Analysis of a Point-of-Care HbA1c Assay: Is It Time for an HbA1c Error Grid?

Journal of Diabetes Science and Technology
2020, Vol. 14(5) 896–897
© 2019 Diabetes Technology Society
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1932296819861983
journals.sagepub.com/home/dst

Jan S. Krouwer, PhD¹

Abstract

In an article in *Journal of Diabetes Science and Technology*, Arnold et al have presented a thorough study of imprecision components for a point-of-care hemoglobin A1c (HbA1c) assay. An interesting and innovative approach is the combination of data from different studies to arrive at a total error estimate. But total error has the oxymoron feature of estimating performance for most (95%) but not all of the results. An HbA1c error grid would provide the severity for results that exceed the 6% requirement. Since this device is intended for Clinical Laboratory Improvement Amendments waived labs and allows for finger-stick samples, monitoring the Food and Drug Administration adverse event database (MAUDE, Manufacturer and User Facility Device Experience) is recommended.

Keywords

error grid, HbA1c, MAUDE, POC, total error

Arnold et al have presented a thorough study of imprecision components for a point-of-care (POC) hemoglobin A1c (HbA1c) assay.¹ This study provides valuable information. The following comments may help future evaluation studies.

There are 3 purposes for an assay—screening, diagnosis, and monitoring. HbA1c is intended for diagnosis and monitoring but the authors in this study mention only diagnosis in their article. This is confusing—would not this assay also be suitable for monitoring?

It is mentioned that precision and total error were investigated. This is also somewhat confusing since if one has evaluated total error and found it to be acceptable, then the precision must also be acceptable because precision is part of total error.

The problem with estimating total error to determine performance is that total error as calculated by the authors (which is the usual way it is calculated) is a probability metric because it involves multiples of the standard deviation (or coefficient of variation) that correspond to percentages of results. Thus, the authors' formula provides the total error for 95% of the results. This leaves 5% of the results without being characterized for performance. To be fair, the venerable NGSP (National Glycohemoglobin Standardization Program) standard also requires only a percentage of results (90%) to be acceptable.² Why should 5% (this study) or 10% (NGSP) be left out of performance specifications? Since total implies everything, how can *total* error only include 95% of the results?

One potential remedy is to create an error grid for HbA1c. If in fact, there are results beyond the currently required total error of 6%, an error grid would have zones that provide the severity of false-positive and false-negative results. And these

zones are not constrained to one HbA1c concentration—the zone values can change as a function of the HbA1c concentration. Since an error grid is so widely accepted for glucose meters, it would seem likely to be accepted for HbA1c. One might need different error grids for HbA1c diagnosis vs monitoring. An error grid for HbA1c was provided recently.³

As mentioned previously, most evaluations use relatively small sample sizes and rarely produce outliers.⁴ This does not invalidate the study—rather one can consider the estimated performance to be for typical results. That is, this is the performance one can expect most of the time (eg 95%) with errors caused by imprecision and average bias. Rare errors or failure to obtain a result often have other causes (user error, manufacturing problems, and hardware issues). A source to examine these errors is the FDA MAUDE database (Food and Drug Administration Manufacturer and User Facility Device Experience database).⁵ The good news is that only 3 events (erratic values) were found for the authors' product in the last 2 years. Since it is implied that this POC assay may achieve Clinical Laboratory Improvement Amendments (CLIA) waived status, the number of user errors may increase especially if the number of finger-stick samples increases. The authors have shown that finger-stick samples are less precise than venous samples. As the authors imply, personnel in CLIA waived labs might have more user error. This is another reason to monitor the MAUDE database.

¹Krouwer Consulting, Sherborn, MA, USA

Corresponding Author:

Jan S. Krouwer, PhD, Krouwer Consulting, 26 Parks Drive Sherborn, MA 01770, USA.

Email: jan.krouwer@comcast.net

Another aspect of the authors' study is the combination data from different studies to arrive at a total error estimate. This is an interesting and innovative approach.

Finally, it is also informative to have one or more graphs of the data (none were provided) in spite of the fact that total error was calculated from different studies.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Jan S. Krouwer <https://orcid.org/0000-0003-2300-076X>

References

1. Arnold WD, Kupfer K, Swensen MH, et al. Fingerstick precision and total error of a point-of-care HbA1c test. *J Diabetes Sci Technol*. 2020;14:890-895.
2. NGSP (National Glycohemoglobin Standardization Program). <http://www.ngsp.org/critsumm.asp>. Accessed May 27, 2019.
3. Jiang F, Hou X, Lu J, et al. Assessment of the performance of A1CNow+ and development of an error grid analysis graph for comparative hemoglobin A1c measurements diabetes. *Technol Ther*. 2014;16:363-369.
4. Krouwer JS, Garrett PE. Getting more information from glucose meter evaluations. *J Diabetes Sci Technol*. 2019;13:1175-1177.
5. FDA MAUDE (Manufacturer and User Facility Device Experience). <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfmaude/search.cfm>. Accessed June 2, 2019.