Patricia Kaiser*, Michael Spannagl, Christel van Campenhout, Yolande Lenga, Carla Siebelder and Cas Weykamp

HbA_{1c}: EQA in Germany, Belgium and the Netherlands using fresh whole blood samples with target values assigned with the IFCC reference system

DOI 10.1515/cclm-2016-0123
Received February 16, 2016; accepted April 7, 2016

Abstract

Background: External quality assessment/proficiency test (EQA/PT) organizers play an important role in monitoring the performance of HbA_{1c} measurements. With increasing quality of the assays, HbA_{1c} is increasingly used for diagnosis of diabetes and the demands on EQA/PT organizers themselves are rising constantly. EQA organizers in Germany (INSTAND), Belgium (WIV/IPV), and the Netherlands (SKML) organized a program with commutable samples and target values assigned with the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) reference system. The aim of this project was to confirm the logistic feasibility of organizing synchronically in the three countries, an accuracy-based EQA program with fresh whole blood, to investigate the performance of HbA_{1c} assays within and across countries and manufacturers, and to review the EQA acceptance limits.

Methods: Throughout 2015, ten fresh whole blood samples were supplied to the participants. Aggregated results were evaluated according to the IFCC model for quality targets at four levels: overall, per country, per manufacturer, and per country per manufacturer.

Results: Robust results in summer and winter demonstrated the feasibility of organizing an EQA with fresh whole blood samples in three countries. The overall performances, as well as the performance for each country were very similar: results fell within the IFCC criteria. Although substantial differences between results from different manufacturers were present, the performances of laboratories using tests of the same manufacturer were strikingly similar in the three countries, suggesting that the quality of HbA_{1c} assays is for the most part manufacturer-related. The improved design of the EQA program also suggested that acceptance limits for performance can be reduced to approximately 8%.

Keywords: EQA/PT; fresh whole blood; HbA_{1c}; IFCC reference system.

Introduction

Diabetes is one of the most prevalent chronic diseases. HbA_{1c} is considered as the key parameter for monitoring and, increasingly, for diagnosis. HbA_{1c} is included in various clinical recommendations and guidelines [1, 2]. It is not surprising that there is a pressure on laboratories and manufacturers to improve the quality of HbA_{1c} assays, especially as HbA_{1c} is considered as the gold standard for diagnosis [3]. External quality assessment (EQA)/proficiency test (PT) organizers play an important role in the quality management. With increasing quality of HbA_{1c} tests, the demands on EQA/PT organizers also increases: small matrix effects in processed samples and target values as derived from the mean of all laboratories are no longer acceptable in the light of good clinical practice. The best approach would be to use fresh whole blood samples with target values assigned with the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) reference system. The major bottleneck however, is a proper logistic organization: to warrant stability of the fresh whole blood (an absolute prerequisite for reliable EQA/PT) the whole chain from blood donation to analysis

***Corresponding author: Dr. Patricia Kaiser,** INSTAND e.V., Reference Laboratory, Düsseldorf, Germany, Fax: +49 211 15921356, E-mail: Kaiser@instand-ev.de
Michael Spannagl: INSTAND e.V., Reference Laboratory, Düsseldorf, Germany
Christel van Campenhout and Yolande Lenga: Scientific Institute of Public Health, Brussels, Belgium
Carla Siebelder and Cas Weykamp: ERL, Queen Beatrix Hospital, Winterswijk, The Netherlands

in all laboratories should be scheduled long in advance and be within a time frame of not more than 4 days [4]. In 2014, pilots of the national EQA organizers in Germany (INSTAND; Society for promoting quality assurance in medical laboratories e.V.), Belgium (WIV/ISP; Scientific Institute of Public Health), and the Netherlands (SKML; Foundation for Quality Assurance in Medical Laboratories) suggested this to be feasible and the decision was made to use fresh whole blood for the EQA starting from 2015. The major objectives of this paper are a) to evaluate data of five consecutive runs (10 samples) to confirm the logistic feasibility of organizing the EQA with fresh whole blood synchronically in our three countries, b) to investigate the performance of HbA_{1c} assays within and across countries and manufacturers, and c) to review the present acceptance limits for performance.

Materials and methods

Sample and logistics

To warrant stability of the samples, the whole logistic train from blood donation to submission of results should be within a working week. Therefore, blood donations (recruited from a pool of diabetic and nondiabetic volunteers; the volume of donations 250 or 500 mL; informed consent; single donations) are on Monday. On the same Monday, the blood was dispensed in 0.3 mL aliquots in vials specifically labeled for the respective national EQA organizers. On Tuesday, the samples were shipped per courier at ambient temperature to the national organizers and these forwarded the samples immediately to the participants at ambient temperature. According to International Standard Organisation (ISO) standard 17043 [5], the temperature during shipment was monitored on a random basis. The participants were asked to cool the samples on receipt, to assay the samples not later than Friday and to submit their results according to the procedure of the national EQA organizer. Results received later than Friday were excluded from evaluation. Reports were prepared and issued at the national level. There were five shipments of two sample sets in January, March, May, July, and August, respectively. Each EQA sample set included one sample in the physiological HbA_{1c} concentration range (in the following termed as "low" level) and a second sample in the pathological HbA_{1c} concentration range (in the following termed as "high" level). The low HbA_{1c} levels were between 32.4 and 41.2 mmol/mol and the high HbA_{1c} levels between 66.8 and 75.7 mmol/mol.

Stability/homogeneity testing and target value assignment

The homogeneity and the stability of the sample materials were tested according to the requirements of ISO standard 17043 [5] and ISO 13528 [6].

For the test of sample homogeneity, 10 EQA test items of sample "low" and 10 EQA test items of sample "high" were selected at random. In the 10 EQA test items of each sample, the total hemoglobin concentrations were analyzed in duplicate by colorimetry with a Sysmex XT-4000i analytical device. Total hemoglobin was the analyte of choice to test homogeneity because sedimentation of erythrocytes during the dispensing in vials from the bulk is the most likely cause of inhomogeneity. The homogeneity was considered to be adequate, if the between-samples standard deviation (s_s) of the 20 measurement results of each sample was less or equal to the 0.3-fold of the standard deviation for proficiency assessment being 1.8%.

To verify the stability of the EQA sample materials during the course of the EQA run, two sets of three EQA test items of sample "low" and two sets of three EQA test items of sample "high" were selected at random on the shipping date. One set each was analyzed in duplicate on the same day ("before" EQA), the other sets were stored at room temperature (20 °C–24 °C) and analyzed in duplicate on the EQA return date ("after" EQA). The test parameter was HbA_{1c}, which was analyzed by the routine system ADAMS™ HA-8160 from Menarini. Stability was considered to be adequate for each sample, if the difference of the means of the measurement results before and after the EQA run was less or equal to the 0.3-fold of the standard deviation for proficiency assessment being 1.8%.

In all five EQA runs, the material fulfilled our criteria for homogeneity and stability. Data for the stability of the EQA samples are given in the Supplementary Table 1.

The target values for HbA_{1c} were determined using the IFCC reference measurement procedure [7]. Measurements were performed in the IFCC network laboratories of INSTAND e.V. (Düsseldorf, Germany) by HPLC-MS [8] and by HPLC-CE in the Queen Beatrix Hospital (Winterswijk, The Netherlands). For calibration IFCC network calibrators (MCA Laboratory, Winterswijk, The Netherlands) were used. From the mean of the measurement results of the reference laboratories the target values for the EQA samples were calculated. The expanded measurement uncertainties for the target values were calculated according to the Guide to the Expression of Uncertainty in Measurement (GUM) [9]. The measurement uncertainties of all the influencing factors (e.g. uncertainty of the target value assignment of the calibrators, imprecision of LC-MS or LC-CE measurements) were estimated and their square sum root extracted by using the GUM Workbench (Metrodata GmbH, Weil am Rhein, Germany). For calculation of the expanded measurement uncertainties, a coverage factor of 2 (confidence interval of 95%) was applied.

Evaluation of EQA scheme (EQAS) results

The national EQA organizers evaluated the results of their participants according to their specific procedures. In Germany the robust mean and standard deviation were estimated according to the robust analysis "algorithm A" for data analysis in PT as given in ISO standard 13528 [6]. Belgium and the Netherlands applied classical statistics to calculate the mean and standard deviation after removing outliers.

The three organizers supplied their numbers of participants, means and standard deviations of each of the 10 samples specified for the major manufacturer groups (Menarini, Tosoh, Roche, Bio-Rad, Sebia, Siemens) to the common database. From those data the aggregated results of the 10 samples were calculated and used as the robust basis for the evaluations with the IFCC model for quality targets.

IFCC model quality targets

Aggregated results of the 10 samples are elaborated and evaluated in the context of the model for quality targets developed by the IFCC Task force HbA_{1c} [10] at four levels: overall, per country, per manufacturer, and per country per manufacturer.

Results

Evaluation of results of individual samples

In Figure 1, the HbA_{1c} EQAS results for two samples from DE, NL, and BE for five different runs split by manufacturers are demonstrated. The figure shows the differences between the measured HbA1c concentrations and the target values assigned with the reference measurement procedure in mmol/mol (denoted as absolute bias). It can be seen that in the five EQAS runs the results of the three manufacturers (Figure 1C, D, and F) vary equally from the target value, with a slightly positive and a slightly negative bias. With respect to the target value results from other manufacturers tend to a positive mean bias of 0.9 mmol/mol (Figure 1A), 2.0 mmol/mol (Figure 1B), and 1.6 mmol/mol (Figure 1E), respectively. For some manufacturers the performance over time is very stable (Figure 1A, B, C) with a mean bias varying in a range of 0.2 mmol/mol, and for others (Figure 1D, E, F) the absolute bias of the participant results to the target value, varies considerably in a range of 0.4 mmol/mol. Some manufacturers tend to have a different extent of absolute deviation from the target value for low and high HbA_{1c}-sample concentrations (Figure 1E, F), with different extents over time. The overall performance of all laboratories per sample is shown in Figure 1G. Results in the respective samples are very similar across countries and manufacturers. This indicates a reproducible robust system: robust samples logistics and targeting throughout the year. This allows to evaluate the data of the three countries on the basis of aggregated results of the 10 samples as is done in the next section.

Evaluation of aggregated results

Table 1 shows the condensed results of the 10 samples at four aggregation levels. At the highest level "Overall" the overall performance is shown: the mean bias of 526 laboratories was 0.6 mmol/mol with a between-laboratory CV of 4.1%. At the second level "Country" results are split per country and it can be seen that the 263 German laboratories had the lowest bias (+0.4 mmol/mol) and the highest between-laboratory CV (4.6%) of the three countries. At the third level "Manufacturer" results are differentiated for each of the manufacturers: e.g. the 156 laboratories using Roche measurement procedures had a mean bias of −0.2 mmol/mol and a between-laboratory CV of 3.9%. And at the fourth level of aggregation "Manufacturer/Country", results are split to both country and manufacturer. It can be seen that the Tosoh users in Germany, Belgium, and the Netherlands had a very similar positive bias of 2.0, 2.1, and 1.8 mmol/mol, respectively. Underlying detailed data are in the Supplemental data, Tables 2–4.

Figure 2 shows the results in perspective of the quality targets model of the IFCC task force for HbA_{1c}. For details of this model please go to reference [10]. The essence of the model is that the performance of HbA_{1c} can be evaluated in terms of the Sigma Metrics and the biological variation concept. Both concepts cover the main sources of analytical error: bias (y-axis) and imprecision (x-axis). In the case of the evaluation of the results of a group of laboratories, the imprecision is expressed as the between-laboratory CV. In the Sigma Metrics approach, the criterion is a total allowable error of 5 mmol/mol with a 2σ risk of not achieving this quality goal (failure in 1 of 20 assays); this criterion is met when the performance (represented by a dot) is within the triangle of the 2σ line and the x- and y-axis. In the biological variation model the criterion for optimum, desirable, and minimum performance is met when the dot is within the yellow, gray, and amber triangle, respectively.

In Figure 2A, the results at the overall, country, and manufacturer level of Table 1 are plotted in this IFCC model. It can be seen that the overall results (All), as well as the results per country [Belgium (BE), the Netherlands (NL), Germany (DE)] are very close to each other, just within the Sigma Metrics criterion, and slightly better than the overall performance seen in a College of American Pathologists (CAP) survey in the US [10]. Four manufacturers are within and two are outside the criterion. None of the country or manufacturer related performances are within even the minimum criterion of the biological variation concept although one method (C) touches this minimum criterion.

In Figure 2B, results are shown per manufacturer per country (the fourth level in Table 1). It can be seen that the performances of users of the same manufacturer are similar in the respective countries. This is especially true for the bias: Tosoh and Bio-Rad users have a relatively high bias and users of Sebia and Roche have a low bias in all countries. This consistency between countries also applies, but to a lesser degree, for the between-laboratory CV: e.g. for Tosoh users the between-laboratory CV is low in all the three countries.

Figure 3 shows a simulation of pass rates of 601 German HbA_{1c} EQAS participants applying different acceptance

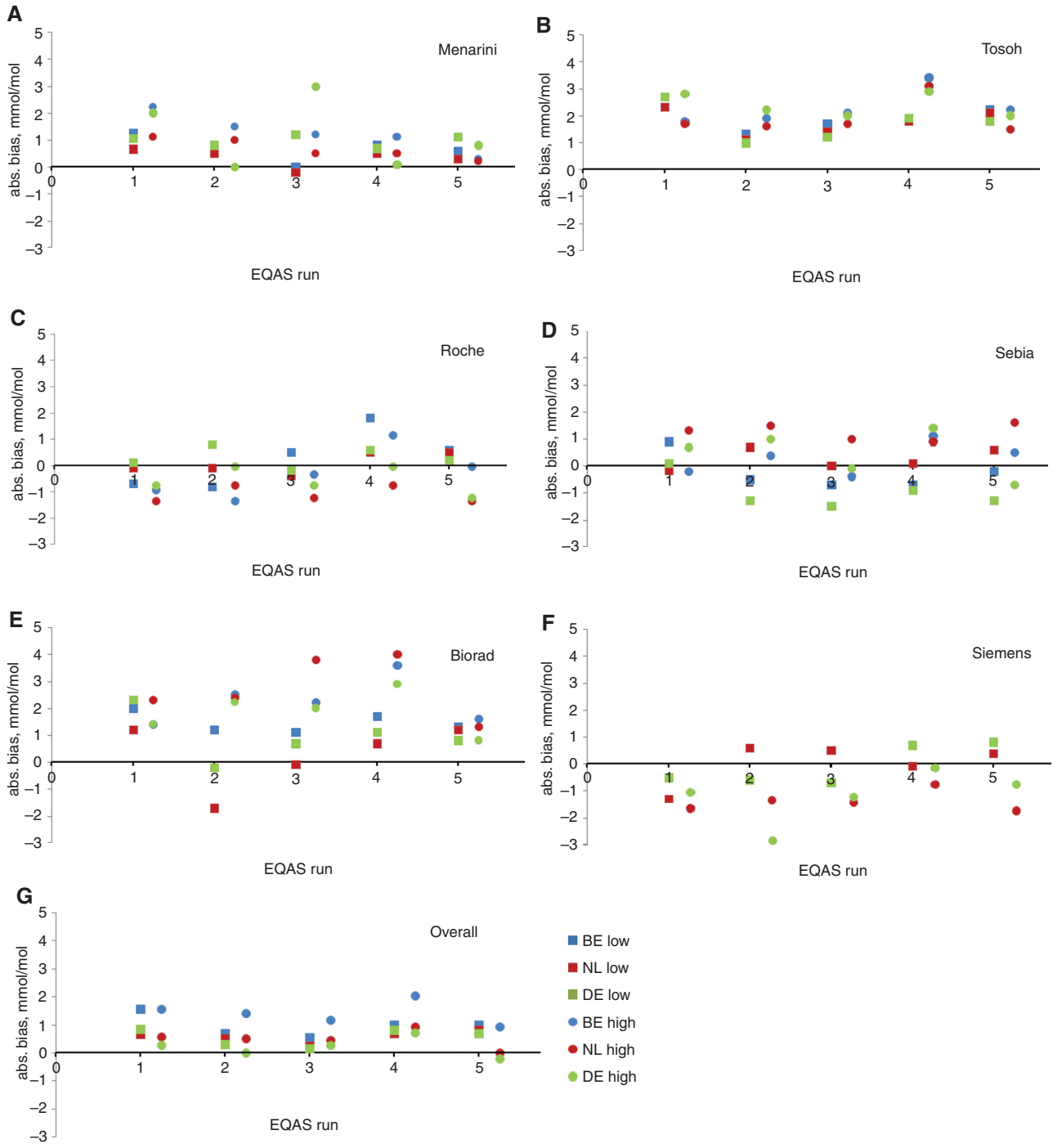

Figure 1: HbA_{1c} results of five EQAS runs for the three countries and different manufacturers.
Deviations of the participant means from the target values are given separately for Germany (DE, green symbols), the Netherlands (NL, red symbols), and Belgium (BE, blue symbols). The dots represent the physiological (low) level and the squares represent the pathological (high) level in the EQAS runs 1–5. In (A–F) the data are shown splitted by manufacturers. (G) Represents the overall mean of all the manufacturers.

limits for evaluation from 18% (the present acceptance limits in Germany according to the Guideline of the German Medical Association) [11] to a tighter limits down to 5% to demonstrate the impact of lowering limits in Germany.

The relative pass rates are given separately for the two samples sent out in this EQAS with HbA_{1c} target values of 37.4 mmol/mol ("sample low") and 70.5 mmol/mol ("sample high"), respectively, and the combined pass rate for both samples ("total"). Allowing a relative deviation of ±18% from the target value, 96.5% of the participants pass the acceptability criterion for sample "low" and 97.8% for sample "high" with a total pass rate of 96%. With a

Table 1: Aggregated results 10 samples.[b]

Group	n	Mean bias (95% CI), mmol/mol	Between-laboratory CV, %
All	526	+0.6[a] (0.5–0.7)	4.1
DE	263	+0.4 (0.3–0.5)	4.6
BE	149	+1.2[a] (1.1–1.3)	3.4
NL	114	+0.5 (0.4–0.6)	3.9
Menarini	125	+0.9[a] (0.8–1.0)	2.8
Tosoh	106	+2.0[a] (1.9–2.1)	2.6
Roche	156	−0.2 (−0.1 to −0.3)	3.9
Sebia	23	+0.1 (−0.1 to +0.3)	3.3
Bio rad	70	+1.6[a] (1.4–1.8)	4.3
Siemens	46	−0.6 (−0.4 to −0.8)	5.1
Menarini			
DE	13	+1.1[a] (0.7–1.5)	3.8
BE	70	+1.0[a] (0.8–1.2)	2.8
NL	42	+0.5 (0.4–0.6)	2.6
Tosoh			
DE	36	+2.0[a] (1.8–2.0)	3.1
BE	42	+2.1[a] (2.0–2.2)	2.3
NL	28	+1.8[a] (1.7–1.9)	2.2
Roche			
DE	127	−0.1 (0.0 to −0.2)	4.0
BE	6	0.0 (−0.4 to +0.4)	2.7
NL	23	−0.5 (−0.2 to −0.8)	3.8
Sebia			
DE	4	−0.3 (−0.7 to +0.1)	2.3
BE	17	0.0 (−0.3 to +0.3)	3.9
NL	2	+0.7 (0.4–0.7)	1.2
Bio-Rad			
DE	49	+1.4[a] (1.2–1.6)	4.5
BE	13	+1.8[a] (1.5–2.1)	3.7
NL	8	+1.5[a] (1.0–2.0)	4.6
Siemens			
DE	35	−0.6 (−0.3 to −0.9)	5.5
BE	–	–	–
NL	11	−0.7 (−0.4 to −1.0)	4.0

[a]statistically significant from target (p=0.05); [b]weighed to n where applicable.

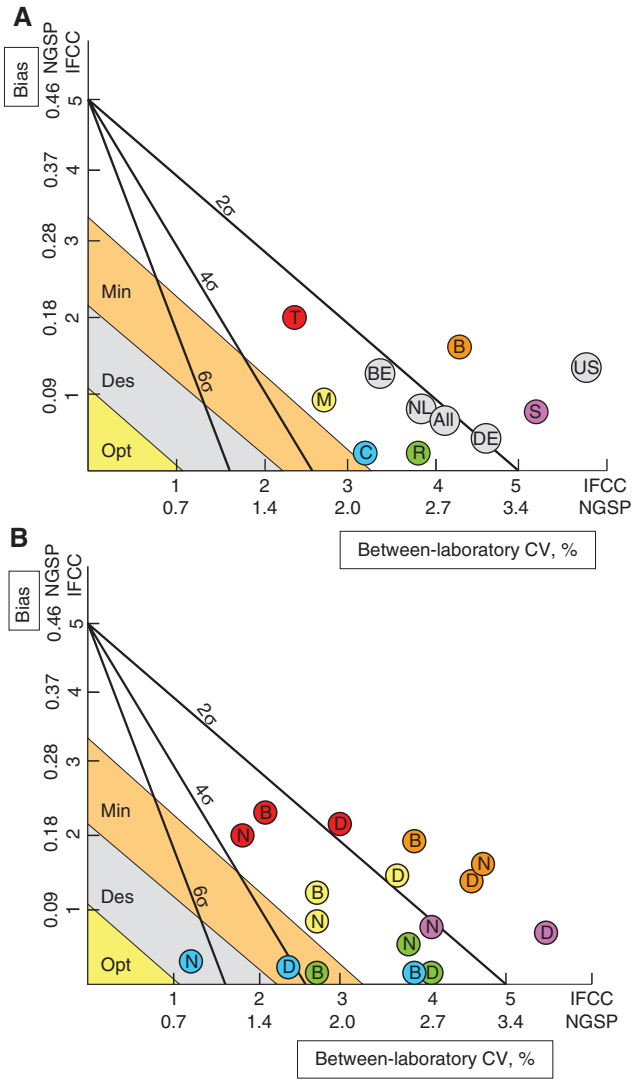

Figure 2: The IFCC quality targets model HbA_{1c} applied at country and manufacturer level.
Imprecision, expressed as between-laboratory CV on the x-axis and bias in mmol/mol on the y-axis. Criteria for quality targets are shown in colors for the biological variation concept (opt, optimum; des, desirable; min, minimum) and with line for the Sigma Metrics concept (2, 4, and 6σ). (A) Performance at the overall, country, and manufacturer level. Grey dots represent countries (DE, Germany; BE, Belgium; NL, Netherlands; All, weighted mean of DE, BE, and NL). Colored dots represent the weighted mean of the manufacturers (T, Tosoh; M, Menarini; C, Sebia; R, Roche; B, Bio-Rad; S, Siemens). (B) Performance at the manufacturer per country level; dots represent manufacturers per country (red, Tosoh; yellow, Menarini; blue, Sebia; green, Roche; orange, Bio-Rad; purple, Siemens; B, Belgium; D, Germany; N, the Netherlands).

tightening of the acceptability limit to ±10% about 88% of the participants still pass the criteria. For tighter acceptability ranges, the total pass range decreases considerably up to 57% for ±5%.

Discussion

Feasibility

The presented study demonstrates that for HbA_{1c} the implementation of EQAS with fresh whole blood samples, targeted with the IFCC reference system, is feasible in different countries with high and stable performance over the duration of several EQASs.

It was quite a challenge to organize an EQA program in three different countries by three independent EQA organizers. Stability of the samples was the prerequisite

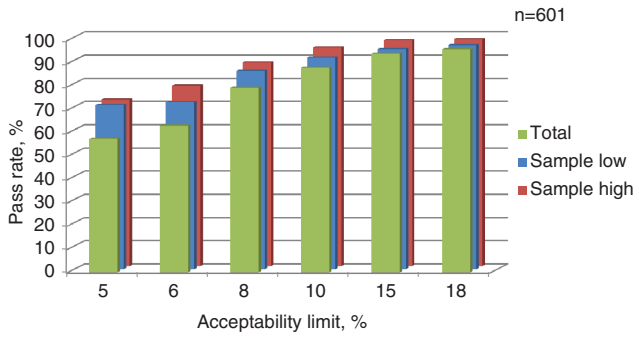

Figure 3: HbA_{1c} EQAS pass rates for different acceptability limits. Percentage pass rates of 601 participants depending on different acceptability limits (± relative % deviation from target value) set as evaluation criterion, given separately for the two EQAS samples (low=37.4 mmol/mol and high=70.5 mmol/mol) and for both (total).

for success and this was warranted by scheduling the whole process from donation of the blood to analysis in the laboratories within one working week. From our logistic experience and from the robust results we conclude that an EQA program with fresh whole blood samples shared by three independent EQA organizers is feasible.

Distribution of fresh whole blood materials in EQAS demands a high standard of the logistics of sample collection, packing, and shipment. Due to the limited stability of the target analyte HbA_{1c} of about 7 days (at ambient temperature) after collection [12], the timeline has stringently to be organized by the EQA provider. In this study, the results obtained at different times of the year from DE, NL, and BE, have been considered and no hints for seasonal difference of the overall performance are noticed. As shown in Table 1, neither the bias to the target values nor the relative standard deviations of the participant results of all three participating countries refer any kind of a seasonal effect. This may be different for other countries with more extreme ambient temperatures during shipment. This issue has to be further investigated.

Performance

Retrospective comparison of the results of the three national programs supplied an interesting view on comparability of performances per country and manufacturer. We evaluated the results in the perspective of the IFCC model for quality targets. It appeared that the mean performance in the three countries is not very different: just within the IFCC criterion and slightly better than in the CAP survey in the US. But there are substantial differences between manufacturers: some are within and some outside the criterion. And these differences between manufacturers are strikingly similar between the countries. From this, we conclude that the quality of an HbA_{1c}-assay in a European laboratory is for the major part manufacturer-related. Overall the performance of laboratories is just within the criterion of the IFCC model but there is room for improvement at the level of manufacturers and at the level of individual laboratories.

Acceptance limits

Evaluation of an EQAS raises the question, which deviation of the participant results from a target value might be acceptable for quality assurance and patient care. With respect to the fact, that HbA_{1c} is no longer used just for long-term monitoring of diabetic patients, but also for diagnosis of diabetes mellitus, these acceptance limits bear increasing significance. In the USA the acceptance limit is set to ±6%. In Europe, this limit varies considerably from ±5% in Scandinavia to more than ±10% in other countries. In Belgium, the Netherlands, and Luxemburg the acceptance criterion is 8% with a gradation of <4%, <6%, and <8% for excellent, good, and acceptable performance, respectively. In Germany, the acceptability criteria for HbA_{1c} EQAS defined by the German Guidelines is set to ±18% deviation from the target value. Use of commutable EQA material and accuracy-based evaluation of the participant results on the basis of SI-traceable target values reduces uncertainty on matrix effects and target values to a minimum and allows to reduce the acceptance limits considerably. This way, results obtained from erroneous calibrated analytical devices could be excluded more safely. From analytical and clinical point of view, tightening of the acceptability limit to ±8% could be reasonable for Germany and would allow to identify poor performing laboratories and diagnostic devices more accurately. Considering that the analytical goals for HbA_{1c} are different when results are expressed in IFCC or National Glycohemoglobin Standardization Program (NGSP) units [13], the suggested ±8% related to mmol/mol units corresponds to the 5%–6% related to NGSP units in the CAP surveys.

Conclusions

By using human fresh whole blood samples, targeted with the IFCC reference measurement procedure, we achieved to organize an EQAS meeting the highest quality standards: a category-1 EQA/PT program [14]. Moreover, we

demonstrated the feasibility to do so in parallel in three countries. It is worth to investigate if the concept can be expanded to more countries: a $EurA_{1c}$ trial as also suggested by Mosca et al. [4]. A critical prerequisite is, of course, reliable logistics.

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

References

1. American Diabetes Association. Standards of medical care in diabetes. 2011. Diabetes Care 2011;34(Suppl 1):S11–61.
2. Sacks DB, Arnold M, Bakris GL, Bruns DE, Horvath AR, Kirkman MS, et al. Guidelines and recommendations for laboratory analysis in the diagnosis and management of diabetes mellitus. Clin Chem 2011;57:e1–47.
3. Nathan D. On behalf of the Expert Committee. International Expert Committee on the role of A1c assay in the diagnosis of diabetes. Diabetes Care 2009;32:1327–34.
4. Mosca A, Paleari R, Carobene A, Weykamp C, Ceriotti F. Performance of glycated hemoglobin (HbA1c) methods evaluated with EQAS studies using fresh blood samples: still space for improvements. Clin Chim Acta 2015;451:305–9.
5. EN ISO/IEC 17043, Conformity assessment –General requirements for proficiency testing. International Organization for Standardization (Geneva, Switzerland).
6. ISO 13528, Statistical methods for use in proficiency testing by interlaboratory comparisons. International Organization for Standardization (Geneva, Switzerland).
7. Jeppson JO, Kobold U, Barr J, Finke A, Hoelzel W, Hoshino T, et al. Approved IFCC reference method for measurement of HbA1c in human blood. Clin Chem Lab Med 2002;40:78–89.
8. Kaiser P, Akerboom T, Molnar P, Reinauer H. Modified HPLC-electrospray ionization/mass spectrometry method for HbA1c based on IFCC reference measurement procedure. Clin Chem 2008;54:1018–22.
9. Evaluation of measurement data – Guide to the expression of uncertainty in measurement. BIPM Joint Comm Guides Metrol 100:2008.
10. Weykamp C, John G, Gillery P, English E, Li Linong, Lenters-Westra E, et al. Investigation of 2 models to set and evaluate quality targets for HbA1c: biological variation and sigma-metrics. Clin Chem 2015;61:752–9.
11. Richtlinie der Bundesärztekammer zur Qualitätssicherung laboratoriums-medizinischer Untersuchungen. Dtsch Arztebl 2014;38:A1583–618.
12. Mosca A, Weykamp C. Feasibility of an EQAS for HbA1c in Italy using fresh blood samples. Clin Chem Lab Med 2014;52:e151–3.
13. Weykamp C, Mosca A, Gillery P, Panteghini M. The analytical goals for hemoglobin A(1c) measurement in IFCC units and national Glycohemoglobin Standardization Program Units are different. Clin Chem 2011;57:1204–6.
14. Miller WG, Jones GR, Horowitz GL, Weykamp C. Proficiency testing/external quality assessment: current challenges and future directions. Clin Chem 2011;57:1677–80.

Supplemental Material: The online version of this article (DOI: 10.1515/cclm-2016-0123) offers supplementary material, available to authorized users.