



ORIGINAL ARTICLE

---

# Clinical Implications of Accuracy Measurements of Continuous Glucose Sensors

Timothy S. Bailey, MD, FACE, CPI

## Abstract

The accuracy of a continuous glucose monitor (CGM) now supports its use by persons with diabetes and clinicians caring for them. This article reviews measures of CGM accuracy, factors contributing to accuracy, comparative accuracy assessment, clinical implications of CGM sensor accuracy, and recent clinical trials that have demonstrated the utility of CGMs.

**Keywords:** Continuous glucose monitoring, Continuous home monitoring, Self-monitoring of blood glucose, Sensors.

SINCE THE 1970s, the ability of patients to monitor their glucose levels has increased dramatically, initially through capillary blood glucose monitors and now through continuous glucose monitors (CGMs).<sup>1</sup> In 1987, the ADA suggested the ambitious accuracy goal of 100% of values within 10% of a laboratory reference for glucose levels between 30 and 400 mg/dL.<sup>2</sup> Despite less than optimal accuracy, diabetes control has been transformed by patient self-monitoring of blood glucose (SMBG) and the decisions it enables. The DCCT<sup>3</sup> was as much a validation of the role of more frequent SMBG as it was of more frequent insulin injections. Microvascular complications were reduced by lower HbA1c levels, accompanied by a rise in hypoglycemia and burden of monitoring. Today, SMBG is accepted to be as integral to diabetes management as is looking through the windshield and at dashboard instrumentation when driving a car.

Today's CGMs allow frequent glucose monitoring through minimally invasive sensors that dwell in the subcutaneous space between 3 and 14 days. The accuracy of early sensors, similar to early SMBG, was less than that of current devices. However, in the near decade-old JDRF trial that used older and less-accurate CGM sensors adjunctively to SMBG, there was a clinically and statistically significant decline in both HbA1c and hypoglycemia.<sup>4</sup> In the predefined subgroup of patients with an initial HbA1c of <7.0%, there was a smaller HbA1c benefit, but significantly less hypoglycemia.<sup>5</sup> Based on improvement in sensor accuracy, regulatory bodies have

removed the adjunctive requirement and allow independent utilization of two CGMs for insulin dosing—the Abbott Freestyle Libre in 2014 in the EU (CE-marked) and the Dexcom G5 in 2016 in the United States.<sup>6</sup> Furthermore, automated sensor-driven changes in insulin delivery have now been approved by FDA in the United States (Medtronic 530G threshold glucose suspend<sup>7</sup> and 670G hybrid closed loop<sup>8</sup> devices) as well as outside of the United States (Medtronic 640G suspend before low glucose device).

## Accuracy

The discussion of CGM accuracy is complicated by the abundance of metrics that have been used to characterize it. Initial efforts to describe accuracy sought to adapt the clinical error grid long applied to SMBG<sup>9</sup> to CGM. This approach originally developed by Clarke to assess the clinical impact of SMBG errors was adjusted to assess CGMs by adding rate of change (ROC) accuracy data to point accuracy and re-named continuous glucose-error grid analysis (CG-EGA).<sup>10</sup> CG-EGA has been used to show differences between sensors.<sup>11</sup> However, CG-EGA is less frequently used today, as it is more complex both to calculate and to understand. Instead, various measures of point accuracy are typically reported, including the original clinical error grid.

CGMs require and benefit from both point and ROC accuracy. However, the benefits of ROC accuracy have not been optimally exploited by CGMs currently available or by

their users. Instead, trend arrows indicate ROC imprecisely. For example, a ROC between 61 and 119 mg/dL per minute is possible when the device displays a single arrow [i.e., indicating an ROC of between 1 and 2 mg/(dL·min)].

ROC arrows may not be well understood by all users. This was demonstrated in a questionnaire by Pettus and Edelman of 166 insulin pump and CGM users who were asked how they would adjust their correction insulin dose if their CGM showed two up arrows. Respondents increased their insulin doses on average by an additional 138% of the dose recommended by their bolus calculator, but responses varied from no additional insulin to a fourfold increase in the dose.<sup>12</sup>

MARD (mean absolute relative difference) has emerged as the preferred metric to assess sensor accuracy. Although it is easy to calculate<sup>13</sup> and is expressed as a single number, MARD is impacted by study design. Therefore, comparing MARDs of sensors from different studies may be misleading. Median absolute relative difference (unfortunately sharing the same abbreviation of MARD) is less-commonly presented and is numerically lower, as it excludes outlier values. Another manner of expressing accuracy is to state the proportion of values that lie within  $\pm 20$  mg/dL (for reference values  $\leq 80$  mg/dL) or within 20% (for reference values  $> 80$  mg/dL).<sup>14</sup> PARD (paired or precision absolute relative difference) compares readings generated by identical CGMs worn simultaneously by subjects and is a useful tool to characterize system performance that supplements MARD.<sup>15</sup>

Although most sensor studies have used a laboratory instrument (e.g., YSI-2300; Yellow Springs International Life Sciences) as the comparator, other studies have used or supplemented this with capillary glucose values.<sup>16–18</sup> Advantages of capillary data include the ability to obtain a greater number of data points, patients living in “real-world” conditions, and that capillary values are those that people with diabetes currently use most frequently to make therapy decisions. SMBG is also the procedure that people with diabetes would most like to replace. The obvious disadvantage of using SMBG is its lower accuracy versus laboratory comparators.

Abbott conducted an accuracy study<sup>17</sup> in patients with types 1 and 2 diabetes treated with injections or insulin pump therapy with a primary endpoint of point accuracy of the CGM as compared with the built-in SMBG meter of the device. With approximately eight data points per day per subject were available for this analysis, 86.7% of sensor results fell within Consensus Error Grid Zone A, and accuracy was maintained for the entire 14-day life of the sensor. Overall MARD was 11.4%. During in-clinic sessions, additional data were generated with more frequent reference (YSI) glucose values. The overall MARD in the clinic alone for sensor results with capillary BG reference and with YSI reference was 12.1% and 12%, respectively. The mean lag time between sensor and YSI reference values was calculated to be 4.5–4.8 min.

There is a widely held belief that an MARD value of 10% represents the safety threshold for the nonadjunctive use of CGM.<sup>19</sup> However, this speculation stems from a perceived flattening of the curve in an *in silico* simulation of hypoglycemia risk at an MARD of 10%. An alternative explanation for the putative change in curve inflection might be diminishing returns from MARD values lower than 10% rather than an objectively characterized threshold for acceptable hypoglycemia risk.

### Contributors to Accuracy

MARD has been shown to vary depending on glucose concentrations.<sup>20</sup> For this reason, Rodbard proposed graphing MARD as a continuous function of glucose concentration. ROC has also been shown to impact MARD.<sup>21,22</sup> Additional factors that impact the calculation of MARD include the absolute number of data points, their distribution, and missing data points.<sup>13</sup>

Calibration, for most devices, is required both initially and at subsequent intervals to compensate for sensor “drift.” Errors in calibration, because of either glucose meter inaccuracy or poor patient technique (i.e., failing to clean finger), directly impacts sensor accuracy. CGMs that do not require calibration potentially avoid these errors. Calibration protocols have been shown to influence MARD.<sup>23</sup> Calibration with a more accurate reference has been shown to reduce MARD.<sup>22</sup>

Software is used to calculate glucose values from measured sensor currents. Although intended to improve the quality of the data, early software sometimes led to artifacts and less accuracy.<sup>24</sup> However, improvements in software alone have also demonstrated significant improvement in MARD.<sup>25</sup>

Variability among manufacturer lots and between day of wear (i.e., first day performance is typically worse than subsequent days)<sup>26</sup> are other confounders when attempting to characterize sensor accuracy. Despite these pitfalls, it is generally appreciated that MARD-characterized accuracy in CGMs has steadily improved over the years.

### Comparative Accuracy

Although today’s CGMs are clearly more accurate than earlier CGMs,<sup>25</sup> reliably assessing the comparative accuracy among today’s CGMs presents a challenge. Few head-to-head studies exist in the literature<sup>11,27</sup> because of limited resources, including sparse on-body space to place comparator sensors.

The comparative assessment of CGM sensor accuracy has also proved complex because of the differing need for calibration (i.e., 0–3 times/day), the choice and reproducibility of accuracy metrics, the choice of reference (e.g., SMBG vs. laboratory instrument), and the choice of evaluation protocol. What was true in 2007 remains true today<sup>28</sup>—despite CLSI guidelines,<sup>29</sup> there is no universally accepted protocol to compare performance among sensors without simultaneous wear head-to-head trials.

### Clinical Effects of Accuracy

Observational data collected in 2014–2015 from the T1D Exchange clinic registry support the benefits of newer devices. In a recent analysis, HbA<sub>1c</sub> levels were significantly lower in patients using CGM than those not using CGM, regardless of whether patients administered insulin through a pump (HbA<sub>1c</sub> 7.7% vs. 8.2%;  $P < 0.001$ ) or through an MDI (7.8% vs. 8.6%,  $P < 0.001$ ).<sup>30</sup>

Traditional point-in-time measures of accuracy developed for SMBG, such as MARD, do not account for the ROC advantages inherent in CGM systems.<sup>31</sup> Understanding the clinical implications of sensor accuracy measurements would be difficult with traditional clinical trial designs. For example, randomizing some subjects to use a less accurate sensor to

detect differences in severe hypoglycemia would be considered unacceptable by most ethics boards. Therefore, in silico simulations have been utilized.

In silico simulations can randomize thousands of virtual patients to sensors of varying accuracy to explore the impact of device characteristics on insulin dosing and glycemia.<sup>25</sup> The recent approval of Dexcom's G5 CGM for nonadjunctive use was largely based on both its lower MARD and two in silico simulations.<sup>6</sup> The first 2-week simulation study utilized a validated physiological model to compare CGM-based and SMBG-based testing time in a range <50 mg/dL and >250 mg/dL. CGM was equivalent to or better than SMBG. CGM also reduced number and duration of hypoglycemic events, particularly in the setting of hypoglycemia unawareness. The second simulation study looked at meal dosing and hypoglycemia. Although setting an excessively low alert threshold, making inappropriate trend adjustments, and calibrating the device less than once a day worsened outcomes, CGM was still associated with lower hypoglycemic risk than SMBG.

### Recent Clinical Trials

The adjunctive use of CGM with type 1 diabetes is now well accepted, particularly in patients using insulin pumps. Two recent trials in patients not optimally controlled have supported its use in type 1 diabetes with multiple insulin injections. The Diamond trial<sup>32</sup> demonstrated a significantly greater (0.6%) decrease in HbA1c and time in hypoglycemia (43 vs. 80 min/day) with the use of CGM plus individualized insulin dosing recommendations. The 1-year duration GOLD trial<sup>33</sup> used a crossover design and demonstrated slightly lower (0.43%) HbA1c levels with CGM and significant education. This effect disappeared in subjects using CGM <70% of the time. In both studies, SMBG testing frequency was decreased but not eliminated with use of CGM (3.6 times/day vs. 4.6 times/day and 2.75 times/day vs. 3.66 times/day).

In practice, many patients already use their CGM data without confirmatory SMBG values for insulin dosing. Two recent clinical trials have been designed to study this nonadjunctive use. One trial (REPLACE) randomized subjects with type 2 diabetes with A1c values 7.5% to 12% on intensive insulin therapy 2:1 to either CGM (Abbott Libre) or SMBG for 6 months.<sup>34</sup> Patients with a history of severe hypoglycemia were excluded and no training to interpret CGM data was provided. Subjects in the SMBG group wore a blinded sensor for the last 2 weeks of the study. In the CGM group, SMBG was nearly eliminated (average of 0.3 checks/day) and Diabetes Treatment Satisfaction Questionnaire (DTSQ) scores improved significantly. HbA1c values were unchanged overall. Hypoglycemia (defined as sensor glucoses <70, <55, and <45 mg/dL) was reduced by 43%, 53%, and 64%, respectively. This might appear counterintuitive as the device has no alarms and subjects were not aware.

Another 6-month trial (IMPACT) was performed in well-controlled (HbA1c <7.5%) type 1 patients.<sup>35</sup> Subjects with hypoglycemia unawareness and CGM experience within 4 months were excluded. Subjects were randomized 1:1 to SMBG or CGM (Freestyle Libre). In this trial, SMBG testing in the CGM group declined to 0.5 tests/day, subjects scanned the CGM an average of 15.1 times/day, and HbA1c was unchanged. However, hypoglycemia (<70 mg/dL) in the CGM group declined by 38% from 3.38 to 2.03 hours/day.

These trials provide important reassurance as to the safety and efficacy of the nonadjunctive use of CGM. They also show a significant hypoglycemia benefit even without alarms. This is not to say that alarms are not valuable, but that their benefit is likely supplementary to that given by CGM data alone. The value of alarms to people with hypoglycemia unawareness has not been well characterized; in fact, such patients are often excluded from clinical trials. Until more data are available, it might be prudent to prescribe CGM with alarms enabled (and insulin pumps with automated insulin suspension) to people with hypoglycemia unawareness.

In summary, the clinical accuracy of CGM has reached a point where replacement of SMBG is feasible. Progress is being made to support less frequent calibration<sup>36</sup> and two devices have secured regulatory approval for nonadjunctive use. However, to completely eliminate SMBG, factory calibration will need to be achieved. Accuracy, combined with innovations in user interface, is enabling a better dashboard for people with diabetes.

### Author Disclosure Statement

**Research Support:** Abbott, Ambra, Ascensia, BD, Boehringer Ingelheim, Calibra, Companion Medical, Dexcom, Elcelyx, Glysens, Janssen, Lexicon, Lilly, Medtronic, Novo Nordisk, sanofi, Senseonics, Versartis, Xeris.

**Consulting Honoraria:** Astra Zeneca, Bayer, BD, Calibra, Lilly, Medtronic, Novo Nordisk, sanofi.

**Speaking Honoraria:** Abbott, Insulet, Medtronic, Lilly, Novo Nordisk, sanofi.

### References

1. Clarke SF, Foster JR: A history of blood glucose meters and their role in self-monitoring of diabetes mellitus. *Br J Biomed Sci* 2012;69:83–93.
2. Colwell JA, Dudley JD, McDonald JM, et al.: Consensus Statement on Self-Monitoring of Blood Glucose. *Diabetes Care* 1987;10:95–99.
3. The Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *N Engl J Med* 1993; 329:977–986.
4. JDRF CGM JDRF randomized clinical trial to assess the efficacy of real-time continuous glucose monitoring in the management of type 1 diabetes: research design and methods. *Diabetes Technol Ther* 2008;10:310–321.
5. Juvenile Diabetes Research Foundation Continuous Glucose Monitoring Study Group: The effect of continuous glucose monitoring in well-controlled type 1 diabetes. *Diabetes Care* 2009;32:1378–1383.
6. FDA advisory panel votes to recommend non-adjunctive use of Dexcom G5 mobile CGM. *Diabetes Technol Ther* 2016;18:512–516.
7. Bergenstal RM, Klonoff DC, Garg SK, et al.: Threshold-based insulin-pump interruption for reduction of hypoglycemia. *N Engl J Med* 2013;369:224–232.
8. Garg SK, Weinzimer SA, Tamborlane WV, et al.: Glucose outcomes with the in-home use of a hybrid closed-loop insulin delivery system in adolescents and adults with type 1 diabetes. *Diabetes Technol Ther* 2017;19:155–163.
9. Clarke WL, Cox D, Gonder-Frederick LA, et al.: Evaluating clinical accuracy of systems for self-monitoring of blood glucose. *Diabetes care* 1987;10:622–628.

10. Kovatchev BP, Gonder-Frederick LA, Cox DJ, Clarke WL: Evaluating the accuracy of continuous glucose-monitoring sensors: continuous glucose-error grid analysis illustrated by TheraSense Freestyle Navigator data. *Diabetes Care* 2004;27:1922–1928.
11. Kovatchev B, Anderson S, Heinemann L, Clarke W: Comparison of the numerical and clinical accuracy of four continuous glucose monitors. *Diabetes Care* 2008;31:1160–1164.
12. Pettus J, Edelman SV: Use of glucose rate of change arrows to adjust insulin therapy among individuals with type 1 diabetes who use continuous glucose monitoring. *Diabetes Technol Ther* 2016;18 Suppl 2:S234–S242.
13. Kirchsteiger H, Heinemann L, Freckmann G, et al.: Performance comparison of CGM systems: MARD values are not always a reliable indicator of CGM system accuracy. *J Diabetes Sci Technol* 2015;9:1030–1040.
14. Christiansen M, Bailey T, Watkins E, et al.: A new-generation continuous glucose monitoring system: improved accuracy and reliability compared with a previous-generation system. *Diabetes Technol Ther* 2013;15:881–888.
15. Obermaier K, Schmelzeisen-Redeker G, Schoemaker M, et al.: Performance evaluations of continuous glucose monitoring systems: precision absolute relative deviation is part of the assessment. *J Diabetes Sci Technol* 2013;7:824–832.
16. Bailey TS, Chang A, Christiansen M: Clinical accuracy of a continuous glucose monitoring system with an advanced algorithm. *J Diabetes Sci Technol* 2015;9:209–214.
17. Bailey T, Bode BW, Christiansen MP, et al.: The performance and usability of a factory-calibrated flash glucose monitoring system. *Diabetes Technol Ther* 2015;17:787–794.
18. Kropff J, Bruttomesso D, Doll W, et al.: Accuracy of two continuous glucose monitoring systems: a head-to-head comparison under clinical research centre and daily life conditions. *Diabetes Obes Metab* 2015;17:343–349.
19. Kovatchev BP, Patek SD, Ortiz EA, Breton MD: Assessing sensor accuracy for non-adjunct use of continuous glucose monitoring. *Diabetes Technol Ther* 2015;17:177–186.
20. Rodbard D: Characterizing accuracy and precision of glucose sensors and meters. *J Diabetes Sci Technol* 2014;8:980–985.
21. Pleus S, Schoemaker M, Morgenstern K, et al.: Rate-of-change dependence of the performance of two CGM systems during induced glucose swings. *J Diabetes Sci Technol* 2015;9:801–807.
22. Kamath A, Mahalingam A, Brauker J: Analysis of time lags and other sources of error of the DexCom SEVEN continuous glucose monitor. *Diabetes Technol Ther* 2009;11:689–695.
23. King C, Anderson SM, Breton M, et al.: Modeling of calibration effectiveness and blood-to-interstitial glucose dynamics as potential confounders of the accuracy of continuous glucose sensors during hyperinsulinemic clamp. *J Diabetes Sci Technol* 2007;1:317–322.
24. Metzger M, Leibowitz G, Wainstein J, et al.: Reproducibility of glucose measurements using the glucose sensor. *Diabetes Care* 2002;25:1185–1191.
25. Facchinetti A: Continuous glucose monitoring sensors: past, present and future algorithmic challenges. *Sensors (Basel)* 2016;16:2093.
26. Castle JR, Ward WK: Amperometric glucose sensors: sources of error and potential benefit of redundancy. *J Diabetes Sci Technol* 2010;4:221–225.
27. Damiano ER, McKeon K, El-Khatib FH, et al.: A comparative effectiveness analysis of three continuous glucose monitors: the Navigator, G4 Platinum, and Enlite. *J Diabetes Sci Technol* 2014;8:699–708.
28. Nichols JH, Klonoff DC: The need for performance standards for continuous glucose monitors. *J Diabetes Sci Technol* 2007;1:92–94.
29. D'Archangelo MJ: New guideline supports the development and evaluation of continuous interstitial glucose monitoring devices. *J Diabetes Sci Technol* 2008;2:332–334.
30. Miller K, Foster N, Tamborlane W, et al.: Continuous glucose monitoring in T1D patients using injections of insulin: a report from the T1D Exchange clinic registry. *Diabetes Technol Ther* 2016;18(S1):A-27. Abstract 069.
31. Kollman C, Wilson DM, Wysocki T, et al.: Limitations of statistical measures of error in assessing the accuracy of continuous glucose sensors. *Diabetes Technol Ther* 2005;7:665–672; discussion 673–664.
32. Beck RW, Riddlesworth T, Ruedy K, et al.: Effect of continuous glucose monitoring on glycemic control in adults with type 1 diabetes using insulin injections: the diamond randomized clinical trial. *JAMA* 2017;317:371–378.
33. Lind M, Polonsky W, Hirsch IB, et al.: Continuous glucose monitoring vs conventional therapy for glycemic control in adults with type 1 diabetes treated with multiple daily insulin injections: the gold randomized clinical trial. *JAMA* 2017;317:379–387.
34. Haak T, Hanaire H, Ajjan R, et al.: Flash glucose-sensing technology as a replacement for blood glucose monitoring for the management of insulin-treated type 2 diabetes: a multicenter, open-label randomized controlled trial. *Diabetes Ther* 2017;8:55–73.
35. Bolinder J, Antuna R, Geelhoed-Duijvestijn P, et al.: Novel glucose-sensing technology and hypoglycaemia in type 1 diabetes: a multicentre, non-masked, randomised controlled trial. *Lancet* 2016;388:2254–2263.
36. Acciaroli G, Vettoretti M, Facchinetti A, et al.: From two to one per day calibration of Dexcom G4 platinum by a time-varying day-specific Bayesian prior. *Diabetes Technol Ther* 2016;18:472–479.

Address correspondence to:  
*Timothy S. Bailey, MD, FACE, CPI*  
*AMCR Institute*  
*625 W. Citracado Parkway Suite 112*  
*Escondido, CA 92025*

*E-mail: tbailey@amcrinstitute.com*