table 3 Uncertainty contributions, bias and 95% confidence intervals for the secondary calibrators from the Orlando 2 study according to GUM

Secondary calibrator	Uncertainty measurement (%)	Uncertainty calibration (%)	Combined uncertainty (%)	Bias (%)	Lower limit ($k=2$) (%)	Assigned value (%)	Upper limit ($k=2$ + bias) (%)
Orlando 6	0.034	0.018	0.038	0.019	6.32	6.40	6.50
Orlando 7	0.027	0.011	0.029	0.019	3.90	3.96	4.04
Orlando 8	0.037	0.025	0.045	0.018	8.39	8.48	8.59
Orlando 9	0.042	0.018	0.046	0.019	5.62	5.71	5.82
Orlando 10	0.017	0.010	0.019	0.019	3.44	3.48	3.54

Fourteen reference laboratories contributed four results per laboratory within the Orlando 2 study. The mean of the results becomes the assigned value of the samples. Table 3 summarizes the uncertainty contributions, the bias, and the 95% confidence intervals for these calibrators.

The calculation of the between-laboratory, between-digest, and within-digest variances showed that the main contribution of variation of the measurements comes from between-laboratory variation, followed by within-digest variation. In order to reduce the influence of the between-laboratory variation it is thus mandatory to have a reasonable number of laboratories to average. In the IFCC network there are more than ten laboratories participating. To give an example: The uncertainty component due to measurement of the assigned value of the Orlando 7 sample is 0.027 HbA1c[%]. For this sample, the between-laboratory variation makes 64%, the between-digest 6% and within-digest 30% of the total variation. If the value would have been assigned by only three laboratories, this uncertainty component would rise to 0.057 HbA1c[%]. If 20 laboratories would have participated in this value assignment, this uncertainty component would be 0.022 HbA1c[%]. Hence it is clear that value assignment by a network, rather than by a limited number of network laboratories, significantly contributes to the low uncertainty of the secondary calibrators. But a further increase of the number of network laboratories will have a relatively small impact on the uncertainty.

The uncertainty of the primary calibrators is transmitted to the secondary calibrators via the uncertainty component due to calibration. If the uncertainty of the primary calibrators could be limited to zero, the combined uncertainty of the secondary calibrators would be equal to the uncertainty component due to measurement.

Simulation study

To validate the model assumptions invoked for the uncertainty calculation according to GUM, a simulation study of the value assignment process of both primary and secondary calibrators has been conducted. Recall that the purpose of uncertainty calculation is the determination of a coverage region, which covers the true value of the measurement with a given confidence level. This means we are interested in the distribution of the differences between true value and measured value. Hence, we calculated in each simulation step the (anticipated) true value of the calibrators as well as the assigned value obtained during the production process. Based on the distribution of the differences, the coverage region for the true value can be obtained.

Primary calibrators

The assigned values of the primary calibrators are obtained by mixing a HbA0 primary standard and a HbA1c primary standard and correcting for the amount of HbA0 in the HbA1c standard and vice versa. Denote with the subscript t the true amounts of the input quantities for the calculation of the assigned values of the primary calibrators. Based on these the true assigned value of each calibrator is obtained by

$$c^t_{\mathrm{pcal}_j} = f\left(c^t_{hb(A0)}, c^t_{hb(A1c)}, w^t_{hb(A0)_j}, w^t_{hb(A1c)_j}, i^t_{A0/A1c}, i^t_{A1c/A0}\right)$$
$$= \frac{w^t_{hb(A1c)_j} \cdot c^t_{hb(A1c)} \cdot \left(1 - \frac{i^t_{A0/A1c}}{100}\right) + \frac{i^t_{A1c/A0}}{100} \cdot w^t_{hb(A0)_j} \cdot c^t_{hb(A0)}}{w^t_{hb(A0)_j} \cdot c^t_{hb(A0)} + w^t_{hb(A1c)_j} \cdot c^t_{hb(A1c)}} \cdot 100.$$

For the simulations the true amounts are derived from the data of the pcal 2004 calibrators, i.e., $c^t_{hb(A0)} = 118.487$ [mg/g], $c^t_{hb(A1c)} = 18.70$ [mg/g]. The true pipetted amounts are modeled as uniform distributed around the target weights, which are calculated to achieve the respective concentration levels. The dispersion of the uniform distribution is based on precision and bias information for the pipettes from manufactures.

The true amount of HbA0 in the HbA1c primary standard is set to 6.59%. Finally, the true amount of HbA1c in the HbA0 standard is modeled uniform distributed in the interval [0%, 0.02%].

However, within the production process of the primary calibrators, only estimates of the true input quantities are available and the assigned values of the primary calibrators are calculated without the correction for the amount of HbA1c in the HbA0 primary standard.

In each simulation step, estimates of the input quantities are obtained as follows:

1. $\bar{c}_{hb(A0)}$, $\bar{c}_{hb(A1c)}$ are drawn from normal distributions with mean $\bar{c}^t_{hb(A0)}$, $\bar{c}^t_{hb(A1c)}$ and standard deviations 0.185 [mg/g], 0.026 [mg/g], respectively.
2. The weighed amounts are modeled normally distributed centered at the true amounts and with standard deviation 0.00005 [g].
3. The impurity estimation is simulated step by step.

Table 4 Absolute and relative differences between the GUM and simulation approach for the uncertainties as well as the lower and upper limits of the coverage regions for $k = 2$

	Differences uncertainties GUM – simulation		Differences lower limits ($k = 2$) GUM – simulation		Differences upper limits ($k = 2$) GUM – simulation	
Primary calibrators	Absolute (%)	In % of assigned value	Absolute (%)	In % of assigned value	Absolute (%)	In % of assigned value
A	0.000		0.00		0.00	
B	0.000	− 0.01	− 0.01	− 0.25	0.01	0.21
C	− 0.001	− 0.01	− 0.01	− 0.15	0.01	0.11
D	− 0.001	− 0.01	− 0.01	− 0.11	0.01	0.07
E	− 0.001	− 0.01	− 0.01	− 0.08	0.00	0.04
F	− 0.001	− 0.01	− 0.01	− 0.07	0.00	0.03
Secondary calibrators						
Orlando 6	− 0.003	− 0.05	0.002	0.03	0.017	0.26
Orlando 7	− 0.003	− 0.08	0.002	0.04	0.017	0.42
Orlando 8	− 0.002	− 0.03	0.001	0.01	0.017	0.20
Orlando 9	− 0.002	− 0.03	0.000	0.00	0.020	0.35
Orlando 10	− 0.004	− 0.11	0.003	0.08	0.017	0.50

Based on the estimated hemoglobin concentrations of the HbA0, HbA1c primary standards, the pipetting volumes for the spiked samples are calculated, denoted as $w^s_{hb(A0)}$, $w^s_{hb(A1c)}$. Afterwards, the assumed ratio of HbA0 to hemoglobin from the HbA1c standard, disregarding the amount of HbA0 in the HbA1c standard, is calculated as

$$r_a = \frac{w^s_{hb(A0)} \cdot \bar{c}_{hb(A0)}}{w^s_{hb(A1c)} \cdot \bar{c}_{hb(A1c)}},$$

whereas the true ratio, i.e., taking this amount into account, would be

$$r^t = \frac{w^s_{hb(A0)} \cdot \bar{c}^t_{hb(A0)} + i^t_{A0/A1c} \cdot w^s_{hb(A1c)} \cdot \bar{c}^t_{hb(A1c)}}{w^s_{hb(A1c)} \cdot \bar{c}^t_{hb(A1c)}}.$$

Measuring the spiked samples in the HPLC method and setting the integrated peak areas in comparison, results in the measured ratio of the spiked samples. The measured ratio is proportional related to the true ratio [11] and clearly subject to measurement error. Therefore it is modeled as $r_m \sim N(c \cdot r_t, \sigma^2_m)$, where the variance is estimated from the data at hand. Based on the data points (r^j_a, r^j_m), where j denotes the number of spiked samples, the linear model $r_m = a + b \cdot r_a$ is estimated. One can show [14] that the respective amount is obtained by interpolating the regression line to the intersection with the x-axis, i.e., $i_{A0/A1c} = \hat{a}/\hat{b}$, where $\hat{a}, \hat{b}$ denote the least-square estimators of intercept and slope. The assigned values of the primary calibrators are then calculated by

$$c_{\text{pcal}_j} = \frac{w_{hb(A1c)_j} \cdot \bar{c}_{hb(A1c)} \cdot \left(1 - \frac{i_{A0/A1c}}{100}\right)}{w_{hb(A0)_j} \cdot \bar{c}_{hb(A0)} + w_{hb(A1c)_j} \cdot \bar{c}_{hb(A1c)}} \cdot 100.$$

Secondary calibrators

To obtain a coverage region of the secondary calibrators, the following steps have been performed: The calibration of the reference method with the primary calibrators was modeled as a method comparison between the true primary calibrator values and assigned values from production. This is possible as the relationship between the true concentrations and signals obtained in the reference method is proportional.

For each primary calibrator, we have a data point $(c^t_{\text{pcal}_j} + \varepsilon_{\text{meas}}, c_{\text{pcal}_j})$. The measurement error is normally distributed with mean 0 and variance estimated from the data at hand. A regression line is fitted to the data points, which will not be centered around the bisecting line as the assigned calibrator values from production carry uncorrected bias. For the secondary calibrator, the true value is given within the simulation and in each simulation step the assigned value of the secondary calibrator is read from the regression line at the point $c^t_{\text{scal}} + \varepsilon_{\text{meas}}$. The measurement error variance is approximated by the standard error of the mean of the measured calibrator values derived from the Orlando 2 study.

To obtain a coverage region for the true values of the secondary calibrators, again, the distribution of the differences $c_{\text{scal}} - c^t_{\text{scal}}$ is considered.

Coverage regions from simulation

The simulation study has been terminated after stable convergence of the uncertainty results. The 2.5% and 97.5% quantiles of the differences for the primary as well as the secondary calibrators were calculated. Based on these quantiles, a 95% coverage region for the true calibrator values can be obtained for the primary calibrators as $[c_{\text{pcal}} - Q_{97.5_{\text{pcal}}}, c_{\text{pcal}} - Q_{2.5_{\text{pcal}}}]$ and for the secondary calibrators $[c_{\text{scal}} - Q_{97.5_{\text{scal}}}, c_{\text{scal}} - Q_{2.5_{\text{scal}}}]$.

Absolute and relative differences between the GUM and simulation approach for the uncertainties as well as the lower and upper limits of the coverage regions for primary as well as the secondary calibrators are summarized in Table 4.

Comparison of GUM and simulation

The uncertainties of the primary calibrators, derived according to GUM, are 0.001[%] smaller than the respective uncertainties from simulation. On detailed inspection of the simulation results, we observe that the uncertainty component from estimating the amount of HbA0 in the HbA1c standard was always higher in the simulation study than the four repeated experiments. Therefore, we modeled the impurity estimation in detail according to GUM and with this approach obtained an estimate for the impurity uncertainty, which was very near to the impurity derived via the simulation approach.

However, as the bias contribution is added in a conservative way to the coverage regions, and as the multiplication factor $k = 2$ is used instead of the 0.975 normal quantile ($k = 1.96$), the coverage regions according to GUM are wider than those from simulation, especially for primary calibrators of the lower levels. Moreover, in comparison to the estimated target values and uncertainties, the observed differences between the simulation approach and GUM are only marginal, hence we conclude that the GUM approach delivers appropriate results for the uncertainties as well as for the coverage regions of the values of the primary calibrators.

Regarding the secondary calibrators, we note that the differences between the estimated uncertainties based on GUM and the uncertainties of the simulation approach were greater than for the primary calibrators, but still the factor of $k = 2$ makes the limits of the coverage regions become wider for the GUM approach than for the simulation approach. The lower uncertainty according to GUM is mostly due to the approximation of the transmission of the uncertainty of the primary calibrators to the secondary calibrators.

The most interesting observation is that the differences in the upper limit of the coverage region are higher than the additional bias, which was added to enlarge the upper limit. This means that the bias of the primary calibrators is not transmitted on the secondary calibrators in a relevant way. Therefore, we decided to ignore the bias contribution when calculating the coverage region for the secondary calibrators in the future. This has a definite practical impact, as the secondary calibrators constitute the basis for manufactures of routine assays for HbA1c testing. Based on the simulation, we see that in the present case the bias of the primary calibrators can be neglected.

Conclusions

In this article, we presented in detail the uncertainty calculation of the IFCC reference material for HbA1c testing. The uncertainty of the assigned value of a calibrator allows the determination of a region, which covers the true value of the calibrator with high confidence Moreover, the calculation of the uncertainty has led to a very instructive review of the standardization process. For example, the impurity estimation process has been improved and the relevance of the traces of HbA1c in the HbA0 primary standard was discovered.

For the secondary calibrators, the importance of the number of network laboratories to maintain the low uncertainty of the secondary calibrators was revealed and it was explained why a further increase of the number of network laboratories will have a relatively small impact on the uncertainty.

In addition, we derived a simple and efficient formula for the transmission of uncertainty and bias of the primary calibrators to the uncertainty of the secondary calibrators. This issue does often occur in standardization processes of diagnostic assays, but is largely unresolved in our eyes. An important aspect of this setting is to take into account the correlation between the different levels of the primary calibrators.

Last but not least, we validated the calculations according to GUM by a simulation study, which led to important improvements of the GUM approach.

Appendix A Variance components

The calculation of the uncertainty of the mean of a set of measurements, gathered in a specific design, requires a careful consideration of the respective design structure. The simplest case of such a design is the random-effects model, which can be written as

$$Y_{ij} = \mu + a_i + \varepsilon_{ij}, \quad i = i, \ldots, I, \quad j = 1, \ldots, J,$$

with Y_{ij} denoted as the measured result obtained e.g., in laboratory i, and repetition j, μ denotes the overall mean, a_i the laboratory-specific effect with mean zero and between-laboratory variance σ_a^2 and ε_{ij} the repetition effect, with mean zero and within-laboratory variance σ_ε^2. As long as no statistical tests are performed, no distribution assumptions need to be made. Note that the variance of an individual measurement is split up into the two variance sources, i.e., $\mathrm{Var}(Y_{ij}) = \sigma_a^2 + \sigma_\varepsilon^2$.

In our particular case, we are interested in the calculation of the variance of the mean $\bar{Y} = \frac{1}{I \cdot J} \sum_{i,j} Y_{ij}$.

However measurements from the same laboratory are no longer independent, as $\mathrm{Cor}(Y_{ij}, Y_{ik}) = \sigma_a^2$, therefore the variance of the mean is not simply the sum of the individual variances, divided by the number of observations, but becomes $\mathrm{Var}(\bar{Y}) = \frac{1}{I}\sigma_a^2 + \frac{1}{I \cdot J}\sigma_\varepsilon^2$. We note that the between-laboratory variation is only reduced by a factor of $1/I$, instead of $1/IJ$. The between-laboratory and within-laboratory variances can be estimated by ANOVA

estimation [16]. Define

$$\mathrm{MSA} = \frac{J}{I-1}\sum_i (Y_{i.} - \bar{Y})^2,$$

$$\mathrm{MSE} = \frac{1}{I(J-1)}\sum_{i,j} (Y_{ij} - \bar{Y}_{i.})^2.$$

An estimator of the between-laboratory variance is given by $\hat{\sigma}_a^2 = \frac{\mathrm{MSA}-\mathrm{MSE}}{IJ}$ and $\hat{\sigma}_\varepsilon^2 = \mathrm{MSE}$.

The interested reader is referred to [16], for the uncertainty calculation in more complicated models, as well as for the estimation of the variance components in cases of unbalanced data.

Appendix B Transformation of uncertainty, correlation and bias from calibrators

The derivation of Eq. (1.3) is based on the following ideas: A (non)-linear calibration curve, derived from I calibrators with assigned values c_i and measured signals s_i, may be approximated between two points $(c_i^*, s_i^*), (c_{i+1}^*, s_{i+1}^*)$ by a straight line $f(a, b, x) = a + bx$. The assigned values c_i have uncertainties $u(c_i)$, maximum bias $\delta(c_i)$ and correlation $\mathrm{Cor}(c_i, c_{i+1}) = \rho_{i,i+1}$.

A value c^*, lying within $c_i^* \le c^* < c_{i+1}^*$, might be approximated by $\hat{c} = \frac{s^*-\hat{a}}{\hat{b}}$, the value read from the straight line, with $\hat{b} = \frac{s_{i+1}^*-s_i^*}{c_{i+1}^*-c_i^*}$ and $\hat{a} = s_i^* - \hat{b}\cdot c_i^*$. An estimator of $\hat{c}$ is therefore given by

$$\hat{c} = g(c_i^*, c_{i+1}^*) = (c_{i+1}^* - c_i^*)\cdot\frac{s^* - s_i^*}{s_{i+1}^* - s_i^*} + c_i^*.$$

Defining by $s = \frac{s^*-s_i^*}{s_{i+1}^*-s_i^*} = \frac{c^*-c_i^*}{c_{i+1}^*-c_i^*}$, we obtain immediately $\frac{\partial g}{\partial c_{i+1}^*} = s$, $\frac{\partial g}{\partial c_i^*} = 1 - s$ and applying Eq. (1.1) the uncertainty component due to the uncertainty of the calibrator value of the read value c, is given by

$$u_{\mathrm{cal}}(c^*) = \dot{s}^2\cdot u^2(c_{i+1}^*) + (1-\dot{s})^2\cdot u^2(c_i^*) + \rho_{i,i+1}\cdot s\cdot(1-s)\cdot u(c_i^*)\cdot u(c_{i+1}^*).$$

The maximum bias of the value $\hat{c}$, can be calculated by regarding $\hat{c} - \hat{c}^t = g(c_i^*, c_{i+1}^*) - g(c_i^{*t}, c_{i+1}^{*t})$, where $g(c_i^{*t}, c_{i+1}^{*t})$ denotes the straight line between the true calibrator values. One can show that

$$\hat{c} - \hat{c}^t = s\cdot d(c_{i+1}^*) + (1-s)\cdot d(c_i^*) \le s\cdot\delta(c_{i+1}^*) + (1-s)\cdot\delta(c_i^*),$$

where $d(c_i)$ denotes the unknown difference between the assigned value of the calibrator and its true value and $\delta(c_i)$ the known maximum bias. Therefore, the maximum bias of the read value is approximated by

$$\delta(c^*) = s\cdot\delta(c_{i+1}^*) + (1-s)\cdot\delta(c_i^*).$$

Acknowledgements This work was done in collaboration with the IFCC Working Group on Standardization of HbA1c, comprising of: Dr. Kor Miedema, Isala Klinieken, Zwolle, The Netherlands (Chair), Prof. Andrea Mosca, Universita degli Studi di Milano, Milan, Italy (Secretary), Dr. Cas Weykamp, Queen Beatrix Hospital, Winterswijk, The Netherlands (Network Coordinator), Randie Little Ph.D., University of Missouri School of Medicine, Columbia MO, United States, David Sacks MD, Brigham & Women's Hospital, Boston MA, United States, Dr. Ian Goodall, A&RMC Austin Campus, Heidelberg VA, Australia, Tadao Hoshino, MD, Ph.D., Institute of Biopathological Medicine, Kanagawa, Japan, Jan-Olof Jeppsson MD, Ph.D., Malmoe University Hospital, Malmoe, Sweden, Dr. Gary John, Norfolk and Norwich University Hospital, Norwich, UK, Gary Myers, PH.D., Centers for Disease Control and Prevention, Atlanta GA, United States, Dr. Rita Paroni, Dr. Renata Paleari and Prof. Donatella Caruso, University of Milano, Milan, Italy.

References

1. The Diabetes Control and Complications Trial Research Group (1993) N Engl J Med 329:977–986
2. Hoelzel W, Miedema K (1996) J Int Fed Clin Chem 9:62–67
3. Jeppsson JO, Kobold U, Barr J, Finke A, Hoelzel W, Hoshino T, et al (2002) Clin Chem Lab Med 40:78–89
4. Directive 98/79/EC (1998) Official Journal L331:1–37
5. EN ISO 17511 (2002) In-vitro diagnostic medical devices - Measurement of quantities in biological samples - Metrological traceability of values assigned to calibrators and control materials. ISO, Geneva
6. EN ISO 18153 (2000) In-vitro diagnostic medical devices - Measurement of quantities in samples of biological origin - Metrological traceability of values for catalytic concentration of enzymes assigned to calibrators and control materials. ISO, Geneva
7. ISO (1993) International vocabulary of basic and general terms in metrology. ISO, Geneva
8. ISO (1995) A Guide to the Expression of Uncertainty in Measurement. ISO, Geneva
9. Cox MG, Harris PM (2004) Uncertainty Evaluation Best Practice Guide No. 6. NPL, Teddington
10. Phillips S, Eberhardt K, Parry B (1997) J Res Nat Inst Stand Technol 102:577–585
11. Finke A, Kobold U, Hoelzel W, Weycamp C, Miedema K, Jeppson JO (1998) Clin Chem Lab Med 36:299–308
12. International Committee for Standardisation in Haematology (1967) Br J Haematol 13–71
13. Miller JC, Miller JN (1986) Statistics for analytical chemistry. Ellis Horwood Limited, Chichester
14. Schubert S (2006) Quality control procedures for standardization systems of diagnostic assays. Diploma Thesis, University of Ulm, Ulm
15. NCCLS Document EP5-A (1999) Evaluation of Precision Performance of Clinical Chemistry Devices; Approved Guideline vol 19 no. 2. NCCLS, Wayne
16. Searle S, Casella G, McCulloch C (1992) Variance components. John Wiley, New York
17. Konnert A, Berding C, Arends S, et al (2006) JTEV 34:128–134