Introduction

The quantitation of blood hemoglobin A1c (HbA1c) provides a measure of blood glucose averaged over the prior 3 to 4 months. As HbA1c is inexpensive and provides a fairly accurate measure of long term (LT) glucose control, it is widely used for diabetes screening, diagnosis, treatment and research. With at least 23 million Americans with diagnosed diabetes (1), the total annual charges for HbA1c testing assuming Medicare reimbursement rates approaches 900 million US dollars.

Derivation of real metrics of long term patient and analytical variation of three hemoglobin A1c assays demonstrates both borderline and highly acceptable analytical performance

George Cembrowski1,2, Junyi Mei3, Renée Guérin4, Mark A. Cervinski5, Christopher McCudden6

1Laboratory Medicine and Pathology, University of Alberta, Edmonton, AB, Canada; 2Cembrowski & Cembrowski Quality Control Consulting, Edmonton, AB, Canada; 3Faculty of Medicine, University of Manitoba, Winnipeg, MB, Canada; 4CSSS Chicoutimi, Chicoutimi, QC, Canada; 5Giesel School of Medicine, Hanover, NH, Canada; 6Department of Pathology and Laboratory Medicine, University of Ottawa, Ottawa, ON, Canada

Contributions: (I) Conception and design: G Cembrowski, J Mei; (II) Administrative support: G Cembrowski; (III) Provision of study materials or patients: R Guerin, M Cervinski, C McCudden; (IV) Collection and assembly of data: J Mei, Cembrowski; (V) Data analysis and interpretation: J Mei, Cembrowski; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Background: Previously, we demonstrated that sequential intra-patient data can yield realistic measures of short term patient biologic and analytic variation. We now determine the long term (LT) combination of biologic and analytic variation associated with three hemoglobin A1c (HbA1c) assays.

Methods: Three sets of patient HbA1c data were analyzed: 2 years of Sebia Capillarys 2 Flex Piercing® HbA1c from a Quebec hospital laboratory, 3 years of Roche Tinia Quant Gen II HbA1c assay measured with Cobas 8000, c502 and Cobas 6000, c501 analyzers by a New Hampshire laboratory and 3.5 years of Siemens Vista HbA1c from an Ontario hospital. We generated graphs of the LT combination of biologic and analytic variation for the three methods and four patient subpopulations: (I) virtually the entire population, (II) high HbA1c, (III) low to normal HbA1c, and (IV) inner 50 percentile (P) HbA1c.

Results: The Sebia and the Roche combinations of LT biologic and analytic variations were consistently lower than those of the Vista. The Vista yielded the highest y intercept and the highest LT biologic/analytic combination variations at 26 weeks. The elevated Vista variations were most obvious for the low-normal and midlevel HbA1c groups.

Conclusions: The Vista assay demonstrated excess analytic variation. HbA1c assays with LT imprecisions exceeding 3% will result in artefactually elevated glycemic variability. The Roche and Sebia assays are fit for purpose. We recommend that post-market evaluations of LT intra-patient HbA1c variation be conducted as part of external quality assessment (EQA) schemes in order to provide clinically relevant perspectives of assay performance.

Keywords: Hemoglobin A1c (HbA1c); imprecision; glycemic variation; Sebia; Roche; Vista

Received: 07 February 2020; Accepted: 12 August 2020.
doi: 10.21037/jlpm-2019-qc-02
View this article at: http://dx.doi.org/10.21037/jlpm-2019-qc-02
Through the workings of the National Glycohemoglobin Standardization Program (NGSP), multiple professional and government organizations provide analytical guidance and assessment to manufacturers and users of HbA1c analytical systems (2). The NGSP’s ultimate goal is to assure that patient HbA1c results compare well to the reference HbA1c method used in the Diabetes Control and Complications Trial (DCCT). The NGSP network is composed of a Central Primary Reference Laboratory that monitors Secondary Reference Laboratories, which in turn work directly with assay manufacturers to harmonize methods and provide comparison data for method certification (3). The NGSP network is monitored semiannually against the International Federation of Clinical Chemistry Laboratory Network (4) via sample comparisons. The effectiveness of the NGSP program in harmonizing HbA1c results is assessed via the College of American Pathologists (CAP) whole blood proficiency testing (PT) survey performed three times a year. As instrument manufacturers have responded to improve bias and precision at the behest of the NGSP, the performance requirements (maximum bias and imprecision) have become stricter.

We have developed a calculus that transforms sequential intra-patient test results into a measure of PreAnalytic variation including biologic variation and ANalytic variation (5). [We call this measure PAAN\textsuperscript{TM} (6)]. Our first, rudimentary application of PAAN demonstrated significant analytic variation in an immunochemical HbA1c assay compared to a high-performance liquid chromatographic assay (7). We and others have used this methodology to study the imprecision of single and multiple analytical systems reporting HbA1c (8,9), blood gases, metabolites such as glucose (6) and electrolyte tests (5,10). Our method does not require any additional laboratory testing. Rather, it involves procuring large series of patient data available in clinical information systems and grouping sequential intra-patient result pairs into period bins that reflect the intervals of time between the sequential tests. For each time interval, the standard deviation of duplicates (SDDs) is calculated for all of the sequential intra-patient test pairs, \((x_i, x_j)\)… \((x_{2i-1}, x_{2i})\)… \((x_{2n-1}, x_{2n})\) within specific testing intervals:

\[
SDD = \sqrt{\frac{\sum (x_{2i} - x_{2i-1})^2}{2n}}
\]  

The initial portion of the SDD vs. time line is linear and represents a Taylor’s series expansion of an exponential equation fitting the data. If the SDD is regressed against the midpoints of the time intervals, the y intercept \((y_0)\) represents the sum of the preanalytic variance \((s_{pa})\), the intra-patient biologic variance \((s_b)\) and the analytic variance \((s_a)\):

\[
y_0 = s_{pa} + s_b + s_a
\]  

An example of preanalytical error in HbA1c is labile HbA1c, characterized by the reversible binding of glucose to hemoglobin as a Schiff base (11). Usually, the contributions of such preanalytic factors are small and are assumed negligible when compared to the biologic and analytic variation. For many analytes, especially those whose concentrations are closely controlled by the body’s homeostatic mechanisms, \(s_b\) is relatively constant. If the contribution of preanalytical error is disregarded, Eq. [2] can be rearranged to yield the biologic variation:

\[
s_b = (y_0^2 - s_a) / 2
\]  

Using this methodology, we have accurately determined short term biologic variations of many constituents of the complete blood count (12) and critical care blood gas and electrolyte panels (5). The biologic variation is important in setting analytic goals. For example, with this approach we have recommended tighter analytical variation goals based on the SDD analysis of serial patient whole blood lactate values (13).

While the short-term biologic variation of HbA1c has been quantitated by many groups, there are few longitudinal studies of HbA1c levels. Recently, HbA1c variation was studied in adolescent and young adults with type 1 diabetes (14). Between ages 10 and 16, HbA1c tends to increase, then plateaus for about 2 years and finally decreases around age 18. There are many HbA1c influences in these young adults including race/ethnicity, income, health insurance and insulin pump usage. High alcohol intake (>140 g/week) has been found to reduce glycemic variability in a general adult population (15). HbA1c increases during winter months and its increase is more obvious at greater distances from the equator (16). We have demonstrated a lack of such seasonal variation in HbA1c measured in equatorially situated Singapore (17). Assessments of the biologic variation of HbA1c in individuals with diabetes depend on the individuals’ glycemic control, the length of the assessment and the study’s season. Both LT HbA1c variation and variation over winters are higher than most published variations. One summertime determination of HbA1c variation cites a biologic variation of 1.7% for patients with type 1 diabetes.
compared to 1.2% in a healthy control population (18). In 1994, towards the end of the DCCT, we studied the HbA1c variation of 29 highly motivated DCCT patients enrolled in the intensive diabetes management arm. Over 3-month periods, the biologic variation of HbA1c was 2.4% (19); for 1 year, the biologic variation was 4.1%.

Our current work focuses on the LT variation of patient HbA1c. Until recently, we determined the SDD from consecutive, intra-patient paired results. As we incorporated only successive test pairs, the SDD was determined from paired results acquired over relatively short time intervals, up to 14 hours for electrolytes, 84 hours for hematology and about 30 days for a very “slow” analyte like HbA1c. However, even with 30 days of HbA1c data we obtained insufficient data pairs to accurately determine longer period SDD.

We now determine the LT SDD from all the sequences of paired HbA1c for every possible time interval. For example, if a patient had HbA1c measured every 18 weeks for 108 weeks (a total of 7 HbA1c’s), the first HbA1c is paired with the next six results: the 2rd (18 weeks interval), the 3rd (36 weeks interval), the 4th (54 weeks interval)… up to the 7th (108 weeks interval) resulting in six test pairs separated from 18 to 108 weeks. The 2nd HbA1c is paired with the five subsequent HbA1c and so on. Over 2 (or 3 years) and with thousands of patient observations, there will be adequate test pairs separated by periods from 1 to 52 weeks, respectively to produce smooth, clinically interpretable HbA1c variation curves. The SDD line represents the average intra-patient variation (including biologic and analytic variation) of the entire patient cohort (5). Significantly different A1c levels (P<0.05) in an individual HbA1c would be indicated by a change of more than 1.96 SDD for the time interval between the two serial tests. As usual, the grouped intra-patient data pairs are transformed with the SDD calculation. Compared to the short term SDD, this maximization of data points enables extension of the SDD line over longer durations. Regression lines may be interpolated with these data and represented by a polynomial which initially is near linear which and then transforms into a curved, gradually increasing line. The regression intercept at time 0 represents a mixture of biologic and analytic variation. In this paper, we compare the LT SDD of three different HbA1c methods and demonstrate diminished analytical accuracy. We present the following article in accordance with the MDAR reporting checklist (available at http://dx.doi.org/10.21037/jlpm-2019-qc-02).

**Methods**

**Patient data**

Prior to data receipt, patient identifiers were replaced by unique codes to preserve the patient links to the HbA1c results and collection date and time. We obtained 40,000 HbA1c results from 19,000 patients analyzed by the Sebia Capillarlys 2 Flex Piercing® assay between October 1, 2012 and October 1, 2014 at Hôpital De Chicoutimi in Chicoutimi Quebec. We obtained 121,000 HbA1c from 53,000 patients analyzed between December 20, 2013 and December 21, 2016 by the Roche Tinia Quant HbA1c Gen II assay on either Roche Cobas 8000, c502 or Roche Cobas 6000, c501 located at the Dartmouth-Hitchcock Medical Center in Lebanon, NH. Finally, 3.5 years of Siemens Vista HbA1c were obtained from The Ottawa Hospital (226,000 results from 129,749 patients obtained between April 1, 2015 and August 20, 2018). At all three testing centers, the vast majority of the HbA1c testing represented outpatients.

**External quality assessment (EQA) data**

The 2017 and 2018 CAP GH5 Survey Data (obtained from the thrice yearly analysis of five pooled fresh specimens) were used to abstract and summarize the imprecisions of low HbA1c samples (<7.0%) analyzed by any of the Roche 500, Sebia and Siemens Vista analytical systems.

**Analysis**

We graphed the methods’ LT intra-patient SDD variation and the corresponding 4th degree polynomial regressions for 4 HbA1c ranges: (I) 1stP to 99thP; (II) 75thP to 99thP (high results, consistent with poor control); (III) 25th to 75th P (middle of the road results); and 1thP to 25thP (the lowest results). We excluded highly abnormal HbA1c (>99thP or <1thP) as these outliers might artefactually elevate the SDD. In our SDD calculations, our shortest time interval was 1 week.

Based on the mixture of biologic and analytic variation for the <25thP population and the 25thP to the 75thP population, we postulated combinations of likely biologic and analytic variations that could explain the observed HbA1c variation at 26 weeks which represents a typical interval for repeating HbA1c. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). Ethics approval was not required as all data were anonymized before being provided to the investigators.
Individual consent for this retrospective analysis, analysis of deidentified patient data, was waived.

**Results**

Figure 1 compares the LT SDD for the different data ranges. Figure 1A encapsulates the variation from virtually all of the patient results, representing the inner 98th percentile of the three tests. For short time intervals between tests (up to 6 months), there is considerable overlap in the sum of the biologic and analytical variations of the Sebia and Roche assays. For longer intervals, far fewer Sebia HbA1c are used in the LT SDD calculation compared to the Roche data. The graphs of the Ottawa patients using the Vista assay are markedly different from those of the Sebia and Roche. Except for the 75th to 99th percentile data, the Vista data yield the highest y intercepts as well as the highest variations at 26 weeks. These high intercepts and the negative slopes are largely due to early repeats of unexpected (presumably high) results. The much higher LT SDD is obvious for HbA1c that are repeated within 26 weeks. There are far fewer early Vista repeats in patients whose results reflect the 75th to 99th percentile indicating physician acceptance of these high values. The differences in variation between the Vista and Roche or Vista and Sebia are most obvious for the low and middle A1c groups. At 26 weeks, the Vista variation exceeds Sebia’s by 25% to 29%.

For all four graphs, the Sebia and Roche assays exhibit...
similar LT SDD. At about 80 weeks, there is an artefactual excursion of the Sebia SDD points and instability in the regression line due to paucity of paired Sebia HbA1c data separated by longer time intervals. The CAP PT group summary reports from 2017 and 2018 indicate that the Sebia is the most precise test [average coefficient of variation (CV) =–2.02%±0.23% standard deviation (SD)] followed by Roche: 2.62%±0.34% and Vista: 2.93%±0.51%. Sebia's imprecision is lower than the Vista (P<0.0004) and the Roche (P<0.0002) assays.

Table 1 shows combinations of likely biologic variation and algebraically matched analytic variation that would generate the LT SDDs at 26 weeks for the low and middle ranges of HbA1c. The magnitude of analytical and biologic variation is highest with the Vista. A 1.5% biologic variation works well for the low HbA1c group [close to the 1.7% from reference (18)] with the Vista imprecision approximately 3% and the Roche and Sebia approximately 2%. For the middle HbA1c group, a biologic variation of 3.0% rounded average of 3 and 12 months variation of well controlled DCCT patients (19).

**Table 1** Combinations of likely intra-patient biological variation CVs and algebraically-derived analytic CVs equal to the instrument's SDD value at 26 weeks (converted to a relative CV by dividing by median HbA1c)

<table>
<thead>
<tr>
<th>Patient subpopulation</th>
<th>Presumptive CV</th>
<th>26 weeks total Sebia CV</th>
<th>26 weeks total Roche CV</th>
<th>26 weeks total Vista CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1stP to 25thP: 4.8% to 5.6%</td>
<td>1.5%</td>
<td>2.7%</td>
<td>2.2%</td>
<td>2.7%</td>
</tr>
<tr>
<td>25thP to 75thP: 5.6% to 7.2%</td>
<td>3.0%</td>
<td>3.8%</td>
<td>2.3%</td>
<td>3.7%</td>
</tr>
</tbody>
</table>

CV, coefficient of variation; SDD, standard deviation of duplicate; HbA1c, hemoglobin A1c.

**Discussion**

Laborators work in a new era of data generation, collection, analysis and utilization. Our laboratory analyzers produce prodigious amounts of data that the laboratorian should transform to enhance quality management, quality assurance and quality control. Our approach to deriving total variation makes far fewer assumptions than those laboratorians invoking the classical model of interaction between biologic variation and analytical variation. Usually, in this model, the analytic variation of an individual assay is exactly known and constant. This premise is overly simplistic as analytic variation can increase over the weeks or months that the assay is in use. This phenomenon is continually demonstrated in the short and longer term precision studies required by the various regulatory agencies. Some of the variation is associated with unstable or aging reagents, imprecise calibration, changing instrument conditions, between reagent lot variation, nonoptimal environmental control and new or distracted staff. In busy laboratories that maintain multiple instruments for reporting HbA1c, the operating conditions of two or three “identical” systems will not be identical (e.g., different column conditions over the lifetime of the separation column). As such, between instrument variation is associated with increased intra-patient variation if a patient's specimen is run on an alternate analyzer. Statistical quality control is usually assessed by comparing one or more quality control observations to an acceptable range. The physician assesses the serial HbA1c in two different ways. She similarly evaluates the serial intra-patient specimens by comparing them to her range of acceptability, but she also compares the new HbA1c to the prior value. The time interval between these consecutive tests may exceed 3 or 6 months, a period that requires exceedingly tight analytic precision to formulate the best medical decisions.

The other tacit assumption of the biologic analytic model is that biologic variation constant is correct and varies little. In many biologic variation studies, there probably exists a powerful self-selection bias that tends to select individuals with higher social status and healthier life styles (20). This bias tends to deselect individuals with lower social economic status or unhealthy lifestyles including physical inactivity, unhealthy diets, smoking and obesity. In the absence of this self-selection, the biologic variation constant (for many tests) would be a broader interval and not a constant. Even the timing of the study probably tends to artefactually reduce the biologic variation. As a rule, the longer the period that biologic variation data are collected, the larger will be the biologic variation. Yearly studies will encompass seasonal variation. As intimated previously, there is a imputed reluctance to perform LT studies and especially in December due to frequent public holidays which would interfere with scheduled blood draws. Interestingly, there
are higher rates of metabolic syndrome in December, associated with increased hyperglycemia, hypertension and hyperlipidemia (21). These reasons might explain the very low s, observed in Carlsen’s type 1 patients (18), a s, approaching that of healthy subjects without diabetes (22).

Figure 1 shows the association of the mixture of analytical and biologic variation and HbA1c level for each assay. Our unique LT SDD HbA1c graphs demonstrate that lower analytic imprecision reduces artefactual patient HbA1c variation. In patients with lower HbA1c levels, (1°P to 25°P and the 25°P to 75°P graphs), analytic error contributes proportionally more to the variation and can either mask patient improvement or falsely demonstrate increased hyperglycemia. Replacement of a suboptimal HbA1c assay allows the clinician to make improved therapeutic decisions. Both the low imprecision Sebia and Roche assays provide superior, actionable information. This work indicates that low imprecision HbA1c assays [CV ≤2.3% (Table 1)] will indicate patient glycemia more accurately than those with higher imprecisions. These findings are similar to those of Lenters-Westra et al. who recommended the limit of imprecision of 2.4% to accurately detect a HbA1c change of 0.5% (23). Study of the Bias/Between Laboratory CV graphs of Weykamp et al. (24) indicates an embarrassment of high CV methods (>2.4%) that are impairing the clinician’s view of the true metrics of glycemia.

Braga and others have written on the role of EQA and post market surveillance of in vitro medical diagnostics (25). One of the principle tasks of today’s EQA organization is the submission of “traceable”, “commutable” mixtures of analytes to participating laboratories for measures of precision and sometimes accuracy. We suggest that an EQA organizations embrace “big data”, resolve the multitude of privacy concerns and begin to analyze health care organizations’ large laboratory data sets. Initially, both the user and vendor might be dismayed by assay quality or the levels of overtesting.

It is our vision that short term and LT SDD calculation software will be eventually be available on all laboratory information systems as well as even on separate analyzers. For analytes like electrolytes, metabolites and blood gases, the short term SDD will quickly provide information about within-day assay quality. The LT SDD will provide estimates of LT intra-patient variation. Not only can the analyses be stratified by HbA1c level, they could be stratified by the patient’s diabetes diagnosis: type 1, type 2 or screening status. The resulting variations will provide the laboratorian and clinician the most realistic view of HbA1c assay performance. In fact, individual patient’s HbA1c variation can be determined from these large data sets and used to set the frequency of HbA1c measurement (26) or even route the patient for follow up testing (27).

In December 2019, Siemens announced the availability of a more accurate HbA1c test for the Vista Analyzer.

Acknowledgements

We would like to thank Dr. Jennifer Taher, Clinical Biochemist at Mount Sinai Hospital in Toronto, for reviewing the manuscript. This work has been conducted in part for the support of the US patent, altering patient care based on long term SDD, publication number: 20190035490 (authors JM and GC).

Finding: None.

Footnote

Provenance and Peer Review: This article was commissioned by the Guest Editor (Mark A. Cervinski) for the series “Patient Based Quality Control” published in The Journal of Laboratory and Precision Medicine. The article was sent for external peer review organized by the Guest Editor and the editorial office.

Reporting Checklist: The authors have completed the MDA reporting checklist. Available at http://dx.doi.org/10.21037/jlpm-2019-qc-02

Data Sharing Statement: Available at http://dx.doi.org/10.21037/jlpm-2019-qc-02

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form, available at http://dx.doi.org/10.21037/jlpm-2019-qc-02. The series “Patient Based Quality Control” was commissioned by the editorial office without any funding or sponsorship. MAC served as the unpaid Guest Editor of the series and serves as an unpaid editorial board member of The Journal of Laboratory and Precision Medicine from May 2019 to April 2021. CM serves as an unpaid editorial member of The Journal of Laboratory and Precision Medicine from Feb 2019 to Jan 2021. Dr. GC reports in addition, Dr. GC has a patent Applied for US patent for long term SDD calculation pending. The other authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all
aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). Ethics approval was not required as all data were anonymized before being provided to the investigators. Individual consent for this retrospective analysis, analysis of deidentified patient data, was waived.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References


