UC Davis
UC Davis Previously Published Works

Title

The Need for Standardization of Continuous Glucose Monitoring Performance Evaluation: An Opinion by the International Federation of Clinical Chemistry and Laboratory Medicine Working Group on Continuous Glucose Monitoring.

Permalink

https://escholarship.org/uc/item/1th5053d

Authors

Pleus, Stefan
Eichenlaub, Manuel
Eriksson Boija, Elisabet
et al.

Publication Date

2024-11-14

DOI

10.1177/19322968241296097

Peer reviewed

The Need for Standardization of Continuous Glucose Monitoring Performance Evaluation: An Opinion by the International Federation of Clinical Chemistry and Laboratory Medicine Working Group on Continuous Glucose Monitoring

Stefan Pleus, PhD[1,2] iD, Manuel Eichenlaub, PhD[2] iD,
Elisabet Eriksson Boija, PhD[1,3], Marion Fokkert, PhD[1,4],
Rolf Hinzmann, MD, PhD[1,5] iD, Johan Jendle, MD, PhD[1,6],
David C. Klonoff, MD, FACP, FRCP (Edin), Fellow AIMBE[1,7] iD,
Konstantinos Makris, PhD[1,8] iD, James H. Nichols, PhD[1,9] iD,
John Pemberton, BSc, RD[1,10], Elizabeth Selvin, PhD, MPH[1,11],
Robbert J. Slingerland, PhD[1,4], Andreas Thomas, PhD[1,12] iD,
Nam K. Tran, PhD[1,13] iD, Lilian Witthauer, PhD[1,14,15] iD, and
Guido Freckmann, MD[1,2] iD; on behalf of the Working Group on
Continuous Glucose Monitoring of the IFCC Scientific Division

Abstract

Metrics derived from continuous glucose monitoring (CGM) systems are often discordant between systems. A major cause is that CGM systems are not standardized; they use various algorithms and calibration methods, leading to discordant CGM readings across systems. This discordance can be addressed by standardizing CGM performance assessments: If manufacturers aim their CGM systems at the same target, then CGM readings will align across systems. This standardization should include the comparator device, sample origin, and study procedures. With better aligned CGM readings, CGM-derived metrics will subsequently also align better between systems.

Keywords

continuous glucose monitoring, IFCC, standardization

[1]Working Group on Continuous Glucose Monitoring, Scientific Division, The International Federation of Clinical Chemistry and Laboratory Medicine, Milano, Italy
[2]Institut für Diabetes-Technologie, Forschungs- und Entwicklungsgesellschaft mbH an der Universität Ulm, Ulm, Germany
[3]Equalis AB, Uppsala, Sweden
[4]Department of Clinical Chemistry, Isala Clinics, Zwolle, The Netherlands
[5]Lampertheim, Germany
[6]School of Medical Sciences, Faculty of Medicine and Health, Örebro University, Örebro, Sweden
[7]Diabetes Research Institute of Mills-Peninsula Medical Center, San Mateo, CA, USA
[8]Clinical Biochemistry Department, KAT General Hospital, Athens, Greece
[9]Vanderbilt University Medical Center, Nashville, TN, USA
[10]Birmingham Women's and Children's NHS Foundation Trust, Birmingham, UK
[11]Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA
[12]Independent Scientific Consulting, Pirna, Germany
[13]Department of Pathology and Laboratory Medicine, University of California Davis Health, Sacramento, CA, USA
[14]Diabetes Center Berne, Bern, Switzerland
[15]Department of Diabetes, Endocrinology, Nutritional Medicine and Metabolism, Inselspital Bern, University of Bern, Bern, Switzerland

Corresponding Author:
Stefan Pleus, PhD, Institut für Diabetes-Technologie, Forschungs- und Entwicklungsgesellschaft mbH an der Universität Ulm, Lise-Meitner-Strasse 8/2, D-89081 Ulm, Germany.
Email: stefan.pleus@idt-ulm.de

Introduction

In response to the increasing use and utility of systems for continuous glucose monitoring (CGM), clinical metrics and targets for CGM data interpretation were published in 2019.[1] These CGM-derived metrics are intended to guide therapeutic recommendations in clinical practice. Over the past few years, multiple studies found that CGM-derived metrics differ between different brands of CGM systems when worn by the same person.[2-7] This is exemplified in panel (a) of Figure 1, where three current factory-calibrated CGM systems from different manufacturers worn by the same person yielded marked differences in times below range (TBR), times in range (TIR), and times above range (TAR) over the course of a single day due to systematic differences between the systems (panels (b)-(d)). A more systematic comparison between Dexcom G5 and FreeStyle Libre found that 11 out of 24 participants had a TBR above 4%, which is the recommended maximum TBR, with one system, and below 4% for the other system, which could have led to diverging therapeutic recommendations.[5]

In a recent response to a study comparing the clinical outcomes of two systems for automated insulin delivery (AID) using different CGM sensors,[8] Messer and colleagues,[9] two of whom worked for a manufacturer of CGM systems at the time of publication, emphasized that "CGM-derived metrics should not be compared between CGM systems as *the instruments of measure are fundamentally different*" (emphasis in original text).

Continuous glucose monitoring-derived metrics are calculated from a large number of individual CGM readings. If CGM-derived metrics are substantially different, then the CGM readings themselves will also be substantially different. As a result, glycemic control and clinical decision-making could differ depending on the specific CGM system being worn.

Potential Causes for Discordant CGM Readings

While physiological factors, such as differences in sensor placement and perfusion into the interstitial space, can cause discordant CGM readings in the same individual, technical factors may be more pronounced and are more likely to be addressed through engineering solutions. These technical factors encompass the sensor architecture, including the overall design, materials used in its construction, manufacturing process, and the sensor chemistry itself. However, differences in the calibration algorithms converting the raw signal to CGM readings are likely the main factor leading to discordant CGM readings. These differences are exemplified in panels (b) to (d) of Figure 1 and can be characterized by a constant offset (bias) possibly in combination with a glucose level-dependent over- or underestimation. Time lags may also differ across sensors due to these technical factors and different algorithms.

Manufacturers of CGM systems assess the performance of their systems based on blood glucose measurements. However, the choice of sample origin, the comparator device, and the study procedures used to obtain validation data can strongly influence the results. For example, it is known that the glucose concentrations can be physiologically different in capillary, venous, and "arterialized"-venous (ie, from a heated hand/ arm) blood samples.[10] In addition, it has been demonstrated that different comparator methods as well as comparator devices from the same brand can exhibit systematic differences, even when the same blood samples are used.[11-14] Therefore, depending on the specific device and sample type (ie, matrix, compartment, and handling) used to establish the calibration algorithms of CGM systems, the CGM readings of different devices/brands can be expected to be different by several percent.

Possible Consequences of Discordant CGM Readings

As stated above, glycemic control and clinical decision-making could differ depending on the specific CGM system being worn. However, people with diabetes should be provided with roughly the same results and information regarding glycemic control, glycemic variability, and hypo-/ hyperglycemia independent from the CGM system being used. Otherwise, therapeutic guidelines that incorporate threshold values or aims for glycemic status may not be effective or even appropriate.

Simply acknowledging that differences between CGM system brands exist is not an adequate solution, because manufacturers do not disclose which kind of glucose concentration readings their CGM systems represent (similar to metrological traceability), for example, capillary-like or venous-like concentrations.[15] The establishment of device-specific targets for CGM-derived metrics is also not a viable alternative, because targets would have to be established for each new CGM system, and potentially each new generation of established CGM systems. In addition, device-specific targets would needlessly complicate the assessment of electronic health records as well as their use by national health systems, payers, and regulators. For AI-driven analytical approaches to electronic health records, having clean and comparable CGM data would also be a benefit.

How to Proceed?

Better harmonization of CGM systems can be tackled through standardization. An example of another analyte measured in diabetes research and clinical practice, where standardization has improved the quality of measurements, is HbA1c.

For years, HbA1c was the de facto approach to assess a persons' glycemic status and therefore the principal parameter when assessing the efficacy of new diabetes treatments. However, the measurement results of different HbA1c assays
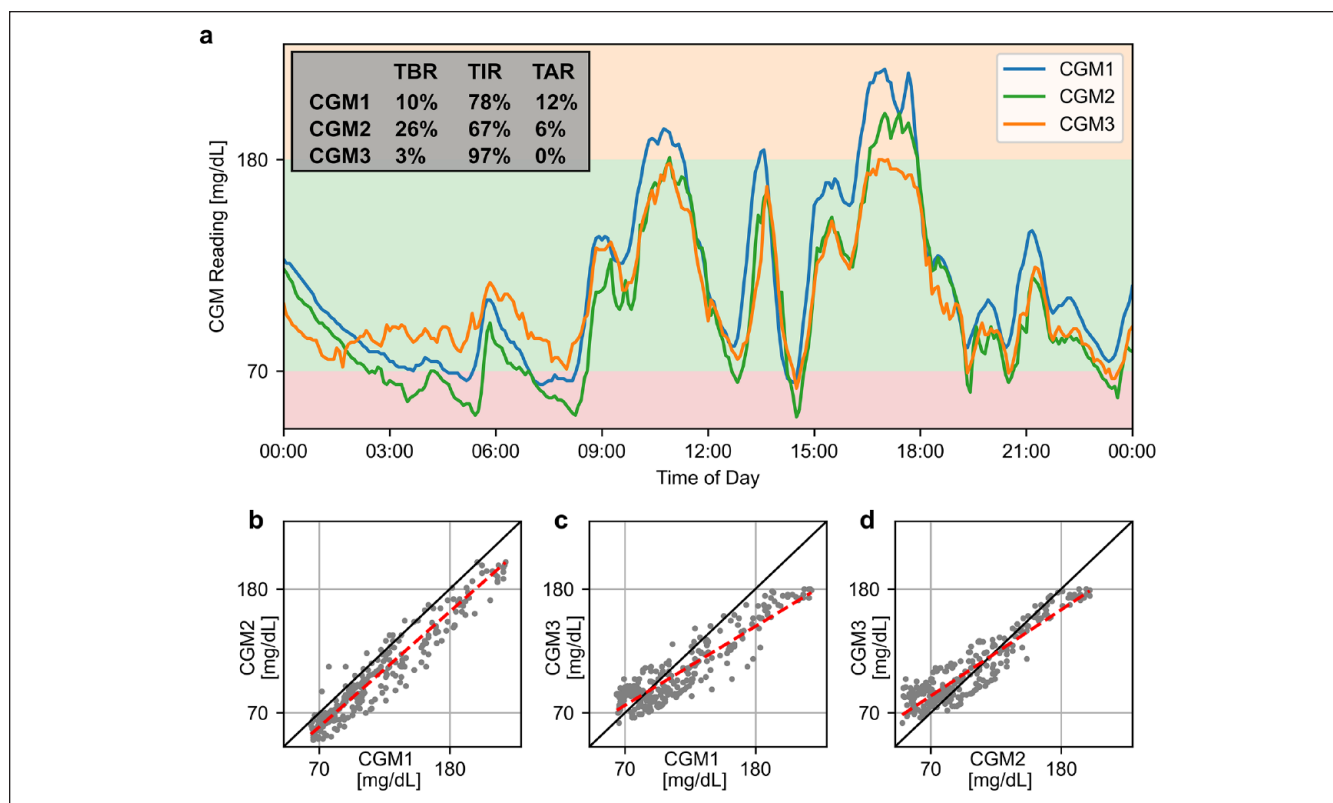
Figure 1. (a) CGM profiles of three factory-calibrated CGM systems from different manufactures simultaneously worn by the same person over the course of a single day. The box in the top left corner provides the respective percentages for the times below range (TBR), times in range (TIR), and times above range (TAR). (b)-(d) Weighted Deming regression plots between readings from different CGM systems illustrating systematic differences in CGM readings (constant offset and/or glucose-dependent over- or underestimation): The black lines show the line of identity, and the red dashed lines display the linear regression fit.

used to lack comparability, so that the widespread implementation of HbA1c targets in clinical practice was hindered. Ultimately, this led to a standardization program of the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) as well as the founding of the National Glycohemoglobin Standardization Program (NGSP),[16,17] and while the road toward HbA1c standardization may have been bumpy, immense progress has been made over the years, so that HbA1c measurements are harmonized across many countries. This harmonization allowed diabetes associations to add HbA1c as a tool for the diagnosis of diabetes.

A similar level of standardization has not been achieved for CGM systems as demonstrated not only by the differences between CGM-derived metrics but also by variations in comparator data origin between manufacturers as reported in a recent review of CGM performance studies.[18] Previous efforts at standardization include the POCT05 guideline[19] published by the Clinical and Laboratory Standards Institute and the acceptance criteria for "integrated" CGM systems established by the US Food and Drug Administration.[20] However, both guidelines lack specific requirements regarding the collection and characteristics of comparator data.[18]

Study procedures are known to affect CGM performance as observed in a study, so acceptance criteria may be helpful but not a solution on their own. It should also be noted that these criteria for integrated CGM systems only apply to CGM systems digitally connected to other devices, like automated insulin dosing systems. Standalone CGM systems, especially adjunctive-use systems are not covered by this designation.

The IFCC has therefore established the Working Group on CGM with the primary goal of developing a standard for the assessment of CGM system performance. One of the main pillars of this standard will be the definition of procedures for the collection of comparator data in CGM performance studies. In particular, the standard will stipulate that the comparator data are obtained with comparator methods fulfilling analytical performance specifications, samples are collected from a specific origin, and that the distribution of comparator data complies with certain requirements which ensure that clinically relevant glucose concentrations and rates of change are represented.[14,21] Manufacturers will aim for optimal performance of their systems in CGM performance studies. Therefore, if the performance of all CGM systems is assessed using the same comparator
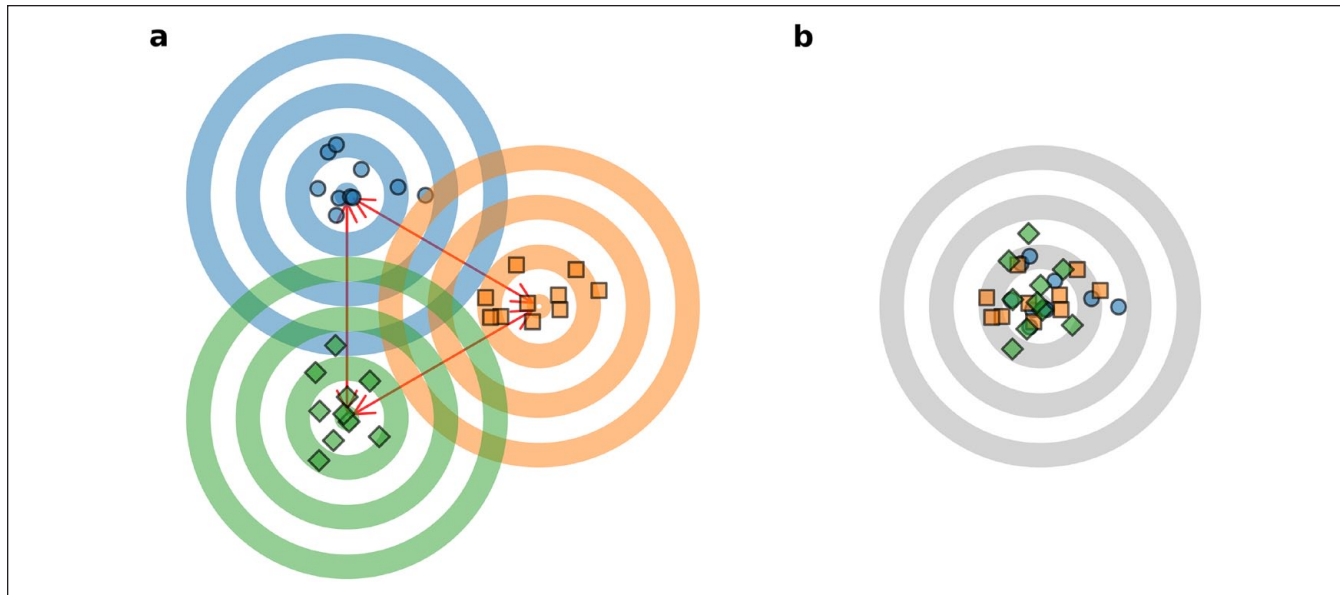
Figure 2. Illustration of the IFCC working group on CGM's approach to minimize discrepancies between CGM readings. Without standardization, each manufacturer uses their individual comparator procedure (target) and while readings of their own CGM system will scatter around their individual target, CGM readings from different systems are substantially discordant (panel (a)). With standardization, scattering of CGM readings around the target will not necessarily change, but readings from every CGM system will scatter around the same target and thus align better (panel (b)).

characteristics, then readings from different CGM systems will approach each other.

This approach used by the IFCC working group on CGM is illustrated in Figure 2. Currently, each manufacturer uses their own target. The bull's eye in the target represents the result obtained with the comparator method chosen by a particular manufacturer. CGM systems can exhibit good performance when compared with their specific comparators, while readings from different CGM systems can differ substantially (panel (a)). By defining a universal target, that is, a standardized procedure to collect comparator data, manufacturers will be encouraged to aim for the same target, leading to an alignment of readings from different CGM systems (panel (b)).

Abbreviations

AID, automated insulin delivery; CGM, continuous glucose monitoring; IFCC, International Federation of Clinical Chemistry and Laboratory Medicine; NGSP, National Glycohemoglobin Standardization Program; TAR, time above range; TBR, time below range; TIR, time in range.

Acknowledgment

The authors acknowledge the contributions of past and present members of the IFCC working group on CGM.

Declaration of Conflicting Interests

ORCID iDs

Stefan Pleus https://orcid.org/0000-0003-4629-7754
Manuel Eichenlaub https://orcid.org/0000-0003-2150-3160
Rolf Hinzmann https://orcid.org/0000-0002-8419-4740
David C. Klonoff https://orcid.org/0000-0001-6394-6862
Konstantinos Makris https://orcid.org/0000-0002-7896-9028
James H. Nichols https://orcid.org/0000-0002-3652-1612
Andreas Thomas https://orcid.org/0000-0002-6549-2793
Nam K. Tran https://orcid.org/0000-0003-1565-0025
Lilian Witthauer https://orcid.org/0000-0001-9459-875X
Guido Freckmann https://orcid.org/0000-0002-0406-9529

References

1. Battelino T, Danne T, Bergenstal RM, et al. Clinical targets for continuous glucose monitoring data interpretation: recommendations from the international consensus on time in range. *Diabetes Care*. 2019;42(8):1593-1603. doi:10.2337/dci19-0028.

2. Kumagai R, Muramatsu A, Fujii M, et al. Comparison of glucose monitoring between Freestyle Libre Pro and iPro2 in patients with diabetes mellitus. *J Diabetes Investig*. 2019;10(3):851-856. doi:10.1111/jdi.12970.

3. Michalak A, Pagacz K, Młynarski W, Szadkowska A, Fendler W. Discrepancies between methods of continuous glucose monitoring in key metrics of glucose control in children with type 1 diabetes. *Pediatr Diabetes*. 2019;20(5):604-612. doi:10.1111/pedi.12854.

4. Pleus S, Kamecke U, Waldenmaier D, et al. Time in specific glucose ranges, glucose management indicator, and glycemic variability: impact of Continuous Glucose Monitoring (CGM) system model and sensor on CGM metrics.

J Diabetes Sci Technol. 2021;15(5):1104-1110. doi:10.1177/1932296820931825.

5. Freckmann G, Pleus S, Schauer S, et al. Choice of continuous glucose monitoring systems may affect metrics: clinically relevant differences in times in ranges. *Exp Clin Endocrinol Diabetes*. 2022;130(5):343-350. doi:10.1055/a-1347-2550.

6. Yeoh E, Png D, Khoo J, et al. A head-to-head comparison between Guardian Connect and FreeStyle Libre systems and an evaluation of user acceptability of sensors in patients with type 1 diabetes. *Diabetes Metab Res Rev*. 2022;38(7):e3560. doi:10.1002/dmrr.3560.

7. Zhou Y, Mai X, Deng H, et al. Discrepancies in glycemic metrics derived from different continuous glucose monitoring systems in adult patients with type 1 diabetes mellitus. *J Diabetes*. 2022;14(7):476-484. doi:10.1111/1753-0407.13296.

8. Singh R, Imberg H, Ahmadi SS, et al. Effects, safety, and treatment experience of advanced hybrid closed-loop systems in clinical practice among adults living with type 1 diabetes. *J Diabetes Sci Technol*. 2024;0(0). doi:10.1177/19322968241242386.

9. Messer LH, Welsh JB, Habif S, et al. "Effects, safety, and treatment experience of advanced hybrid closed-loop systems in clinical practice among adults living with type 1 diabetes." *J Diabetes Sci Technol*. 2024;18(5):1265-1266. doi:10.1177/19322968241257003.

10. Pleus S, Link M, Hinzmann R, et al. Effect of arterialization on venous blood glucose concentrations and implications for observed continuous glucose monitoring accuracy. *Diabetes Technol Ther*. 2024;26(4):238-245. doi:10.1089/dia.2023.0489.

11. Twomey PJ. Plasma glucose measurement with the Yellow Springs Glucose 2300 STAT and the Olympus AU640. *J Clin Pathol*. 2004;57(7):752-754. doi:10.1136/jcp.2003.013417.

12. Freckmann G, Baumstark A, Jendrike N, et al. Impact of two different reference measurement procedures on apparent system accuracy of 18 CE-marked current-generation blood glucose monitoring systems. *J Diabetes Sci Technol*. 2022;16(5):1076-1088. doi:10.1177/1932296820948873.

13. Bailey TS, Klaff LJ, Wallace JF, et al. Fundamental importance of reference glucose analyzer accuracy for evaluating the performance of blood glucose monitoring systems (BGMSs). *J Diabetes Sci Technol*. 2016;10(4):872-875. doi:10.1177/1932296816634356.

14. Pleus S, Eichenlaub M, Gerber T, et al. Improving the bias of comparator methods in analytical performance assessments through recalibration. *J Diabetes Sci Technol*. 2024;18(3):686-694. doi:10.1177/19322968221133107.

15. Freckmann G, Nichols JH, Hinzmann R, et al. Standardization process of continuous glucose monitoring: traceability and performance. *Clin Chim Acta*. 2021;515:5-12. doi:10.1016/j.cca.2020.12.025.

16. Little RR, Rohlfing CL. The long and winding road to optimal HbA1c measurement. *Clin Chim Acta*. 2013;418:63-71. doi:10.1016/j.cca.2012.12.026.

17. Hinzmann R, Schlaeger C, Tran CT. What do we need beyond hemoglobin A1c to get the complete picture of glycemia in people with diabetes. *Int J Med Sci*. 2012;9(8):665-681. doi:10.7150/ijms.4520.

18. Freckmann G, Eichenlaub M, Waldenmaier D, et al. Clinical performance evaluation of continuous glucose monitoring systems: a scoping review and recommendations for reporting. *J Diabetes Sci Technol*. 2023;17(6):1506-1526. doi:10.1177/19322968231190941.

19. Clinical and Laboratory Standards Institute. *POCT05-Performance Metrics for Continuous Interstitial Glucose Monitoring*. Second edition. Wayne, PA: Clinical and Laboratory Standards Institute; 2020.

20. United Stated Code of Federal Regulations. Integrated continuous glucose monitoring system. https://www.ecfr.gov/current/title-21/section-862.1355

21. Eichenlaub M, Pleus S, Rothenbühler M, et al. Comparator data characteristics and testing procedures for the clinical performance evaluation of continuous glucose monitoring systems. *Diabetes Technol Ther*. 2024;26(4):263-275. doi:10.1089/dia.2023.0465.