Are Hemoglobin A_{1c} Point-of-Care analyzers fit for purpose? The story continues.

HbA_{1c} POCT: the story continues

Erna Lenters-Westra[1,2*] and Emma English[3]

[1]Isala, Department of Clinical Chemistry, Dr. Van Heesweg 2, 8025 AB Zwolle, The Netherlands.

[2]European Reference Laboratory for Glycohemoglobin, Location Zwolle, The Netherlands.

[3]University of East Anglia, Faculty of Medicine and Health, Norwich Research Park, NR4 7TJ, United Kingdom.

Corresponding Author: Erna Lenters-Westra
Isala
Department of Clinical Chemistry
Dr. Van Heesweg 2
8025 AB Zwolle
The Netherlands
Phone: +31-384247803
E-mail: w.b.lenters@isala.nl

Words:3673
Tables:4
Figures:2
Supplemental Tables:5
Supplemental Figures:6

Keywords: : HbA_{1c}, Hemoglobin A_{1c}, POCT, quality, diabetes, Hb-variant

Abbreviations: Point-of-Care (POC); Clinical and Laboratory Standards Institute (CLSI); International Federation of Clinical Chemistry and Laboratory Medicine (IFCC); National Glycohemoglobin Standardization Program (NGSP), Secondary Reference Measurement Procedures (SRMP); External Quality Assessment (EQA); European Reference laboratory for Glycohemoglobin (ERL); System Usability Scale (SUS).

Abstract

Background Point of care (POC) analyzers are playing an increasingly important role in diabetes management but it is essential that we know the performance of these analyzers in order to make appropriate clinical decisions. Whilst there is a growing body of evidence around the more well-known analyzers, there are many 'new kids on the block' with new features, such as displaying the presence of potential Hb-variants, which do not yet have a proven track record.

Methods The study is a comprehensive analytical and usability study of six POC analyzers for HbA_{1c} using Clinical and Laboratory Standards Institute (CLSI) protocols, international quality targets and certified International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) and National Glycohemoglobin Standardization Program (NGSP) Secondary reference Measurement Procedures (SRMP). The study includes precision (EP-5 and EP-15), trueness (EP-9), linearity (EP-6), sample commutability (fresh, frozen and lyophilised), interference of Hb-variants (fresh and frozen samples).

Results Only two of the 6 analyzers performed to acceptable levels over the range of performance criteria. Hb-variant interference, imprecision or variability between lot numbers are still poor in 4 of the analyzers.

Conclusions This unique and comprehensive study shows that out of 6 POC analyzers studied only 2 (The Lab 001 and Cobas B101) met international quality criteria (IFCC and NGSP), 2 (A1Care and Innovastar) were borderline and 2 (QuikReadgo and Allegro) were unacceptable. It is essential that the scientific and clinical community are equipped with this knowledge in order to make sound decisions on the use of these analyzers.

Introduction

Diabetes is a global health burden and a leading cause of morbidity and mortality worldwide. It is estimated that up to 50% of people with diabetes are currently undiagnosed, and there is an urgent need for rapid, accurate and timely diagnostic testing to identify those both with and at risk of the disease [1].

Point of care analyzers play an increasingly important role in wide range of clinical settings and there is increasing desire from clinicians to have access to more POC tests and a wider range of tests [2]. There is belief that POC enables faster clinical decision making, increased rapport with patients and reduced referrals to secondary care and subsequent healthcare costs. Over 50% of primary care physicians surveyed by Horwick et al (2014) wanted increased access to HbA_{1c} POC testing, bespeaking a clear demand for HbA_{1c} POC [3]. However, Jones et al (2013) also highlighted an apparent nervousness amongst primary care physicians around the accuracy of POC testing [4]. Coupling the desire for increased HbA_{1c} POC availability and the prudent concerns on quality it is essential that we understand how well POC HbA_{1c} analyzers perform.

Understanding the quality of POC testing has been a topic of key interest for over a decade with the stark message of 6 out of 8 analyzers not meeting the accepted quality criteria in 2010 [5]. Since this seminal study, there have been numerous evaluations of POC HbA_{1c} performance with a focus around the more common analyzers [6]. External quality assessment (EQA) provides a snap shot of 'real world' data on the performance of POC analyzers, although only a fraction of analyzers in use are currently enrolled in EQA schemes [7, 8].

Whilst there is a growing body of evidence around the more well-known analyzers, there are many 'new kids on the block' with new features, such as displaying the

presence of potential Hb-variants, which do not have a proven track record. Clinicians and laboratory scientists need robust and rigorous evaluation data on performance, acceptability and usability of POC analyzers to support informed decision making around use of POC testing analyzers.

The evaluation of POC analyzers is not without issue. Whilst many HbA_{1c} POC analyzers are scaled down versions of laboratory analyzers they have their own unique differences which require adaptations in order to complete a comprehensive evaluation. One key issue is that several POC analyzers are not compatible with frozen or lyophilized blood samples, meaning conventional evaluation protocols cannot be directly applied.

Whilst there are numerous method comparisons published, it is important to note that these are often single comparisons to routine laboratory methods (which will have their own imprecision and bias to consider), which provide insight into local performance but do not provide a robust picture of performance against internationally accepted secondary reference measurement procedures (SRMPs) or international quality criteria [9,10].

This study aims to understand the performance of a range of POC HbA_{1c} analyzers using a rigorous evaluation protocols which examines issues such as; interference from Hb-variants with fresh and frozen samples, sample compatibility (fresh, frozen and lyophilised) and system usability whilst comparing analytical performance to IFCC and NGSP SRMPs.

Materials and Methods

Analyzers evaluated The six POC analyzers included in this study and their key characteristics, are summarized in Table 1. The choice of analyzer was both

manufacturer led (Allegro, QuikReadgo and The lab 001) and investigator led (Cobas B101, InnovaStar and A1Care). The manufacturers of the last three analysers were approached by the authors as previous evaluations had highlighted some performance issues. An initial familiarization protocol was undertaken with each instrument and the results were shared with the manufacturers to enable them to decide if they wished to continue to a full evaluation. This supports a collaborative approach to working with manufacturers with the aim to improve quality. In some cases when a product is new in development, feedback at an early stage will enable further development before a product is brought to full evaluation, saving time and resources [11]. Six analyzers were fully evaluated, however two new analyzers were not ready yet for a full evaluation.

Imprecision study (EP-5 and EP-15) The CLSI EP-5 protocol was used to investigate assay imprecision. It is known that some POC methods have a bias with frozen material and as it is not known if frozen samples have an impact on the imprecision, also EP-15 was performed with two fresh patient samples (HbA_{1c} values of 48 and 75 mmol/mol). Both samples were analyzed five-fold for 5 days. CVs were also calculated on the basis of the duplicates of the fresh patient samples in the EP-9 protocol.

Method comparison (trueness; EP-9) The CLSI EP-9 protocol was performed with 40 fresh patient samples and the data was used to investigate the bias between each instrument and 4 SRMPs (n=40, 8 samples per day for 5 days, duplicate measurements). Values were assigned with 4 IFCC and NGSP Certified SRMPs [12,13]:

- Roche Tina-quant Gen.3 HbA_{1c} on Cobas c513, immunoassay, IFCC and NGSP certified (Roche Diagnostics);
- Premier Hb9210, affinity chromatography HPLC, IFCC and NGSP certified (Trinity Biotech);
- Tosoh G8, cation-exchange HPLC, IFCC certified (Tosoh Bioscience);
- Abbott Enzymatic method on Alinity, IFCC and NGSP certified (Abbott Diagnostics).

Linearity (EP-6) Linearity was assessed using the CLSI EP-6 protocol. After adjustment for Hb concentration, patient samples with a low HbA_{1c} value and a high HbA_{1c} value were mixed in incremental amounts to generate a series of equally spaced samples over a broad HbA_{1c} concentration range. Eleven samples were analyzed in duplicate in one day. The samples were made fresh and then frozen at -80 °C degrees until analysis. Whilst some analyzers display a bias with frozen samples this is generally a consistent bias and therefore these can still be used to assess linearity.

The difference between the fitted values of the best polynomial line and the regression line for the 11 samples were compared. CLSI states for EP-6 that goals for linearity should be derived from goals for bias, and should be less than or equal to these goals [14]. The IFCC Task Force on Implementation of HbA_{1c} Standardization has set an TAE of 10% at an HbA1c concentration of 50 mmol/mol (19). Taking into account the whole clinical relevant range, we have set a TAE of 6 mmol/mol with a nonlinearity budget of 50% (=3 mmol/mol). If the deviation exceeds allowable nonlinearity (3 mmol/mol) the data was considered nonlinear.

Hemoglobin variants AS, AC, AE, AD, elevated A2 and HbF

Twenty patient samples of each heterozygous Hb variant, from our frozen whole blood biobank, were measured on each of the different POC analyzers. Values were assigned using an IFCC calibrated boronate affinity HPLC (Premier Hb9210). For samples with increased HbF, HbA_{1c} values were assigned using an IFCC calibrated cation-exchange HPLC (Menarini HA8180V, Diabetes Mode, (frozen) and Tosoh G8 (fresh)). Percentage HbF (3.5 – 42.0%) was determined using the Sebia Capillarys 2 Flex Piercing Hemoglobin program.

In addition to the frozen samples, 16 HbAS, 7 HbAC, 5 HbAD, 9 HbAE and 4 HbF (9.1, 20.5, 20.8 and 27.5%) fresh Hb-variant samples were also analyzed on each analyser as two of the analyzers (InnovaStar and QuikReadgo) showed a bias with frozen samples.

Any bias observed due to the presence of variants is a compound of both the bias in normal samples (identified by the EP-9 protocol) and the bias associated with the variant. In order to account for this, the results were adjusted for the bias found during EP-9 (Premier Hb9210 for fresh and frozen Hb-variant), thus any residual bias would be due to the Hb-variant. For the two analyzers that also display a bias with frozen samples, the bias correction was done using the 24 frozen EQA samples rather than the EP-9 data. Whilst this is not a perfect solution it avoids a two-step correction.

For an Hb variant to be considered as not causing a clinically relevant interference, the results of the Hb variant should fall within a defined scatter line of ±10% (SI units) of the regression line derived from the comparison of the test instrument and the Premier Hb9210 with the nonvariant samples (HbAA).

Schiff Base, Icteric samples, different Hemoglobin concentrations

In order to create labile samples (Schiff Base) 12, 16 and 20 mg/ml glucose was added to aliquots of high, medium and low HbA_{1c}, EDTA samples. Icteric samples were generated by removing the plasma of a non-icteric sample and replacing with plasma with 219, 236 and 258 µmol/L bilirubin (icteric sample), again at three different HbA_{1c} levels. Similarly addition or removal of plasma was used to create a range of samples with varying hemoglobin levels. The samples were stored frozen at –80 °C until analysis. A mean relative difference of ± 10% (in SI units) pre and post treatment of the samples, was considered a significant interference.

EQA Programs (assessing sample commutability) In order to assess sample commutability, samples from both the IFCC Certification Program for manufacturers [15] and the European Reference Laboratory for Glycohemoglobin (ERL) EQA Program [16] were used to provide data on frozen and lyophilized samples respectively.

System Usability Scale (SUS)

This study included a SUS score generated by the two technicians who performed the evaluation study. SUS is a simple technology diagnostic tool consisting of ten questions which gives a global view of subjective assessment of the usability of the device tested [17].

A SUS score >81 can be considered as excellent, between 71 and 80 as good, between 52 and 70 okay and <51 poor [18].

Defining the quality criteria

International Quality Standards This study used the previously published global guidance on acceptable quality and performance criteria for HbA1c testing from the IFCC Task Force on Implementation of HbA_{1c} Standardization [19].

NGSP Manufacturer Certification Criteria Thirty six out of 40 results must be within 5% (relative) of an individual NGSP SRMP to pass certification [20].

Statistical analysis Calculations were performed using Microsoft® Excel 2016 (Microsoft Corporation). Statistical analyses were performed using Analyse-It®, version 5.40 (Analyse-It Software) and EP Evaluator Release 12 (Data Innovations).

Results

Imprecision (EP-5 and EP-15) Table 2 displays the CVs derived from both EP-5 and EP-15 and the duplicates from the EP-9 protocol. Only the Lab 001 and the InnovaStar achieved the performance criteria of <3% CV (SI units) (<2% in NGSP units) across all protocols and both high and low HbA_{1c} levels [21, 22]. The Lab 001 actually achieved <2 % CV (SI units) showing very low levels of imprecision. However the Allegro failed to achieve <3% CV (SI units) in any protocol at either level, with CVs as high as 4.2%, showing unacceptable levels of imprecision. The A1Care had mixed results, performing better with fresh samples and at higher HbA_{1c} values. The B101 also had mixed results, with better performance at higher HbA_{1c} values. The QuikReadgo met the criteria in both EP-5 and EP-15 with little difference between fresh and frozen samples. However, the performance with the duplicates from EP-9 was mixed with one lot failing and one lot passing.

Method comparison (trueness; EP-9) Table 3 details the results of the method comparison study and also NGSP pass/fail rate. From these data it is clear that all analyzers suffered some degree of bias, with some such as the Allegro showing this across multiple levels and between lot numbers. The data for the individual POC analyzers versus the individual SRMPs in NGSP units is available in supplemental table 1. From this table and table 3 it can be seen that only the Cobas B101 passed the NGSP criteria with two lot numbers compared to all four individual SRMPs and that the QuikReadgo and the Allegro failed the NGSP criteria with both lot numbers for all four individual SRMPs. Figure 1 shows the regression lines for each POC device versus the mean of the SRMPs. All analyzers suffered some degree of bias. The Lab 001 and the Cobas B101, had the least bias across the lots and the HbA1c range and all other POC analyzers had a statistically significant difference either between the lot numbers or at different HbA_{1c} levels or both and showed a large dispersion around the deming regression line compared to the mean of the SRMPs.

Linearity (EP-6) Supplemental table 2 details the results of the linearity study. The maximum deviation is shown between the fitted values of the best polynomial line and the regression line for the 11 samples. If the deviation exceeds allowable nonlinearity (3 mmol/mol) the data was considered nonlinear. Based on this criteria all POC analyzers were linear except for the Cobas B101 and the InnovaStar. However, the detection limit of the InnovaStar was >30 mmol/mol. Excluding the lowest sample for the calculations showed that the InnovaStar was linear. The HbA_{1c} result of the highest sample for the Allegro was above the detection limits (> 130 mmol/mol) therefore the linearity was assessed with 10 samples instead of 11.

Hemoglobin variants AS, AC, AE, AD, elevated A2 and HbF Table 4 shows the mean relative difference of the frozen and fresh Hb-variants samples and supplemental figure 1 to 6 shows the graphs of the interference of Hb-variants for frozen and fresh Hb-variants for the different POC analyzers. All methods, except for the A1Care, had an interference with one or more of the Hb-variants with frozen or fresh samples (mean relative difference was >10%). All Hb-variants were detected and correctly identified by the Lab 001 via a S-, C-, D-, E- or F-window.

Schiff Base, Icteric samples, different Hemoglobin concentrations

None of the POC analyzers showed an interference for schiff base, icteric samples or different hemoglobin concentrations. Supplemental tables 3 to 5 show the data.

Sample commutability

To investigate the impact of sample type in relation to different clinical applications, fresh (EP-15 and EP-9), frozen (EP-5 and IFCC certification program samples) and lyophilized (ERL EQA scheme) samples were compared. Figure 2 (panel A and B) show the data from EP15 and EP-9 (red circles) representing all fresh samples, IFCC certification samples (blue circles) representing all frozen samples and the ERL EQA (green circles) representing lyophilized samples. In addition panel B shows EP-5 and EP-9 data (grey circles) in order to assess the impact of using EP-5 versus EP-15 for routine method evaluations. The EP-5 and EP-15 studies were both used to compare performance with fresh and frozen samples and for the majority of analyzers the EP-15 evaluation provided sufficient data to assess performance.

The data clearly demonstrates that lyophilized material was only commutable with The Lab 001, all other POC analyzers showed a large positive bias when analyzing

lyophilized material. Frozen material was not commutable with the InnovaStar and the QuikReadgo. When using fresh patient samples all, except the Allegro, passed the IFCC criteria of having a sigma >2 at an HbA_{1c} concentration of 50 mmol/mol. Conversely the Allegro actually performed better when using frozen samples instead of fresh samples.

SUS scores

Table 1 shows the SUS scores of the different POCT analyzers. The usability of all the POCT analyzers was good to excellent (mean SUS score > 80) except for the QuikReadgo (mean SUS score was 60).

Discussion

What progress has been made?

In the study of 2014 the InnovaStar showed an interference with fresh patient samples, which was likely due to the instrument being calibrated using frozen samples [23]. The previous publication led the manufacturer to switch to fresh patient samples, which are available from the ERL, to calibrate their cartridges resulting in lower bias in fresh samples [24].

The Lab 001 device is new to the market and the sigma graphs show excellent performance however there was a small bias at higher HbA_{1c} levels. Paradoxically, had the imprecision of the instrument been higher than the bias would not have been detected as the confidence intervals would be wider. This device shows that the field of POCT has moved on and quality improvements are possible.

There are still significant issues with the performance of some analyzers

The key issues we still see are: a) lot to lot variability, b) high imprecision, and c) significant interference from variants. Four out of the 6 evaluated analyzers still do not demonstrate acceptable and/or consistent performance.

The InnovaStar had poor performance between lot numbers that was not acceptable. The new to the market QuikReadgo also showed a statistically significant difference between the lot numbers, and failed to meet the NGSP criteria with either lot number when compared to any of the 4 SRMPs. The Allegro also showed similarly poor performance. It should be noted that both the Allegro and A1Care were evaluated in a previous study (data not presented) in which the results were acceptable, to good. As the data for certain elements of the current study was acceptable it is likely that there is an inconsistency in the manufacturing chain that needs to be identified. The considerable variability in performance across the analyzers, albeit less than in earlier studies, shows that there is still work to be done [6].

A key issue with POC analyzers still appears to be interference from variant hemoglobins. A complicating factor when evaluating Hb-variants is the fact that some analyzers (QuikReadgo and Innovastar) are not compatible with frozen samples. This study addresses this issue with the use of fresh Hb-variant samples. However it was not possible to obtain as wide a range or number of fresh Hb-variant samples for investigation as would be desirable. Explaining the interferences seen in these analyzers, which are nearly all immunoassay, is difficult, why would frozen and fresh samples perform so differently? Why are they causing an interference at all when the epitopes that the antibodies bind to do not contain the mutations that cause the Hb variant? HbE for example is an inherited single base mutation at codon 26 of the beta-globin gene, leading to substitution of glutamic acid (46 amino acid of the beta chain) for lysine which should not, theoretically, interfere with the antibodies used in the POC

analyzers. A possible explanation could be that the mutation causes folding of the hemoglobin molecule in such a way that position 46 is then very close to the first 4 amino-acids of the beta chain and interacts with the antibodies used in the immunoassay [25]. Alternatively this may be due to differences in the immobilization of the antibodies which lead to differences in the surface chemistry and thus the binding of the antibody.

The Lab 001 suffers from interference with fresh HbAS samples, but less so with frozen samples which are easier to obtain for method development. This does potentially pose a problem for patient samples, however this is mitigated by the fact that Lab 001 as a capillary electrophoresis method is capable of identifying the presence of a variant – unlike most other POC analyzers.

It is important to clarify that some of the manufacturers do clearly state that the presence of variants may alter the HbA_{1c} results however these claims and the findings of this study do not always correlate.

It is not all about analytical performance

Whilst many evaluations focus on analytical performance, it is important to consider the wider context of the use of POC analyzers. The usability/user-friendliness of each of the analyzers was assessed and found to be variable.

A crucial factor in the practical, clinical usability of an instrument is how long it takes to generate a result, with the benefit of providing real time results often touted as a key selling point for POC analyzers. The time from a ‘cold start’ (turning the power on and warming reagents if needed) to a result was assessed. The range of time needed was wide at ~3.0 minutes for The Lab 001 to ~16.5 minutes for the InnovaStar. This is

important information for users of the analyzers as they may have little advance warning that a test may need to be undertaken.

Key messages

This complex and detailed evaluation provides a comprehensive overview of 6 HbA_{1c} POC analyzers. Whilst there are areas of excellence in performance there are still significant areas for improvement with the performance of some being unacceptable. It is possible for an analyzer to meet certification criteria for the IFCC and/or NGSP and perform well in one evaluation and then perform very poorly in subsequent evaluations. From a clinical and scientific perspective this is alarming. It is essential that performance of an analyzer is stable, especially with increased use for both monitoring and diagnosis of people with diabetes. Four of the analyzers in this study showed highly variable performance which is not acceptable.

It is unclear why such discrepant results are seen when fresh or frozen samples are used, especially as this is not commonly seen with routine laboratory analyzers. Whilst many evaluations and method development often necessitate the use of frozen samples, it is rare that a POCT would be used with anything other than fresh samples. We have shown here that there can be marked differences in performance with each sample type.

One way to identify variability in performance is through the use of EQA schemes. The authors strongly advocate the use of EQA to identify ongoing performance issues, and although POC analyzers are often exempt from the need to participate in EQA it is a valuable and powerful tool for monitoring performance. A caveat to this is that POC analyzers may not be able to utilize the frozen or lyophilized samples often used in EQA schemes. None of the manufacturers claim in their information for users that

lyophilized samples can be used and this is supported by the data from the ERL-EQA samples (see figure 2), EQA program leads need to be cognoscente of this issue and work towards providing commutable samples for POC analyzers.

As discussed earlier, the interference from Hb-variants in a number of the analyzers is perplexing. The disparity in results seen between fresh and frozen samples is of concern as many manufacturers will likely develop their methods using frozen samples but in the ‘real world’ setting where fresh samples are used, the variants pose a potential unseen problem. Not all manufacturers are accurate in their claims for Hb-variant performance.

Author Contributions All authors confirmed they have contributed to the intellectual content of this paper and have met the following three requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Acknowledgements The authors wish to thank Agnes den Ouden for her support with the analysis of the samples. They also wish to thank DiaSys Diagnostic Systems, Holzheim, Germany, Roche Diagnostics, Almere, The Netherlands, Arkray Europe, Amsterdam, The Netherlands, i-SENS, Seoul, South Korea, Aidian, Espoo, Finland and Nova Biomedical, Waltham, US for their financial support and for donating reagents and analyzers used in this study.

Role of the sponsor:

The manufacturers provided all analyzers and reagents at no cost and financially supported this study. The evaluating protocol was designed by the authors. The manufacturers played no role in the review and interpretation of data or preparation or approval of manuscript and had no rights of refusal for publication of the data.

References

1. IDF diabetes Atlas (Accessed August 2020) (https://www.diabetesatlas.org/en/).
2. Schols AMR, Dinant G, Cals WL Point-of-care testing in general practice: just what the doctor ordered? British Journal of General Practice 2018;68:362-3.
3. Howick J, Cals JWL, Jones C, et al. Current and future use of point-of-care tests in primary care: an international survey in Australia, Belgium, The Netherlands, the UK and the USA. BMJ Open 2014;4:e005611. doi: 10.1136/bmjopen-2014-005611.
4. Jones CH, Howick J, Roberts NW, et al. Primary care clinicians' attitudes towards point-of-care blood testing: a systematic review of qualitative studies. BMC Fam Pract. 2013;14:117.
5. Lenters-Westra E, Slingerland RJ. Six of eight Hemoglobin A1c point-of-care analyzers do not meet the general accepted analytical performance criteria. Clin Chem 2010;56:44-52.
6. Hirst JA, McLellan JH, Price PP, English E, Feakins BG, Stevens RJ and Farmer AJ. Performance of point-of-care HbA_{1c} test analyzers: implications for use in clinical practice- a systematic review and meta-analysis. Clin Chem Lab Med 2017;55:167-80.
7. NGSP website note CAP survey data. http://www.ngsp.org/CAP/CAP20a.pdf (Accessed August 2020).
8. EurA1c Trial Group. EurA1c: The European HbA_{1c} Trial to Investigate the Performance of HbA_{1c} Assays in 2166 Laboratories across 17 Countries and 24 manufacturers by Use of the IFCC Model for Quality Targets. Clin Chem 2018;64:1183-92.

9. Nathan DM, Griffin A, Perez FM, Basque E, Do L, Steiner B. Accuracy of a point-of-care hemoglobin A1c assay J Diabetes Sci Technol. 2019;13:1149-53.
10. Lenters-Westra E, English E. Investigation of Quality of POCT HbA_{1c} analyzers: What Are Our Next Steps? J Diabetes Sci Technol. 2019;13:1154-57.
11. Lenters-Westra E. Independent evaluation using fresh patient samples under real clinical conditions is vital for confirming the suitability and marketability of any new HbA_{1c} assay. An example. Clin Chem Lab Med 2018;56:e157-e159.
12. Approved laboratories of the IFCC network laboratories for HbA_{1c} 2020. https://information.ifcchba1c.org/sustainable-implementation/reference-laboratories/list-reference-laboratories (Accessed August 2020).
13. NGSP Network Laboratory Members. www.ngsp.org/network.asp (Accessed August 2020).
14. CLSI Document EP6-A. Evaluation of the linearity of quantitative measurement procedures: a statistical approach; Approved guideline. CLSI, 940 West Valley Road, Suite1400, Wayne, PA 19087-1898 USA, 2003.
15. Information IFCC certification program. https://www.ifcchba1c.org/UserData/Files/Certificate%20Information.pdf (Accessed August 2020).
16. Information ERL EQA program. http://www.euroreflab.com/ (Accessed August 2020).
17. Information System Usability Scale (SUS). https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html (Accessed August 2020).
18. Interpretation of SUS scores. https://measuringu.com/interpret-sus-score/ (Accessed August 2020).

19. Weykamp CW, John WG, Gillery P, English E, Ji L, Lenters-Westra E, Little RR, Roglic G, Sacks DB, Takei I. Investigation of two models to set and evaluate quality targets for HbA_{1c}: Biological Variation and Sigma-metrix. Clinical Chemistry 2015;61:752-759.
20. NGSP manufacturer performance criteria. http://www.ngsp.org/critsumm.asp (Accessed August 2020).
21. Sacks DB, Arnold M, Bakris GL, Bruns DE, Horvath AR, Sue M et all. Guidelines and recommendations for Laboratory Analysis in the Diagnosis and Management of Diabetes Mellitus. Diabetes Care 2011;34:e62-e99.
22. Weykamp CW, 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;57:1204–1206.
23. Lenters-Westra E, Slingerland RJ. Three of Seven Hemoglobin A1c Point-of-Care Analyzers Do Not Meet the Generally Accepted Analytical Performance Criteria. Clinical Chemistry 2014;60:1062-72.
24. Information for manufacturers for support and services by ERL. https://information.ifcchba1c.org/sustainable-implementation/manufacturers/support-and-services.aspx (Accessed August 2020).
25. Lenters-Westra E, Strunk A, Campbell P, Slingerland RJ. Can the Afinion HbA_{1c} Point-of-Care instrument be an alternative method for the Tosoh G8 in the case of Hb-Tacoma? Scan J Clin and Lab Invest 2016;18:1-6.

Legends

Figure 1: EP-9 data for each of the POC analyzers when compared to the mean of the SRMPs with two lot numbers. Panel A - the A1Care, B - the Lab 001, C - the Cobas B101, D - the QuikReadgo, E - the InnovaStar, F - the Allegro

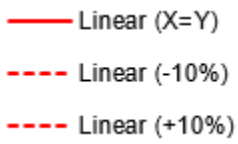

Figure 2: Sigma metrics graphs showing the impact of different sample types on the ability to meet the International quality performance criteria. Figure 2A shows EP- 15 and EP- 9 (red) values compared to frozen (blue) and lyophilized samples (green). Figure 2B shows EP-5 (grey) data compared to frozen and lyophilized samples showing minimal difference in performance between EP-5 and EP-15 protocols. Panel A - the A1Care, B - the Lab 001, C - the Cobas B101, D - the QuikReadgo, E - the InnovaStar, F - the Allegro

Analyzer	Manufacturer	Sample Volume (µl)	Analysis Time (min)	Time from a cold start to first result (min)	Method Principle	Weight (kg)	Dimensions (W x H x D mm)	Operating temperature	Storage temperature cartridges	Hb-variant visible?	Mean SUS score ([a]) N=2
A1Care	I-Sens	2.5	4.2	≈13.7	Enzymatic	3.8	290 x 250 x 130	10 - 32 °C	1 - 30 °C	No	88
Cobas B101	Roche Diagnostics	2	6.0	≈9.5	Immuno-assay	2.0	135 x 184 x 234	15 - 32 °C	2 - 30 °C	No	95
InnovaStar	DiaSys	10	6.5	≈16.5	Immuno-assay	4	200 x 150 x 170	15 - 35°C.	2 - 8 °C	No	80
The Lab 001	Arkray	1.5	1.5	≈3.0	Capillary electrophoresis	10	220 x 298 x 330	10 - 30 °C	2 - 30 °C	Yes	91
QuikReadgo	Aidian	1	6.0	≈7.4	Immuno-assay	1.7	145 x 155 x 270	15 - 35°C.	2 - 8 °C, 2 months at 18 -25 °C	No	60
Allegro	Nova Biomedical	1.5	6.5	≈9.2	Immuno-assay	9.1	203 x 381 x 381	15 - 32 °C	2 - 8 °C	No	84

[a]= System Usability Scale

Table 1: Analyzers included in the study and their key characteristics.

	CV (%) in SI units (sample value)	CV (%) in NGSP units (sample value)
A1Care		
EP-5 Frozen samples	**3.5** (46.6 mmol/mol) 2.7 (71.3 mmol/mol)	**2.4** (6.41%) **2.1** (8.68%)
EP-15 Fresh samples	**3.1** (46.3 mmol/mol) 1.9 (79.8 mmol/mol)	1.9 (6.38%) 1.5 (9.44%)
Lot number A[a]	**3.2**	2.3
Lot number B[a]	2.9	2.0
Cobas B101		
EP-5 Frozen samples	**3.1** (44.7 mmol/mol) 1.7 (71.4 mmol/mol)	**2.2** (6.24%) 1.3 (8.69%)
EP-15 Fresh samples	**3.6** (44.2 mmol/mol) 1.2 (81.5 mmol/mol)	**2.3** (6.20%) 0.9 (9.61%)
Lot number A[a]	2.3	1.6
Lot number B[a]	1.6	1.1
InnovaStar		
EP-5 Frozen samples	2.0 (48.7 mmol/mol) 2.4 (75.5 mmol/mol)	1.3 (6.61%) 1.8 (9.06%)
EP-15 Fresh samples	1.5 (45.1 mmol/mol) 1.4 (84.0 mmol/mol)	1.0 (6.28%) 1.1 (9.84%)
Lot number A[a]	2.4	1.7
Lot number B[a]	1.5	1.0
The Lab 001		
EP-5 Frozen samples	1.4 (46.1 mmol/mol) 1.8 (72.0 mmol/mol)	0.9 (6.37%) 1.5 (8.74%)
EP-15 Fresh samples	1.7 (45.4 mmol/mol) 1.1 (82.4 mmol/mol)	1.1 (6.30%) 0.8 (9.69%)
Lot number A[a]	1.8	1.2
Lot number B[a]	1.6	1.1
QuikReadgo		
EP-5 Frozen samples	2.5 (50.9 mmol/mol) **3.0** (82.7 mmol/mol)	1.9 (6.79%) **2.1** (9.71%)
EP-15 Fresh samples	2.5 (43.5 mmol/mol) 2.1 (86.8 mmol/mol)	1.5 (6.14%) 1.8 (10.08%)
Lot number A[a]	**3.5**	**2.5**
Lot number B[a]	2.7	1.9
Allegro		
EP-5 Frozen samples	**4.2** (44.3 mmol/mol) **4.1** (70.0 mmol/mol)	**2.8** (6.23%) **2.8** (8.56%)
EP-15 Fresh samples	**4.2** (43.6 mmol/mol) **3.8** (85.4 mmol/mol)	**3.0** (6.16%) **3.0** (9.96%)
Lot number A[a]	**3.4**	**2.4**
Lot number B[a]	**3.6**	**2.6**

[a]based on the duplicates in EP-9

Table 2: Imprecision results based on EP-5, EP-15 and on the duplicates in EP-9. Red: fail performance targets (CV <3% in SI units and <2% in NGSP units).

	Bias at different HbA_{1c} levels (mmol/mol)			
	30 mmol/mol (95% CI)	**48 mmol/mol (95% CI)**	**75 mmol/mol (95% CI)**	**Pass NGSPcriteria? X out of 4 SRMP***
A1Care				
Lot A vs mean SRMP	**31.7** (30.62 to 32.74)[a]	**47.9** (47.29 to 48.57)	**72.3** (70.92 to 73.69)[a]	
Lot B vs mean SRMP	**31.6** (30.42 to 32.73)[a]	**48.9** (48.15 to 49.60)[a]	**74.8** (73.67 to 75.98)	
Lot A vs Lot B	**30** (29.5 to 31.2)	**49** (48.3 to 49.4)*	**77** (75.5 to 77.5)[a]	
Lot A				2/4
Lot B				1/4
Cobas B101				
Lot A vs mean SRMP	**30.1** (28.81 to 31.47)	**47.**6 (46.82 to 48.34)	**73.7** (72.95 to 74.55)[a]	
Lot B vs mean SRMP	**29.3** (28.29 to 30.29)	**47.2** (46.63 to 47.85)[a]	**74.2** (73.41 to 74.93)[a]	
Lot A vs lot B	**29.0** (28.0 to 30.7)	**48.0** (46.9 to 48.4)	**75.0** (74.2 to 76.2)	
Lot A				4/4
Lot B				4/4
InnovaStar				
Lot A vs mean SRMP	**30.4** (29.63 to 31.21)	**47.3** (46.88 to 47.81)[a]	**72.7** (71.57 to 73.87)[a]	
Lot B vs mean SRMP	**28.6** (27.73 to 29.52)[a]	**46.2** (45.68 to 46.80)[a]	**72.7** (71.88 to 73.45)[a]	
Lot A vs Lot B	**28.1** (27.73 to 28.53)[a]	**46.7** (46.49 to 46.94)[a]	**74.6** (74.04 to 75.13)	
Lot A				4/4
Lot B				2/4
The Lab 001				
Lot A vs mean SRMP	**29.7** (28.73 to 30.58)	**47.1** (46.58 to 47.55)[a]	**73.2** (72.30 to 74.06)[a]	
Lot B vs mean SRMP	**30.0** (29.44 to 30.60)	**47.4** (47.07 to 47.71)[a]	73.4 (72.75 to 74.13)[a]	
Lot A vs Lot B	**29.9** (29.47 to 30.36)	**47.8** (47.57 to 48.08)	**74.7** (74.26 to 75.12)	
Lot A				3/4
Lot B				4/4
QuikReadgo				
Lot A vs mean SRMP	**28.4** (27.17 to 29.69)[a]	**46.5** (45.84 to 47.21)[a]	**73.7** (71.65 to 75.66)	
Lot B vs mean SRMP	**30.3** (28.70 to 31.92)	**48.3** (47.30 to 49.36)	**75.4** (73.79 to 76.93)	
Lot A vs Lot B	**32.0** (30.7 to 33.0)[a]	50.0 (49.1 to 50.5)[a]	**77.0** (74.8 to 78.6)	
Lot A				0/4
Lot B				0/4
Allegro				
Lot A vs mean SRMP	**32.4** (30.89 to 34.00)[a]	**50.0** (49.01 to 51.01)[a]	**76.4** (74.27 to 78.45)	
Lot B vs mean SRMP	**30.3** (28.48 to 32.03)	**48.4** (47.22 to 49.56)	**75.6** (74.06 to 77.12)	
Lot A vs Lot B	**28.0** (26.5 to 29.0)[a]	**46.0** (45.6 to 47.1)[a]	**74.0** (72.8 to 75.5)	

Lot A				0/4
Lot B				0/4

[a] Statistically significant differences were observed for bias when 30 and/or 48 and/or 75 mmol/mol are not within 95% confidence interval limits.

Table 3: Bias at different HbA_{1c} levels for each of the different POC analyzers using two different lot numbers, compared with the mean of the SRMPs. Bias between the two lot numbers. Results for NGSP show how many times the criteria were met out of a possible 4 comparisons.

	HbAS (frozen n=20, fresh n=16)	HbAC (frozen n=20, fresh n=7)	HbAD (frozen n=20, fresh n=5)	HbAE (frozen n=20, fresh n=9)	A2 (frozen n=15, fresh n=0)	HbF[a] (frozen n=15, fresh n=4)	Manufacturer claims for interference if present
A1Care							
Frozen Fresh	2.1 -0.4	1.8 -1.1	-0.7 -2.2	5.5 4.7	2.9	>9.3%	Hb F >10%
Cobas B101							
Frozen Fresh	-0.1 -0.1	0.2 1.8	**9.9** 6.2	**11.1** **10.9**	6.3	>9.5%	None
InnovaStar							
Frozen Fresh	1.3 -2.2	**9.5** **14.0**	3.3 4.5	1.8 7.1	2.8	>5.4%	AE and elevated F
The Lab 001							
Frozen Fresh	-6.6 **-10.1**	-6.1 -0.5	-0.3 -0.5	-1.1 -3.1	0	>42%	Hb F >30%
QuikReadgo							
Frozen Fresh	2.0 -1.7	**18.5** **22.1**	**-11.5** -6.1	-2.2 5.7	**-10.8**	>3.5%	AC and F >7%
Allegro							
Frozen Fresh	**-9.5** **-10.1**	7.0 8.5	**1.1** **9.2**	6.2 7.0	-1.7	>3.5%	None

[a]= % of HbF at which a significant negative bias results

Table 4 Mean relative difference (%) of the common Hb-variants compared to the assigned value after correction for bias in non-variant samples (number of samples for frozen and fresh samples). Red: equals at or near 10% difference. Green: this variant would not be seen to affect the value if only evaluated using frozen samples. The manufacturer's claims for are also listed.