Performance of Hemoglobin A\textsubscript{1c} Assay Methods: Good Enough?
Randie R. Little\textsuperscript{1}\textsuperscript{*}

Two very interesting articles on hemoglobin A\textsubscript{1c} (Hb A\textsubscript{1c}) method performance appear in this issue of \textit{Clinical Chemistry}. Woodworth et al. \textit{(1)} use Hb A\textsubscript{1c} as an example of how the risk of reporting inaccurate results can be estimated to guide individualized QC strategies. They assess the risk of reporting unreliable Hb A\textsubscript{1c} results with 6 different methods in 4 academic medical centers; 5 are laboratory instruments and 1 is a point-of-care (POC) method. Patient-weighted metrics were calculated on the basis of 1–2\textsubscript{a} QC rules with an average of 1 QC for every 100 samples. The authors found considerable differences in the risk of reporting unreliable results among the methods used at these sites, and that only 1 method, the Capillarys\textsuperscript{2} Flex Piercing, achieved a patient-weighted \( \sigma \) metric \( \geq 3 \) at a total error allowable (TE\textsubscript{a}) of 6%. This total error limit is also used for College of American Pathologists (CAP) proficiency testing and National Glycohemoglobin Standardization Program (NGSP) certification and is considered optimal for clinical care at this time. This means that for all but the aforementioned method, “maximum QC (three levels, three times per day) should be performed to achieve the necessary error detection.” Of note was that for some methods there was a substantial amount of bias (up to 5.8% relative for values within the reference interval), and CV values were higher than the recommended 2% for either the high or low QC for 3 of the methods studied \textit{(2)}. The maximum number of Hb A\textsubscript{1c} results out of 100 expected to be unreliable from an out-of-control condition at TE\textsubscript{a} of 6% ranged from 0.60 to 71.48 \textit{(1)}, and for 2 methods the number of results expected to be unreliable was \( > 19 \) of 100 even when the method would be considered in control (mainly due to high bias).

The second article on Hb A\textsubscript{1c} method performance by Lenters-Westra and Slingerland \textit{(3)} includes an evaluation of 7 Hb A\textsubscript{1c} POC methods. This study also finds large differences in performance among methods, with CV values varying from \(< 1\% \) (better than most of the laboratory methods in the Woodworth study \textit{(1)}) to \( > 3\% \). Precision evaluations in both studies followed the Clinical and Laboratory Standards Institute (CLSI) EP-5 protocol, so the CV values reported in the 2 studies should be directly comparable. Overall, the CV values for the POC methods overlapped those for the laboratory methods, and the CV values for the DCA Vantage (POC method) included in both studies were similar, albeit slightly above the recommended 2% for within-laboratory variability. Overall, the imprecision of all the methods evaluated in both studies varied considerably, with CV values ranging from 0.8% to 3.2% when evaluated across the measurement range.

In both studies, bias was calculated on the basis of comparison to NGSP/IFCC Reference Laboratories, so the reference point in both studies is also the same. Although the way in which bias is reported is not the same (percentage bias at 2 levels vs mean absolute Hb A\textsubscript{1c} bias across the entire range), clearly, in both studies, there are large biases for some methods and almost no bias for others. All the methods evaluated in both studies were NGSP certified (on the basis of comparison data submitted by the manufacturer) at the time of each study, so one would hope that they would be able to perform at the same level in the laboratory. However, Lenters-Westra and Slingerland \textit{(3)} found that 4 of the 7 methods evaluated in their laboratory would fail the certification criteria. It is worth noting that these evaluations were performed in a laboratory with experienced personnel; because these methods are all CLIA waived, there are no mandates that the end users be trained laboratory personnel or perform proficiency testing. Clearly, as noted in previous studies \textit{(4, 5)}, some methods that can perform well enough to pass NGSP certification when testing is performed by the manufacturer do not consistently achieve the same level of performance in the field. Also of note is that in both studies, bias rather than imprecision seemed to be the major factor when methods did not perform well; this would seem to indicate that lot-to-lot variations between reagents and/or calibrators may play a significant role. Although there were no substantial differ-

---

\textsuperscript{1} Diabetes Diagnostic Laboratory, Department of Pathology and Anatomical Sciences, University of Missouri School of Medicine, Columbia, MO.

\textsuperscript{*} Address correspondence to this author at: Diabetes Diagnostic Laboratory, Department of Pathology & Anatomical Sciences, University of Missouri School of Medicine, 1 Hospital Dr. M766, Columbia, MO 65212. Fax 573-884-4748; e-mail littler@health.missouri.edu.

Received May 23, 2014; accepted May 28, 2014.

Previously published online at DOI: 10.1373/clinchem.2014.225789

© 2014 American Association for Clinical Chemistry

\textsuperscript{2} Nonstandard abbreviations: Hb A\textsubscript{1c}, glycated hemoglobin; POC, point-of-care; TE\textsubscript{a}, total allowable error; CAP, College of American Pathologists; NGSP, National Glycohemoglobin Standardization Program; CLSI, Clinical and Laboratory Standards Institute.
ences between tested lots for the POC methods evaluated in either of the present studies, such differences have been observed in previous studies (4, 6–9). It would have been interesting to see if the laboratory methods evaluated in the Woodworth et al. study (1) would have passed NGSP certification. Clearly, many methods, including a few POC methods, do perform well in laboratories, as seen by data from the CAP GH-2 proficiency surveys.

Both articles show a wide range of performance among Hb $A_{1c}$ methods. The current NGSP criterion is stringent; at least 37 of 40 sample results (for samples in the 4%–10% Hb $A_{1c}$ range) must be within 6% of an NGSP network laboratory for a method to be certified, and method certification must be renewed annually. However, certification is performed under optimal conditions by the manufacturer, typically using 1 lot of reagents. Thus, NGSP certification shows that a Hb $A_{1c}$ method is capable of excellent performance, but cannot guarantee that the method will consistently achieve the same performance in the hands of end users. A method may show little bias for the lot of calibrators/reagents used during certification, but if quality assurance processes at the manufacturer level cannot consistently maintain this level of performance from 1 lot of reagents or calibrators to the next, substantial lot-to-lot variations can be observed. Proficiency testing assesses the performance of large numbers of laboratories, usually using several different lots of reagents, and thus provides a measure of both within- and between-method variability as well as bias from a reference-assigned value. An important consideration in proficiency testing, especially for accuracy-based assessments such as the CAP GH-2 Hb $A_{1c}$ survey, is use of matrix-appropriate materials to minimize the potential for matrix effects that can occur when processed materials (e.g., lyophilized blood) are used. The CAP survey for Hb $A_{1c}$ uses fresh whole blood material with NGSP value assignments. This allows for accurate evaluation of within- and between-method precision and bias. This GH2 survey is provided twice a year with Hb $A_{1c}$ pooled whole blood samples at 3 Hb $A_{1c}$ levels. The CAP LN15 Hb $A_{1c}$ Calibration Verification/Linearity survey is also provided twice a year with the same sample type and value assignments (6 levels are provided). The most recent GH2 survey included >3000 laboratories using >30 different methods (10).

In selecting a Hb $A_{1c}$ method, laboratories can first determine if a method is NGSP certified (potential for optimal performance) and then evaluate CAP data (both current and past surveys) to assess how well the method actually performs in the field. If a certified method shows large CV values on >1 CAP survey, it suggests that the method will not perform well over time in an individual laboratory. Similarly, a method with a large bias on several surveys will likely show the same bias in an individual laboratory. For example, the bias observed by Woodworth et al. for the Tosoh G8 is consistent with what recent CAP surveys have shown for this method (1), and the consistent good performance of the DCA Vantage seen on several CAP surveys is also evident in both of the current studies as well as in previous observations (11).

Choosing a method that is NGSP certified and has performed well on proficiency testing is very important but still does not guarantee optimal performance in every laboratory. As discussed in Woodworth et al. (1), each laboratory must develop optimal QC practices and also evaluate their method continuously. Although each laboratory may not be able to accurately calculate their risk of reporting unreliable results (as in Woodworth et al.), they must ensure that they have good QC practices, that their CV values are consistently <2%, and they have minimal bias on the basis of comparison with a reference, either by direct comparison with NGSP (e.g., individual laboratory certification and monitoring) or through the CAP GH2 and/or LN15 surveys or a comparable accuracy-based Hb $A_{1c}$ survey. Lesters-Westra and Slingerland (3) recommend that users of POC methods be required to perform proficiency testing; if this proved to be feasible, it would be a major step in ensuring that these methods are producing results that are consistently accurate. The fact that Hb $A_{1c}$ is now recommended for diagnosis of diabetes and prediabetes (12), as well as monitoring of mean glycemia, reinforces the requirement that Hb $A_{1c}$ results be accurate. This need for accuracy has led to the tightening of criteria for passing both NGSP certification and CAP proficiency testing. Hb $A_{1c}$ is an important test for the diagnosis and ongoing monitoring of a disease that is increasing in prevalence; laboratories must be vigilant in ongoing quality assurance and monitoring of their Hb $A_{1c}$ assays to achieve the levels of performance demanded by the clinical community to ensure optimal patient care.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 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.

Authors’ Disclosures or Potential Conflicts of Interest: Upon manuscript submission, all authors completed the author disclosure form.

Disclosures and/or potential conflicts of interest:

Employment or Leadership: None declared.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: None declared.
Expert Testimony: None declared.
Patents: None declared.
Other Remuneration: Roche, Bio-Rad, Arkray, and Abbott.
References