EXTERNAL QUALITY ASSESSMENT OF GLYCATED HEMOGLOBIN IN IRAN:
COMPARISON OF FIVE DIFFERENT COMMERCIAL METHODS WITH
TWO DIFFERENT TOTAL ALLOWABLE ERRORS

EKSTERNÁ PROCENA KVALITETA GLIKOLIZIRANOG HEMOGLOBINA U IRANU:
POREĐENJE PET RAZLIČITIH KOMERCIJALNIH METODA SA
DVE RAZLIČITE UKUPNE DOZVOLJENE GREŠKE

Fereshteh Atabi1, Reza Mohammadi1,2

1Department of Biochemistry and Pharmacology, Islamic Azad University,
Tehran Medical Sciences Branch, Tehran, Iran
2Department of Biochemistry of External Quality Assessment Program,
Iranian Association of Clinical Laboratory Doctors, Tehran, Iran

Summary
Background: Glycated hemoglobin (HbA1c) measuring has a critical role in the monitoring and diagnosis of diabetes. So, the analytical performance of its measuring method must be acceptable. Clinical laboratories should continuously monitor the performance of their commercial methods, both by using proper internal quality control (IQC) and by participating in external quality assessment schemes (EQAS).

Methods: In January and August 2016, two different freshly prepared commutable patient QC samples were sent to over 1000 laboratories, but 682 and 925 different laboratories which were used five common commercial methods for measuring HbA1c, included in this study during 23th and 24th runs of the external quality assessment program (EQAP), respectively. Target values for total group and also for peer groups were calculated. The performance of each method and laboratory were determined according to two different allowable total errors (TEa), including ±6% and ±20%, which are suggested by the National Glycohemoglobin Standardization Program (NGSP) and Reference Health Laboratory of Iran, respectively.

Results: Considering TEa of ±20% in evaluating HbA1c commercial methods and laboratory performances, pass rates were from 97% to 98% during EQAP-23 and EQAP-24, respectively. However, when using TEa of ±6%, the pass rates were significantly lower.

Kratak sadržaj
Uvod: Merenje glikoliziranog hemoglobina (HbA1c) ima presudnu ulogu u praćenju i dijagnostikovanju dijabetesa. Zbog toga, metode za njegovo merenje moraju imati prihvatljive analitičke performanse. Kliničke laboratorije moraju stalno da prate performanse svojih komercijalnih metoda, kako korisčenjem odgovarajuće interne kontrole kvaliteta (IQC) tako i kroz učešće u šemama spoljne procene kvaliteta (EQAS).

Metode: Tokom 23. i 24. kruga programa spoljne procene kvaliteta (EQAP), u januaru i avgustu 2016, dva različita sveže pripremljena komutabilna QC uzorka pacijenata poslata su u 682 i 925 različitih laboratorija koje koriste pet uobičajenih komercijalnih metoda za merenje HbA1c. Izračunate su ciljne vrednosti za ukupnu grupu i srodne grupe (peer groups).

Rezultati: Performanse svake metode i laboratorije određene su u odnosu na dve različite dozvoljene ukupne greške (TEa), naime ±6% i ±20%, predložene od strane Nacionalnog programa za standardizaciju glikohemoglobina (NGSP) i Referentne zdravstvene laboratorije Irana. Kad se radi od TEa od ±20% u evaluaciji komercijalnih metoda za HbA1c i laboratorijskih performansi, stope prolaznosti (pass rates) bile su od 97% do 98% tokom EQAP-23 i EQAP-24. Međutim, kada...
Introduction

According to the critical role of the level of glycated hemoglobin (HbA1c) in the monitoring and diagnosis of diabetes (1, 2) the acceptable precision and accuracy of different commercial methods in clinical laboratories for measuring HbA1c is essential (3). There are many different methods such as ion-exchange chromatography, capillary electrophoresis, boronate affinity chromatography, immunoassay and enzymatic methods, which separate glycated hemoglobins from nonglycated ones, based on existing differences in structure, charge, affinity or immunity (2).

The manufacturers should produce commercial methods that their results could be traceable to the higher-order references and be able to achieve acceptable analytical goals of measurement (3). Analytical goals must be defined in such a way that the test could save its clinical usefulness. On the other hand, it is the responsibility of clinical laboratories to monitor the performance of their methods, both by using a proper internal quality control (IQC) and participating in external quality assessment schemes (EQAS) (3). In addition, before routinely using any commercial methods for analyzing patient specimens, each laboratory should verify method producer claims or validate method performance according to the laboratory analytical goals (4). In each above experiment, including IQC, EQAS, and method validation, judgement of method performance acceptability depends on the established allowable total error (TEa). TEa is a maximum tolerable error that, if it occurs, will not impair the result’s clinical usefulness (5–7).

In order to significantly reduce differences between results obtained by various commercial methods, standardization of the methods of HbA1c measurement has been necessary. This can be achieved by the International Federation of Clinical Chemistry (IFCC) and National Glycohemoglobin Standardization Program (NGSP) (5). The American Diabetes Association (ADA) recommends that laboratories use only NGSP-certified HbA1c methods and participate in the College of American Pathologists (CAP) fresh sample proficiency testing survey (6, 7). Using NGSP certified methods is not the only step in achieving a precise and accurate analytical HbA1c measuring method. Continuous monitoring of the performance of the method is needed. In this regard, defining analytical goals with which method performance must be evaluated, has profound effects on detection of analytical errors.

The exact number of errors made can be quantified by employing sigma metrics in the laboratory. According to selected TEa along with method bias and imprecision, we can calculate sigma of the method which is a valuable criterion for assessing the quality of the analytical phase. A high sigma value indicates good performance. A sigma value of 2 is defined as the minimum allowable value for routine performance (8).

Unfortunately, there are few reports about the performance of clinical laboratories measuring HbA1c in EQAS and proficiency testing (PT) of different countries. In the United States, CAP publishes these data regularly and some reports can be found from some other countries, including Italy and China. In this study, for the first time in Iran, we evaluated the effects of two different analytical goals, defined as TEa, on interpretation of EQAS HbA1c results. TEa encompasses the imprecision and bias of a single test measurement and is used to evaluate laboratory tests performance in EQAS (9). We also evaluated the analytical performance of the methods by calculating the sigma value of each method according to different TEa.

Materials and Methods

Sample Preparation

Pooled whole bloods were collected from diabetic patients who suffered from polycythemia and should have regular phlebotomy to reduce their red blood cell count. The bloods were collected in EDTA containing bags and then four or five isogroup bloods were mixed in order to have enough blood (about 1500 mL) for filling each HbA1c vial with 1.0 mL pooled whole blood. In this preparation, only EDTA, with no preservatives was added as a coagulant which is recommended by method producer. So in comparison to routine blood collection for HbA1c measuring,
the matrix of the blood was not changed and we can consider them commutable.

Explanation of the Procedure

During twenty third and twenty fourth runs of the external quality assessment program (EQAP-23 and EQAP-24), in January 2016 and August 2016, two different freshly prepared commutable patient whole blood samples in EDTA-containing vials were sent to 966 and 1393 participant laboratories, respectively, at 4 °C to 8 °C. In order to protect the blood samples against possible degradation and leakage during mailing, the plastic tubes were filled with the whole blood and sealed with stoppers and inserted into a triple container. The first container is a cold box with adequate ice bag for cooling the filled tubes. The second one is the Styrofoam container which acts as a thermal insulation and a water barrier and the last one is a carton on which the names and the addresses of the providers and participants together with the information about the mailing cautions have been recorded.

Using commutable samples, the results of different commercial methods can be compared (10). Before sending to participating laboratories, homogeneity of whole blood vials and after sending, the stability of these vials was assessed and confirmed. These assessments and confirmation were done according to WHO requirements (11). For homogeneity, about one percent of total vials, including ten and fourteen vials were randomly selected during EQAP-23 and EQAP-24, respectively, and their HbA1c determined by Tosoh G8 Auto HPLC in duplicate. Then the coefficient of variation (CV) of the results was calculated and compared with one fourth of an allowable total error of 6%, or 1.5%. For stability, one whole blood vial in each run was refrigerated by the end of the program and then the HbA1c was measured in duplicate and the average of the results compared with the homogeneity mean value ± modified allowable total error (mTEa%).

Although there are more than ten commercial methods for monitoring of HbA1c in Iran. This study was focused on common methods for which the number of using laboratories was at least twenty. So their statistical analysis could be valid, and their performance had also been evaluated (12). These included Biosystem, NycoCard, Pars Azmon, Pishtaz Teb, and Roche kits which their assay principles were cation-exchange chromatography, boronate-affinity chromatography, immunoturbidimetry, enzymatic, and immunoturbidimetry, respectively.

In EQAP-23, 682 participating laboratories used the desired commercial methods. These laboratories were grouped in five peer groups, including Biosystem, NycoCard, Pars Azmon, Pishtaz Teb, and Roche with 172, 233, 120, 137, and 20 participating laboratories, respectively. In EQAP-24, 925 participating laboratories used the desired commercial methods. These laboratories were grouped in five peer groups, including Biosystem, NycoCard, Pars Azmon, Pishtaz Teb, and Roche with 209, 337, 149, 205, and 25 participating laboratories, respectively. Each participant laboratory had to examine the sent control material as a routine patient sample according to the instructions of commercial method manufacturers and had to calibrate and control its measuring method by the calibrators and control materials, as an internal quality control, provided by method manufacturers. After measuring HbA1c, the results were sent to EQAP and statistical analyses were done. According to the used commercial method, the results were grouped into five peer groups. Then, mean, SD, and CV of each peer group and also total results were calculated. In EQA, the mean of each peer group is used as target value to evaluate each laboratory performance. In this regard, it is necessary to delete outliers which are out of Mean ± 2SD or 3SD (13). In EQAP, Mean ± 2.5SD is used. After deleting outliers, calculation of mean and SD was repeated until there were no outliers. The last calculated mean, termed as weighted mean, was used as the target value. Statistical analysis was done by SPSS 20 and Med Calc 13.3.1 software.

Target values for total group and also for peer groups were calculated. In Iran, laboratory performance is evaluated according to standard deviation interval (SDI). SDI is commonly calculated by the following formula (6–14):

\[
SDI = \frac{\text{Laboratory result} - \text{Peer group target value}}{\text{Peer group SD}}
\]

but adjusted SD instead of peer group SD is used. Adjusted SD is calculated by the following formula:

\[
\text{Adjusted SD} = \frac{\text{CCV} \times \text{Peer group target value}}{100}
\]

in which, CCV (Chosen Coefficient of Variation) is defined by the Reference Health Laboratory of Iran and equals 10% for HbA1c methods. SDI ≤ 2 is considered as acceptable result (12). This CCV and acceptability criteria represent TEa of 20%; i.e. SDI = 2 shows that the result is 2 SD far from mean target value and SD equals 10% of mean target value, so TEa = 2 SD = 2 × 10% = 20%.

For calculating the relative bias of the mean value of each commercial method, in comparison to the mean value of all groups, we subtracted peer group mean value from the all groups mean value.

\[
\text{Relative bias} = \frac{\text{Mean of commercial method} - \text{Mean of all groups}}{\text{Mean of all groups}}
\]

where SDI = 2 and peer group SD is used. The relative bias of the mean value of each commercial method in comparison to the mean value of all groups was calculated. Hence, the bias of the mean value of each commercial method in comparison to the mean value of all groups was calculated.
Total analytical error (TAE) has two parts, including bias and imprecision which are representative of systematic and random errors, respectively. For calculating bias% of each commercial method, first the TAE% of each method is calculated by the following formula:

\[
\text{TAE}\% = \frac{(\text{peer group mean value}) - (\text{all groups mean value})}{\text{all groups mean value}} \times 100.
\]

TAE% is commonly calculated as bias % + 2 CV% (15, 16). The components of analytical errors of a single measurement result include random and systematic errors. In comparison with a single measurement, when two or more replicate measurements are made, the mean of the obtained values is more likely to be closer to the true or target value, since repeated measurements decrease the effect of the random error of the mean. Ultimately, when an infinite number of measurements are made, the effect of random error on the mean will be eliminated (16). The effect of a random error component of TAE is inversely related to the square root of the number of measurements (16, 17). If a test is repeated, its random error will decrease by \(\frac{1}{\sqrt{n}}\), in which \(n\) is the number of repeats (17). Thus, according to the number of measurement repeats, the effects of random error (imprecision) and systematic error (bias) on TAE can be calculated. In calculating bias%, we used the following formula to eliminate the effect of random error on TAE according to the number \((n)\) of participant laboratories in each peer group:

\[
\text{Bias}\% \text{TAE}\% = \frac{\sqrt{n}}{2 + \sqrt{n}} \times \text{TAE}\%
\]

We can define modified total analytical error \(\text{mTEa}\) for when the test is repeated \(n\) times. In this way, we can determine TEa for situations when the test is repeated, for example 2, 4, 10 or 20 times.

\[
\text{mTEa}\% = \frac{2 + \sqrt{n}}{3 \times \sqrt{n}} \times \text{TEa}\%
\]

this formula was derived from \(\text{mTEa}\% = \text{Bias}\% + 2 \text{CV}\% / \sqrt{n}\) when test is repeated \(n\) times.

Also, we calculated pass rates of laboratories having acceptable results during EQAP-23 and EQAP-24 with TEa of ±20% suggested by the Reference Health Laboratory of Iran and TEa of ±6% suggested by NGSP and CAP (5, 18). In this comparison, we used total mean as target value.

Commercial methods quality, as the sigma value of each method performance, was calculated according to the above two different TEa along with Bias and CV of each method (9):

\[
\text{Method quality (sigma)} = \frac{\text{TEa}\% + \text{Bias}\%}{\text{CV}\%}
\]

Results

In EQAP-23, and EQAP-24, the CVs of homogeneity studies were 0.8% and 0.9% respectively.

### Table I

Weighted mean (target value), SD, CV, TAE, Bias%, and method quality according to two different TEa in EQAP-23 and EQAP-24.

<table>
<thead>
<tr>
<th>Commercial Methods</th>
<th>No. labs</th>
<th>Mean (%</th>
<th>SD</th>
<th>CV (%)</th>
<th>TAE (%)</th>
<th>Bias (%)</th>
<th>Method quality (sigma)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EQAP-23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biosystem</td>
<td>172</td>
<td>7.93</td>
<td>0.83</td>
<td>10.5</td>
<td>- 0.50</td>
<td>- 0.44</td>
<td>0.61</td>
</tr>
<tr>
<td>NycoCard</td>
<td>233</td>
<td>8.05</td>
<td>0.52</td>
<td>6.5</td>
<td>+ 1.00</td>
<td>+ 0.89</td>
<td>0.79</td>
</tr>
<tr>
<td>Pasr Azmon</td>
<td>120</td>
<td>7.99</td>
<td>0.73</td>
<td>9.1</td>
<td>+ 0.25</td>
<td>+ 0.21</td>
<td>0.63</td>
</tr>
<tr>
<td>Pishtaz Teb</td>
<td>137</td>
<td>7.78</td>
<td>0.59</td>
<td>7.6</td>
<td>- 2.38</td>
<td>- 2.04</td>
<td>1.06</td>
</tr>
<tr>
<td>Roche</td>
<td>20</td>
<td>8.12</td>
<td>0.32</td>
<td>3.9</td>
<td>+ 1.88</td>
<td>+ 1.30</td>
<td>1.19</td>
</tr>
<tr>
<td>Total</td>
<td>682</td>
<td>7.97</td>
<td>0.62</td>
<td>7.8</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>EQAP-24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biosystem</td>
<td>209</td>
<td>7.66</td>
<td>0.90</td>
<td>11.8</td>
<td>+ 0.13</td>
<td>+ 0.11</td>
<td>0.50</td>
</tr>
<tr>
<td>NycoCard</td>
<td>337</td>
<td>7.63</td>
<td>0.53</td>
<td>7.0</td>
<td>- 0.26</td>
<td>- 0.24</td>
<td>0.90</td>
</tr>
<tr>
<td>Pasr Azmon</td>
<td>149</td>
<td>7.70</td>
<td>0.61</td>
<td>7.9</td>
<td>+ 0.65</td>
<td>+ 0.56</td>
<td>0.69</td>
</tr>
<tr>
<td>Pishtaz Teb</td>
<td>205</td>
<td>7.60</td>
<td>0.57</td>
<td>7.5</td>
<td>- 0.65</td>
<td>- 0.57</td>
<td>0.88</td>
</tr>
<tr>
<td>Roche</td>
<td>25</td>
<td>7.88</td>
<td>0.39</td>
<td>5.0</td>
<td>+ 3.01</td>
<td>+ 2.15</td>
<td>0.78</td>
</tr>
<tr>
<td>Total</td>
<td>925</td>
<td>7.65</td>
<td>0.62</td>
<td>8.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Abbreviations: TAE, total analytical error; TEa, allowable total error; HbA1c, glycated hemoglobin; EQAP-23, twenty-third run of external quality assessment program; EQAP-24, twenty-fourth run of external quality assessment program.
### Table II Acceptable range and pass rates of HbA\textsubscript{1c} in EQAP-23 and EQAP-24 with different allowable total error (TEa).

<table>
<thead>
<tr>
<th>Commercial Methods</th>
<th>No. Labs</th>
<th>Reported Range (%)</th>
<th>TEa = 6%</th>
<th>TEa = 20%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Acceptable range (%)</td>
<td>Pass Rates</td>
</tr>
<tr>
<td>EQAP-23</td>
<td></td>
<td></td>
<td></td>
<td>N</td>
</tr>
<tr>
<td>Biosystem</td>
<td>172</td>
<td>5.90–10.00</td>
<td>7.49–8.45</td>
<td>86</td>
</tr>
<tr>
<td>NycoCard</td>
<td>233</td>
<td>6.80–9.20</td>
<td></td>
<td>163</td>
</tr>
<tr>
<td>Pars Azmon</td>
<td>120</td>
<td>6.20–9.80</td>
<td>62</td>
<td>52</td>
</tr>
<tr>
<td>Pishtaz Teb</td>
<td>137</td>
<td>6.40–9.10</td>
<td>92</td>
<td>67</td>
</tr>
<tr>
<td>Roche</td>
<td>20</td>
<td>7.69–8.69</td>
<td>17</td>
<td>85</td>
</tr>
<tr>
<td>Total</td>
<td>682</td>
<td>5.90–10.00</td>
<td>420</td>
<td>62</td>
</tr>
<tr>
<td>EQAP-24</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NycoCard</td>
<td>337</td>
<td>6.40–8.90</td>
<td></td>
<td>224</td>
</tr>
<tr>
<td>Pars Azmon</td>
<td>149</td>
<td>6.20–9.10</td>
<td>86</td>
<td>58</td>
</tr>
<tr>
<td>Pishtaz Teb</td>
<td>205</td>
<td>6.20–9.00</td>
<td>132</td>
<td>64</td>
</tr>
<tr>
<td>Roche</td>
<td>25</td>
<td>7.01–8.45</td>
<td>17</td>
<td>68</td>
</tr>
<tr>
<td>Total</td>
<td>925</td>
<td>5.56–9.80</td>
<td>555</td>
<td>60</td>
</tr>
</tbody>
</table>

Abbreviations: HbA\textsubscript{1c}, glycated hemoglobin; EQAP-23, twenty-third run of external quality assessment program; EQAP-24, twenty-fourth run of external quality assessment program.

**Figure 1** Performance during EQAP-23. Error bars represent ±2SD interval from the means (filled squares). The continuous horizontal line represents the total mean. Dotted and dashed horizontal lines show the allowable interval for total allowable errors of 6% and 20, respectively.
which were less than the allowable CV% of 1.5%. So, the homogeneity of the vials containing whole blood samples was confirmed. In the case of stability, the results of stored vials at 4 °C by the end of the program were 8.20% in EQAP-23 and 7.95 in EQAP-24, which were in the 6% acceptable limit of related homogeneity target values, 7.88 in EQAP-23 and 7.70 in EQAP-24, respectively. The results of Mean (target value), SD, CV, TAE, bias, method quality (sigma level with TEa of 6% and 20%) of each peer group and also target value, SD, and CV of total group for participating laboratories using the desired commercial methods in EQAP-23 and EQAP-24, have been shown in Table I. According to CAP allowable bias% of absolute 0.3%, all mean methods results were within the acceptable range in both EQAP-23 and EQAP-24.

As shown in Table II, when TEa of ±20% was used to evaluate HbA1c method performance, about 97% and 98% of participant laboratories had acceptable performance during EQAP-23 and EQAP-24, respectively. But when this evaluation was performed according to TEa of ±6% pass rates decreased significantly to 62% and 60%, respectively.

Data related to HbA1c methods and laboratory performance from EQAP-23 and EQAP-24 are demonstrated as multiple variable graphs in Figures 1 and 2, respectively. In these graphs, positions of different method means in respect to total mean, distribution of data around related means, error bars representing ±2SD interval, and allowable intervals for TEa of 6% and 20% are shown.

### Discussion

Reliability of methods for measurements of blood HbA1c has been studied extensively throughout the world (19–24). This reliability for methods which are commonly used in Iran has also been studied previously (12). This study and the present study have three important differences in relation to the other studies. First, in the other studies one method, usually an HPLC method, was chosen as a reference or comparative method and other methods, as test methods, were compared with that one. In this study, we had no access to a reference or comparative method which has no analytical error and whose results could be reliable for comparing the results of other methods. To resolve this problem, commutable fresh patient blood specimens are used instead of noncommutable commercial quality materials and means of all groups are used as target values. It seems this approach is a better approach to obtain results with the least error. Second, in the most of other studies, methods were grouped according to principles of HbA1c separation, including ion-exchange chromatography, affinity chromatography, immunoturbidimetry, enzymatic, etc. But, in this
study, methods were grouped according to commercial methods. In spite of using similar principles, products of different manufacturers may have different performance. For example, in this study, both Roche and Pars Azmon kits used immunoturbidimetric methods, but their performance (bias%, CV%, and sigma level) was completely different. Third, in this study, we used a formula to account the effect of the number of repeating measurements in reducing random error and calculating analytical bias properly. This may be the first time that such a formula has been used in the calculation of bias, TAE, and mTEa.

Allowable total error (TEa) is a parameter used to define acceptable analytical performance which defines the amount of error that can be tolerated without invalidating the medical usefulness of the analytical result. If TAE is less than TEa, then the performance of the test is considered acceptable. If this error is larger than the TEa, corrections should be made to reduce the analytical error or the method should be replaced. This process ensures that laboratory test results give accurate information to physicians to manage their patients effectively (25).

According to the Diabetes Control and Complications Trial (DCCT), HbA1c results <7.0% show good glycemic control and results >8.0% show poor glycemic control. In order to properly classify a patient with an HbA1c value of 7.5% will be incorrectly classified an HbA1c value of 5.6% (sample 1) and 7.8% (sample 2), respectively. With a TEa of 6.0%, pass rates were 70% and 77% of samples 1 and 2, respectively. With a TEa of 6.0%, pass rates were 70% and 77% of samples 1 and 2, respectively. All methods had a mean bias of ≤2.8% with respect to the target values, with the exception of Tosoh (bias of +6.1 and +5.8%, for samples 1 and 2, respectively) (27). Our results are comparable to theirs. Our bias is somewhat better and our inter-laboratory CVs% and pass rates are worse than theirs.

Conclusion

According to our study and comparison with the results of other related studies in different countries, we can conclude that (1) the calculated sigma and
HbA1c performance methods in Iran are comparable to the other countries and still there is a need to improve these method performances all over the world (2). Our laboratory pass rate is low and inter-laboratory CV% is high and these must be improved by using closed and NGSP-certified methods. And we need to gradually tighten the TEa from the improper ±20% to proper ±6% and the Iranian clinical laboratories that have not achieved acceptable results must improve their performance methods gradually, or replace their methods with an acceptable HbA1c measuring method (3).

Acknowledgments
This work was supported by Iran External Quality Assessment Program (EQAP).

Conflict of interest statement
The authors stated that they have no conflicts of interest regarding the publication of this article.

References


Received: September 24, 2017
Accepted: December 25, 2017