Opinion Paper

Federica Braga* and Mauro Panteghini

Standardization and analytical goals for glycated hemoglobin measurement

Abstract: Glycated hemoglobin (HbA1c) plays a key role in diagnosing diabetes and monitoring the glycemic state. To guarantee the reliability of its measurement at the global level, the IFCC has defined a reference measurement system, based on the definition of the measurand as hemoglobin molecules having a special hexapeptide in common, which is the stable adduct of glucose to the N-terminal valine of the hemoglobin β-chain. In addition to the traceability of HbA1c results to the reference system, the establishment of analytical goals to make HbA1c measurements clinically reliable becomes crucial. However, allowable goals will depend on the assay specificity (i.e., selectivity) and, consequently, on units in which HbA1c results are expressed [mmol/mol for IFCC-aligned systems or % for National Glycohemoglobin Standardization Program (NGSP) converted numbers]. In this regard, analytical goals derived from biological variability studies in which the determination of HbA1c has been carried out by an assay providing the same selectivity for the measurand as defined by the IFCC are recommended. Only these targets should be used for evaluating the performance of commercial assays traceable to the IFCC system and of clinical laboratories using them. Analytical systems following different calibration hierarchies (e.g., the NGSP-aligned assays) will require different analytical goals, possibly derived from clinical outcome data.

Introduction

Glycated hemoglobin (HbA1c) is the product of non-enzymatic reaction of condensation between the aldei-dic group of glucose and the amino residues of some amino acids, mainly terminal valines, of the hemoglobin β-chains. The monitoring of its blood concentrations is central in the assessment of the degree of glyco-metabolic control in diabetic patients and in the risk prediction of vascular complications of these subjects [1]. In addition, HbA1c determination has recently been recommended for the diagnosis of diabetes mellitus [2].

As HbA1c test plays a crucial role in the monitoring and diagnosis of diabetes, it is essential that its clinical use be supported by standardized results, i.e., accurate and equivalent among different commercial methods and clinical laboratories using them. In general, this can be achieved by the implementation of a reference measurement system, based on the concept of metrological traceability, together with the achievement of a clinically acceptable level of measurement uncertainty [3]. The diagnostic manufacturers should implement analytical systems that produce results traceable to the higher-order references and able to fulfil the analytical goals of measurement uncertainty, established on the basis of clinical application of the test [4]. Finally, it is the responsibility of clinical laboratories to continuously monitor the performance of commercial methods in use, both by the implementation of a proper internal quality control (IQC), checking the daily alignment of the analytical system and evaluating the assay long-term imprecision, and the participation in appropriately organized external quality assessment schemes (EQAS) [4]. In this paper, we have analyzed the standardization process of HbA1c, with special attention to approaches for establishing the analytical goals of the measurement in light of the changes that have occurred with the implementation of IFCC standardization.
IFCC reference measurement system for HbA₁c and its implementation

To significantly reduce differences among results obtained by various commercial methods, in 1995 the IFCC established a working group with the aim of standardizing the measurement of HbA₁c at global level [5]. Starting from the measurand definition (i.e., hemoglobin molecules having a special hexapeptide in common, which is the stable adduct of glucose to the N-terminal valine of the hemoglobin β-chain), the IFCC working group has developed a complete reference measurement system for HbA₁c, through the implementation of two equivalent reference methods (HPLC/mass spectrometry- and HPLC/capillary electrophoresis-based) highly specific for the measurand as defined, the characterization of primary and secondary calibrators and the organization of an international network of laboratories performing one or both reference procedures [6]. Particularly, the IFCC HbA₁c network cooperates in assigning values to a panel of whole blood samples of proven commutability, which should be used by diagnostic manufacturers to calibrate their internal selected procedure and, through a protocol transferring the trueeness, to finally assign traceable values to product (i.e., commercial) calibrators used in the field procedures and patient samples. As a consequence, the results obtained on routine samples by clinical laboratories using the aligned analytical systems will be traceable to the reference measurement system, thus obtaining the standardization of HbA₁c measurement.

Before the establishment of IFCC reference measurement system for HbA₁c, national and regional harmonization programs of the measurement were already in place. The most popular has been the U.S. National Glycohemoglobin Standardization Program (NGSP), created to harmonize HbA₁c results through the implementation of assay traceability to the ion-exchange HPLC method, originally employed in the Diabetes Control and Complications Trial (DCCT) [7]. In comparison with the previous situation, the definition of the IFCC reference measurement system has brought about from the metrological point of view important and innovative improvements, summarized in Table 1. A first changing aspect has been the unequivocal definition of the ‘HbA₁c’ measurand. In the NGSP system, HbA₁c was roughly defined as the area under the curve of the corresponding chromatographic peak obtained with the before mentioned method, which had certainly not in itself the inherent characteristics of a reference method, measuring HbA₁c by a simple chromatographic separation based on the difference in isoelectric points between HbA₀ and all glycated hemoglobins. As mentioned, the reference methods recommended by IFCC are instead procedures highly specific for the measurand as it has been defined [8]. In order to distinguish the two metrological systems and to avoid confusion of values if expressed in similar units the measurement unit of the IFCC system was changed from percentage to mmol/mol [9].

In practice, the transition from one system to another has, however, been not easy, especially because all the main data supporting the clinical use of HbA₁c have been obtained using assays aligned to the NGSP system [4, 10]. It was thus sought to establish a kind of ‘clinical relationship’ between NGSP and IFCC systems in order to transfer decision limits and the clinical experience gained with the NGSP-aligned assays to the new IFCC system. From this need arose, by a series of experiments repeated over the years, the so-called IFCC-NGSP ‘master equation’ (ME) \[ \text{NGSP (\%)} = 0.09148 \times \text{IFCC (mmol/mol)} + 2.152 \], which expresses the correlation between the two systems and has allowed the conversion of analytical and clinical data from one system to another [10, 11].

Table 1 Main differences from the metrological point of view between the US National Glycohemoglobin Standardization Program (NGSP) and the IFCC reference measurement system for HbA₁c.

<table>
<thead>
<tr>
<th>Feature</th>
<th>NGSP system</th>
<th>IFCC system</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurand definition</td>
<td>Area under the curve of the chromatographic peak</td>
<td>Molecules of hemoglobin having a special hexapeptide in common, which is the stable adduct of glucose to the N-terminal valine of the hemoglobin β-chain (βN1-deoxyfructosyl-hemoglobin)</td>
</tr>
<tr>
<td>Reference method</td>
<td>Ion-exchange HPLC (Biorex 70 column)</td>
<td>Proteolytic digestion (by endoproteinase glu-C) of the N-terminal hemoglobin β-chain, HPLC separation of glycated and non-glycated N-terminal hexapeptides, and their quantification using mass spectrometry or capillary electrophoresis</td>
</tr>
<tr>
<td>Units for result</td>
<td>%</td>
<td>Millimoles of glycated hemoglobin β-chain per moles of hemoglobin β-chain (mmol/mol) [in agreement with the International System (SI) of measurement]</td>
</tr>
</tbody>
</table>
The use of the ME to convert individual results of HbA\textsubscript{1c} from mmol/mol to % (and vice versa) presents, however, some limitations that should be considered by users of commercial assays. It is known that moving along the metrological traceability chain, i.e., using methods that are less and less accurate, but of greater practicality, and calibrators having matrix similar to that of biological samples, a progressive increase in the combined standard uncertainty of the measurement is observed. The conversion of results from mmol/mol to % adds an additional (although small) source of uncertainty to the result when expressed as a percentage. The independent variable \(x\) of ME represents the HbA\textsubscript{1c} result obtained by the commercial method traceable to the IFCC reference system. This value already has its own uncertainty, which is mainly due to the sum of the uncertainty of the commercial method used to obtain it and uncertainties accumulated in the higher parts of the traceability chain. If this result is converted to NGSP units (%) by the use of ME, to the described combined uncertainty one has also to add that relating to slope and intercept estimates of the regression equation. Geistanger et al. \[12\] estimated the average increase of uncertainty that occurs with the use of such conversion, showing that the relative standard uncertainty added to HbA\textsubscript{1c} results with the ME transformation may vary from 0.05\% to 0.20\%, depending on the HbA\textsubscript{1c} concentration. To better understand this issue, Figure 1 shows in detail the traceability chain of the HbA\textsubscript{1c} measurement and the associated combined standard uncertainty. In the example, the values of uncertainties attributable to primary and secondary calibrators are those indicated in the original work of Jeppsson et al. \[8\] (0.63\% and 1.00\%, respectively) and the uncertainty attributable to the manufacturer’s product calibrator is that declared by Roche Diagnostics for the C.f.a.s. calibrator employed on the Cobas Integra system for the Tina-quant HbA\textsubscript{1c} gen. 2 immunoturbidimetric method, equal to 1.18\% \[13\]. Finally, the uncertainty estimate of the measurement obtained with the routine procedure derives from the yearly average imprecision (expressed as CV) of the above-mentioned method (2.0\%), used in the clinical laboratory of the ‘Luigi Sacco’ hospital in Milan (Italy), divided by 2.

The second limitation is inherent to the different analytical selectivity of the two measuring systems

---

**Figure 1** HbA\textsubscript{1c} reference measurement system and associated combined standard uncertainty.
(NGSP and IFCC), which is confirmed by the presence of a significant intercept in the ME (2.152). If the used methodology has different specificity for the measured analyte, one can expect that also the biological variability, a property closely associated with the characteristics of the analyte itself, significantly changes [14]. And, if the biological variability changes, the analytical goals derived from it may be different.

**Analytical goals for HbA\textsubscript{1c} measurement**

In addition to the definition of the reference measurement system, it is crucial to define the analytical goals for the HbA\textsubscript{1c} measurement to make its determination clinically usable and to ensure that the measurement error does not prevail. Otherwise this may obscure the clinical information supplied by the result [15, 16]. In the conference held in 1999 in Stockholm, under the auspices of IFCC and International Union of Pure and Applied Chemistry (IUPAC), a hierarchy of sources for deriving the analytical goals of a laboratory measurement and defining quality specifications in Laboratory Medicine was defined [17]. Using the established consensus, the best scientific approach of defining analytical performance goals should rely on (in a hierarchical order): 1) the evaluation of the effect of analytical performance on clinical outcome in the specific clinical setting (e.g., clinical misclassification); 2) data based on components of biological variation of the analyte; 3) data based on experts’ opinion and published recommendations; and, if the previous information is lacking, 4) goals set by regulatory bodies and EQAS organizers based on the current state of the art of the measurement.

In principle, analytical goals derived from experimental studies assessing the clinical impact of methodological performances are difficult to obtain in practice, with the result that there are few examples available in laboratory medicine [18]. For HbA\textsubscript{1c}, a simulation based on the DCCT results has been suggested [19]. As in this study patients in poor glycemic control had HbA\textsubscript{1c} concentrations >8.0%, while those in good glycemic control had values <7.0%, it has been estimated that to properly classify an individual with an HbA\textsubscript{1c} value of 7.5%, the measurement error should not exceed ±0.5% (as absolute value of HbA\textsubscript{1c}), amounting to a relative total error (TE) of ±6.7% (0.5%/7.5%) [20]. Indeed, if the measurement error is greater, a patient with an HbA\textsubscript{1c} of 7.5% would be indifferently classified in both glycemic control categories (good or poor) and this obviously would not be acceptable. This approach has been endorsed in Italy by the A\texttextsubscript{1c} Delegates Working Group (GLAD) in their document published in 2010 [16].

As direct experimental data on the impact of HbA\textsubscript{1c} measurement quality on clinical outcome are lacking, it is possible to derive analytical goals from biological variation components of the analyte [21]. It is now usual in our profession to derive this information from the database available on the website www.westgard.com/biodatabase1.htm, where there is a dedicated section showing, for each analyte, intra- (CV\textsubscript{I}) and inter-individual (CV\textsubscript{G}) biological variability components, derived from data available in the literature, and the corresponding desirable targets for imprecision, bias and TE. In order to check the reliability of this information, we systematically analyzed the results of all the experimental studies on HbA\textsubscript{1c} biological variation available in literature [14]. According to Fraser and Harris [22], our evaluation criteria were the type of enrolled subjects, frequency and number of blood sampling, way of sample storage until analysis (if any), analytical selectivity of employed methods and approaches used for statistical derivation of data. Quite unexpectedly, the analysis showed that none of available studies could be considered scientifically reliable to use their biological variability results for the derivation of analytical goals for the HbA\textsubscript{1c} measurement [14].

The third level in the IFCC/IUPAC hierarchy of strategies to derive analytical goals is occupied by recommendations given by international and national expert groups. This is evidently a more subjective approach if compared with the previous ones, usually based on consensus. For HbA\textsubscript{1c}, both an Australian working group and the Italian GLAD used this strategy to recommend a target for analytical imprecision (as CV) of <2.0% [16, 23].

The analytical goals established by EQAS organizers and those derived from state of the art of the measurement fill the lowest position in the hierarchy, being undoubtedly the source more tautological and less suitable for the improvement of measurement quality. Carobene et al. [24] have clearly shown that if analytical goals are exclusively defined by state of the art of performances obtained from EQAS, the criteria for acceptability of results adopted by different EQAS providers would be different. Furthermore, it is important to consider that control materials used in many EQAS are usually not commutable, i.e., they do not behave in the same way of biological samples [25]. In general, all those issues lead to a wide heterogeneity and a poor scientific background of the acceptability criteria employed in different EQAS [26]. This is certainly a problem, especially considering the importance that these programs should have in evaluating and monitoring the performance of clinical laboratories.
With regard to HbA1c measurement, the situation is further complicated by previously mentioned difference in analytical selectivity of IFCC and NGSP systems. Weykamp et al. [27] have elegantly demonstrated with a simple mathematical reasoning that analytical goals for HbA1c related to NGSP and IFCC systems are different. This consideration is applicable to all cases in which the results of a measuring system are related to those of another one by means of a conversion equation having an intercept significantly different from zero, which reflects a different analytical selectivity between the two systems. In their simulation, the authors converted by ME the percentage (NGSP) results of a study on HbA1c biological variation to mmol/mol (IFCC) unit and estimated, for both systems, the two biological variability components, i.e., CVI and CVG. The results showed that these components were higher using the IFCC system (2.9% and 7.3%, respectively) than for NGSP system (1.6% and 4.1%, respectively), giving rise, consequently, to different analytical goals.

In the derivation of data on the HbA1c biological variation it is therefore essential to focus on the selectivity of the employed analytical method. All studies evaluated in the above-mentioned review [14], except one, assayed a measurand different from that defined by IFCC, also including in the measurement hemoglobins glycated on sites different from N-terminal valine of β-chain. The only study using an immunoturbidimetric assay, selective for the measurand as defined by IFCC, unfortunately did not meet many of other established quality criteria [28]. Considering the evident lack of definitive data, we have, therefore, experimentally revaluated the biological variability of HbA1c, using an assay of which we had previously checked the perfect alignment to the IFCC reference system [29]. Table 2 shows the obtained results compared with those available at www.westgard.com/biodatabase1.htm (accessed November 2012), derived from more heterogeneous and less controlled sources. As previously noted [14], the data related to the HbA1c biological variability reported in this database are apparently obtained from the mean of results of some studies available in literature. Moreover, it is impossible to understand the criterion of the study evaluation and data selection, so that the reported data have been an object of perplexity [14, 30, 31], which, for HbA1c, we think have been, at least partly, the reason of a change in the associated CVI in this database, which from 3.4% in 2010 has been recently (2012) reduced to 1.9%, likely on the basis of the expressed criticisms.

The impact of the issue related to the analytical selectivity for the measurand can be highlighted also for analytical goals derived from clinical outcome, on the basis of the previously reported simulation [20]. Indeed, if NGSP decision levels in percentage (7.0% and 8.0%) are replaced with the corresponding values in mmol/mol, obtained using ME (53 and 64 mmol/mol), to avoid a misclassification the HbA1c of a subject with 58.5 mmol/mol should be measured with a maximum TE of ±9.4% (5.5/58.5), a goal larger than that calculated using data expressed in NGSP units (±6.7%).

### Performance evaluation of HbA1c methods and of laboratories using them

After defining the reference measurement system (i.e., the full traceability chain) and analytical goals, both essential to make the measurement clinically acceptable, and implementing by diagnostic manufacturers analytical systems (including platform, reagents, calibrators and control materials for checking the system alignment) that meet the established traceability and quality requirements, it is responsibility of end users (i.e., clinical laboratories) and, more in general, of our profession to verify and monitor the performance of commercially available systems and of laboratories that perform measurements in clinical setting [3, 26]. From this point of view, the availability of

<table>
<thead>
<tr>
<th>Reference</th>
<th>CVI</th>
<th>CVG</th>
<th>Imprecision (CVI)</th>
<th>Bias (%)</th>
<th>Total error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>O</td>
<td>D</td>
<td>M</td>
<td>O</td>
<td>D</td>
</tr>
<tr>
<td>Braga et al. [29]</td>
<td>2.5%</td>
<td>7.1%</td>
<td>0.63</td>
<td>1.25</td>
<td>1.88</td>
</tr>
<tr>
<td><a href="http://www.westgard.com">www.westgard.com</a></td>
<td>1.9%</td>
<td>5.7%</td>
<td>0.50</td>
<td>0.95</td>
<td>1.45</td>
</tr>
</tbody>
</table>

CVI, intra-individual biological variation; CVG, inter-individual biological variation; D, desirable quality level; M, minimum quality level; O, optimum quality level.
a IQC material allowing to check and confirm the system alignment in daily practice (CE-marked devices should produce results virtually unbiased compared to the established reference system) and of an appropriately structured EQAS (true value in commutable materials), able to provide objective information on the analytical quality of measurements performed by participating laboratories, are essential requirements [32]. However, as already mentioned, current EQAS programs are often inadequate for assessing the traceability of HbA_1c results obtained by clinical laboratories. Particularly, the commutability of employed control materials is not validated and only the mean (or other indicators of central tendency) of specific method- or commercial system group’s results is used for the performance evaluation of participating laboratories, without assigning values (and corresponding uncertainty) to EQAS materials by the reference measurement procedure. From this point of view, it should be noted that the accredited laboratories listed in the database of Joint Committee for Traceability in Laboratory Medicine (JCTLM), which are able to deliver a reference measurement service to interested customers according to the specified analytical requirements for IFCC reference procedures, may provide results within tolerable uncertainty [33]. The assignment of the true value to EQAS commutable materials, along with the application of the clinically acceptable TE limits, is the only approach that allows an objective evaluation of the analytical quality of HbA_1c measurements by a competence classification of clinical laboratories participating in the program based on the accuracy of their performance. An example of the effectiveness of this approach has been recently provided for the measurement of serum creatinine [34].

An outlook

In this paper we have tried to clarify the most important issues related to the HbA_1c measurement after the implementation of standardization to the IFCC reference measurement system. Overall, a greater attention to scientific aspects and to quality of the used approaches is undeniably evident, with a more correct application of the metrological approach (beginning with the ‘measurand’ definition), the definition of more evidence-based analytical goals and the more objective evaluation of laboratory performances. This does not mean that the optimal situation has been reached. For instance, by analyzing the combined standard uncertainty of the current traceability chain for HbA_1c, as depicted in Figure 1, it is clear that the relative combined standard uncertainty associated with the measurement of a biological sample (approx. 2.0%), which corresponds to an expanded uncertainty equal to approximately 4.0% [obtained by multiplying the combined standard uncertainty by a coverage factor of 2 (95% level of confidence)], is still $>2$ times the minimum acceptable target that, for unbiased results, would be $\leq 1.88\%$ (minimum quality level goal for imprecision in Table 2). As already highlighted for other important laboratory parameters requiring high analytical quality [35], further advances are probably needed, from one hand to reduce uncertainty associated with higher-order metrological references (reference materials and procedures) and on the other hand to increase the precision of commercial HbA_1c assays. One possibility would be to reduce uncertainties related to the certified purities/concentrations or mass fractions of HbA_1c and HbA preparations used as primary references (reference materials and procedures) and on the other hand to reduce the uncertainty associated with the metrological chain for HbA_1c, as depicted in Figure 1. As mentioned several times in this paper, the implementation of standardization to the IFCC reference measurement system should also be associated with the univocal definition of limits of...
clinical acceptability of the measurement error. In this regard, according to data of the study on biological variability previously mentioned [29], we would like to recommend (approximating) an analytical goal for imprecision (total CV) of <2.0% along with a maximal allowable bias of ±2.8%. These limits can reasonably represent a good compromise as minimum quality specifications for the HbA1c measurement. The first one can also represent the expanded uncertainty budget on which to base the performance specifications of reference procedures. By convention, these procedures should work with a measurement uncertainty at most equal to one third of the total budget [40] (so, in the case of HbA1c, around 0.7%). The maximal allowable bias represents the allowable systematic deviation limit of the calibration of commercial methods above which a realignment process to the higher-order metrological references of the calibrator’s value should be implemented by the manufacturer. An allowable maximum TE of ±6% to be applied in EQAS programs would be consistent with the above proposed limits for imprecision and bias. EQAS need a ‘Copernican’ change to contribute to ensure and possibly improve the reliability of HbA1c results. As previously mentioned, no EQAS programs currently available for HbA1c fulfil all the requirements necessary for an objective evaluation of the performance of individual participating laboratories (Table 3). The fact that at least some of such programs meet these criteria must become a priority.

Acknowledgments: The authors thank Prof. Andrea Mosca (University of Milan) for his contribution to the discussion of many aspects of this work.

Conflict of interest statement

Authors’ conflict of interest disclosure: The authors stated that there are no conflicts of interest regarding the publication of this article.

Research funding: None declared.

Employment or leadership: None declared.

Honorarium: None declared.

Received January 30, 2013; accepted April 18, 2013; previously published online May 13, 2013

Table 3 Requirements for the applicability of the External Quality Assessment Scheme (EQAS) results in the evaluation of the performance of participating laboratories in terms of standardization and traceability of the HbA1c measurement.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Aim</th>
</tr>
</thead>
<tbody>
<tr>
<td>EQAS materials value-assigned with IFCC reference procedure by an accredited reference laboratory</td>
<td>To check the alignment of commercial HbA1c methods used by participating laboratories to the IFCC reference measurement system</td>
</tr>
<tr>
<td>Proved commutability of EQAS materials</td>
<td>To allow transferability of participating laboratory performance to the measurement of patient samples</td>
</tr>
<tr>
<td>Use of an analytical goal for allowable total error equal to ±6%</td>
<td>To verify the suitability of laboratory HbA1c measurements in clinical setting</td>
</tr>
</tbody>
</table>

References


26. Mosca A. Some practical advices on how to implement the international standardization of glycated hemoglobin measurement in Italy. Biochim Clin 2011;35:36–41.


