# Measures of Accuracy for Continuous Glucose Monitoring and Blood Glucose Monitoring Devices

**Guido Freckmann, MD[1], Stefan Pleus, MSc[1]** (ID),
**Mike Grady, PhD[2], Steven Setford, PhD[2], and Brian Levy, MD[3]**

## Abstract

Currently, patients with diabetes may choose between two major types of system for glucose measurement: blood glucose monitoring (BGM) systems measuring glucose within capillary blood and continuous glucose monitoring (CGM) systems measuring glucose within interstitial fluid. Although BGM and CGM systems offer different functionality, both types of system are intended to help users achieve improved glucose control. Another area in which BGM and CGM systems differ is measurement accuracy. In the literature, BGM system accuracy is assessed mainly according to ISO 15197:2013 accuracy requirements, whereas CGM accuracy has hitherto mainly been assessed by MARD, although often results from additional analyses such as bias analysis or error grid analysis are provided. The intention of this review is to provide a comparison of different approaches used to determine the accuracy of BGM and CGM systems and factors that should be considered when using these different measures of accuracy to make comparisons between the analytical performance (ie, accuracy) of BGM and CGM systems. In addition, real-world implications of accuracy and its relevance are discussed.

Many people with diabetes are required to regularly check their current glucose concentrations. Depending on their diabetes regimen, the glucose results can then be used to make therapeutic decisions, such as insulin dosing.[1]

Currently, patients with diabetes may choose between two major types of system for glucose measurement: blood glucose monitoring (BGM) systems measuring glucose within capillary blood and continuous glucose monitoring (CGM) systems measuring glucose within interstitial fluid (ISF). While BGM is the standard and more established approach for informing diabetes therapy, CGM has seen a rapid increase in users in recent years. Most CGM systems currently in use can be categorized as either real-time CGM (rtCGM) or as intermittently scanned CGM (iscCGM; also referred to as "flash glucose monitoring"). Whereas rtCGM systems typically provide new glucose values every 5 minutes, provided that the device used for display is in range of the wireless transmission, iscCGM systems require the user to scan the sensor to obtain current glucose values, either on a specific reader or app-enabled smartphone. For many years, CGM systems were only intended for adjunctive use, that is, to supplement BGM information when making diabetes management decisions. However, there are now CGM systems manufactured by Dexcom and Abbott that are intended for nonadjunctive use, that is, to replace BGM in

many therapeutic situations particularly in terms of using the information from the CGM system for insulin dose adjustments.[2,3]

Although BGM and CGM systems offer different functionality, both types of system are intended to help users achieve improved glucose control. BGM, being an episodic measurement process initiated by the end-user, is recommended to be performed at least 3 or 4 times per day by intensive insulin-using people with type 2 diabetes[4] and actually performed on average 5 to 6 times per day by people with type 1 diabetes.[5] Guidance recommends people with type 1 diabetes perform self-monitoring of blood glucose (SMBG) between 4 and 10[6] or even 6 and 10[1] times per day. In contrast, CGM can provide a more comprehensive picture of glycemia, since these systems gather data on a more continuous basis (because of the higher measurement

[1]Institut für Diabetes-Technologie, Forschungs- und Entwicklungsgesellschaft mbH an der Universität Ulm, Ulm, Germany
[2]LifeScan Scotland Ltd, Inverness, Scotland, UK
[3]LifeScan Inc, Wayne, PA, USA

**Corresponding Author:**
Stefan Pleus, MSc, 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

**Table 1.** Measurement Accuracy Criteria of ISO 15197:2013 and FDA Requirements for Integrated CGM (iCGM) Systems.

| Glucose concentrations | ISO 15197:2013 | FDA's iCGM requirements[b] |
|---|---|---|
| Overall | ≥ 95% within ±15 mg/dl or ±15%[a] | >87% within ±20% |
| <70 mg/dl | NA | >85% within ±15 mg/dl |
| | | >98% within ±40 mg/dl |
| | | no value >180 mg/dl |
| 70-180 mg/dl | NA | >70% within ±15% |
| | | >99% within ±40% |
| >180 mg/dl | NA | >80% within ±15% |
| | | >99% within ±40% |
| | | no value <70 mg/dl |
| Additional requirements | ≥ 99% of within consensus error grid zones A and B | ≤ 1% of glucose rates of change >1 mg/(dl*min) if true rate of change <−2 mg/(dl*min) |
| | | ≤ 1% of glucose rates of change <−1 mg/(dl*min) if true rate of change >2 mg/(dl*min) |

[a]At least 95 % of measured glucose values should be within either ±15 mg/dl of the averaged comparison values at glucose concentrations <100 mg/dl or within ±15% at glucose concentrations ≥100 mg/dl.
[b]The percentage of measured glucose values within a defined difference value or within a defined percentage value is calculated and the lower one-sided bound of the 95% confidence interval of this percentage must be in excess of the given requirement value. For studies where relatively low numbers of paired data-points are collected, a correspondingly wider confidence interval will apply. In such cases, a higher percentage of values must meet the accuracy requirement in comparison to a study where a larger dataset is collected.

frequency) and store the ISF glucose values several times per hour in the reader, receiver or associated app on a connected device. In addition to current glucose values, CGM systems can provide retrospective glucose trend graphs and an estimation of the current rate of glucose change (eg, rising or falling in mg/dl per min increments), derived from previously recorded glucose data. Currently, additional measurement functionality is available in some episodic (BGM) systems, for example the ability to measure analytes such as ketones, by simply using a different reagent strip in the same episodic meter.

Another issue in which BGM and CGM systems differ, is measurement accuracy. For BGM systems, guidance documents[7,8] or international standards[9] provide recommendations and stipulations on how to assess measurement performance as well as defining minimum accuracy requirements. For CGM systems, even those intended for nonadjunctive use (eg, insulin dosing decisions), minimum accuracy requirements have, until very recently, not been defined, and recommendations on CGM accuracy assessments, although in the process of being updated, are around 10 years old.[10] The US Food and Drug Administration (FDA) has recently outlined a new 510K (premarket approval) route for some CGM systems, designated as "integrated CGM" (iCGM) with additional special controls governing accuracy (see Table 1). Integrated CGM systems are required to transmit glucose measurement data to digitally connected devices, although in practice they may be used without such devices. These new recommendations are intended to improve awareness and transparency regarding CGM accuracy requirements and help bridge a knowledge gap in terms of how the scientific community compare and contrast BGM and CGM data.

Simulation analyses were used to assess clinical relevance of CGM accuracy in the past,[11] and they may help in establishing suitable accuracy criteria. It might, for example, be argued that the risk associated with allowing up to 2% of CGM results with "true" glucose concentrations of <70 mg/dl, but exhibiting deviations of >±40 mg/dl (Table 1) may be unacceptable.

In addition, the recommended methods for accuracy assessments and performance parameters are inconsistent between BGM and CGM systems.

When considering BGM systems, accuracy is mainly influenced by differences between the integrated measurement method and the comparison method against which they are calibrated (typically, both methods measure within blood, ideally capillary blood). CGM systems, however, are also influenced by the time delay between glucose changes in the interstitial fluid compartment and the compartment (ie, blood) in which comparative measurements are obtained. The intention of this review is to provide a comparison of the different approaches used to determine the accuracy of BGM and CGM systems and factors that should be considered when using these different measures of accuracy to make comparisons between the analytical performance (ie, accuracy) of BGM and CGM systems. In addition, real-world implications of accuracy and its relevance are discussed.

## Assessing Measurement Performance

The following section is limited to those assessments most commonly reported in the literature as they pertain to BGM and CGM accuracy performance. It should be acknowledged that there are many more ways to assess measurement performance, such as linear regression analysis or calculation of

correlation coefficients. However, these assessments, while valid, are less often reported than those detailed here.

All the following assessments rely on comparing individual BGM/CGM measurement values with the corresponding values of the comparison (reference) method. Often, laboratory analyzers are used as the comparison method, although, depending on limitations that may be posed by the particular study setting, BGM systems may serve as the comparison method, particularly in studies that involve an element of "home-testing" by the subject.

### Accuracy Limits of ISO 15197:2013

The International Organization for Standardization (ISO) published standard ISO 15197:2013, describing requirements for BGM systems for self-testing in managing diabetes mellitus and provides extensive guidance on how to assess measurement accuracy, along with other guidelines for design verification and performance validation. ISO 15197:2013 was harmonized with the regulations of the European Union as EN ISO 15197:2015. This harmonization had no impact on the requirements and procedures in ISO 15197:2013; changes were made to the foreword and an informative annex added.

ISO 15197 states that the difference between measurement results obtained with the BGM system and with the comparison method from the same sample must fall within defined limits for a certain percentage of samples. According to ISO 15197:2013, and as described in Table 1, at least 95% of results for each of three different reagent system lots shall fall within ±15 mg/dl of the comparison method result at glucose concentrations <100 mg/dl (ie, based on the difference between the paired values) and within ±15% at glucose concentrations ≥100 mg/dl (ie, based on the relative difference between the paired values).[9] In addition, at least 99% of pooled results shall fall within zones A and B of the consensus error grid.[9]

In some publications, additional accuracy criteria are reported with regard to these difference and/or relative difference values between the test system and comparison method results. Examples include more stringent accuracy criteria than those defined within ISO 15197:2013 (eg, ±5 mg/dl and ±5%, or ±10 mg/dl and ±10%, or more lenient accuracy criteria of ±20 mg/dl / ±20% or ±30 mg/dl / ±30%, all calculated against the same cut-off glucose concentration).[9,12-14]

### Mean Absolute Relative Difference (MARD)

MARD analysis is widely used to describe accuracy of CGM systems,[13-15] although it is also sometimes applied to BGM systems,[16-18] particularly where it is desired to make comparisons between the two types of system. MARD is calculated by averaging the absolute values of relative differences between CGM/BGM system measurement results and corresponding comparison method results. In this case, "absolute" means

each individual relative difference value is considered a positive value, irrespective of whether the calculated difference with respect to the comparison result is positive or negative. Sometimes, the mean absolute difference (MAD) may be calculated below specific glucose concentration thresholds.[14,19] In this case, absolute difference as opposed to absolute relative difference values are used to calculate the mean.

MARD results are sometimes reported as "aggregated MARD," that is, the average of all individual pairs between CGM/BGM system measurement results and corresponding comparison method results within a given study. Aggregated MARD data may be reported alongside, especially in the case of CGM systems, the MARD values for individual sensors or individual subjects. Aggregated MARD can make the whole spectrum of possible individual differences across multiple sensors more visible, whereas the sensor- or subject-specific MARD better reflects the expected variability of sensor performance in regular use.[14]

Furthermore, CGM MARD results are sometimes stratified by glucose concentration range[10] (akin to how BGM data is analyzed for system accuracy) or, uniquely, by sensor glucose rate of change, by virtue of the essentially continuous measurement of glucose values.[20]

### Systematic Measurement Difference (Bias)

Another common approach to describing the accuracy of a measurement system is to calculate bias. Bias may be defined as the systematic difference between measurement results from the system under investigation and the comparison method. A laboratory analyzer such as a YSI (Yellow Springs Instrument) glucose analyzer is often used as the comparison method for BGM systems while for CGMs a laboratory analyzer, or in many cases a BGM (particularly in studies performed in a nonclinical setting), is used as the comparison method to substantiate accuracy. Bias differs from MARD in that the bias calculation incorporates the directionality of the difference, whether positive or negative compared to the value of the comparison method.

In a seminal publication, Bland and Altman, in response to limitations posed by the use of correlation methodologies in comparing two quantitative measurement systems, suggested plotting individual differences between results of the of the investigated systems and the comparison method against their mean value.[21] Other publications advocate comparison of paired-result differences against the comparison result alone. Therefore, care should be taken in interpreting those publications stating the "relative bias" according to Bland and Altman, where the differences are divided by the mean of the paired results as opposed to others that consider differences versus the comparison result alone. In addition, Bland and Altman suggested calculation of "limits of agreement," defined as mean ± 2 × standard deviation (or, for normally distributed data, mean ± 1.96 × standard deviation), within which 95% of differences are expected to lie for

different systems to be considered similar in terms of quantitative performance.

### Error Grid Analysis Methods

Error grid analysis differs from the purely numerical ISO and paired-value difference approaches described above, in that the clinical risk associated with the distribution of results between the investigated system and the comparison method is assessed. Pairs of glucose measurement results are categorized by risk scores (eg, surveillance error grid [EG][22]) or risk zones (eg, consensus EG[9] and continuous glucose EG[23]).

ISO 15197:2013, for example, stipulates use of the consensus EG which provides 5 different risk zones from "no effect on clinical action" (zone A) to "altered clinical action that could have dangerous consequences" (zone E).

A major difference between consensus EG and surveillance EG analysis is that in the more recently developed surveillance EG, risk scores are assigned to pairs of measurement results on a continuous scale, thus eliminating sudden changes in clinical risk at borders between risk zones.

The continuous glucose EG was specifically designed to be used for comparisons of CGM values and BGM values obtained with sufficiently high frequency (one value every 10 to 15 minutes).[23] Within such short time frames, consecutive CGM results cannot be viewed as statistically independent samples, so that additional analyses, for example, regarding rates of change in glucose concentration are considered. The continuous glucose EG assesses rate-of-change accuracy, that is, how well changes between subsequent CGM results match changes between paired subsequent comparison results. This is shown in a separate grid from point accuracy which compares how well individual CGM results match paired comparison results. In addition, the zones boundaries of the point accuracy grid are shifted to account for the time delay between blood and interstitial glucose concentration changes.[23]

## Comparing Accuracy of BGM Systems and CGM Systems

In the literature, BGM system accuracy is assessed mainly according to ISO 15197:2013 accuracy requirements, whereas CGM accuracy has hitherto mainly been assessed by MARD. Therefore, comparing accuracy of BGM and CGM systems is difficult since accuracy, as defined by ISO 15197:2013 requirements, cannot be directly compared or translated into MARD values. Therefore, estimations of ISO percentages based on MARD values, or vice versa, have to rely on mathematical models about the distribution of measurement results and the likelihood of obtaining specific sets of measurement results (probabilistic models).[24,25]

Even when the same assessment approach is used (eg, MARD), results from studies of different types of system may not be comparable. For example, in a CGM study, the number of results obtained from the same subject wearing the same CGM system is much larger (up to 288 sensor values per day) than in BGM studies. Therefore, the datasets that are obtained for each system cannot be easily reconciled for a meaningful comparison to be made. In addition, MARD may depend on the study setting (eg, number of results and distribution of glucose concentrations),[26,27] so that for some parameters, results from different studies of the same type of system may not be comparable. Although difficult to directly compare, for guidance purposes, the typical overall MARD for commercially available CGMs labeled for nonadjunctive use ranges from 10 to 12%,[13-15] whereas commercially available high-quality BGMs can achieve MARD results below 5%.[17,18]

Furthermore, certain parameters, may not allow differentiation between imprecision and bias. In the case of MARD, for example, a low (ie, good) MARD indicates that both imprecision and bias are small (ie, good), whereas a high (ie, poor) MARD does not provide information as to whether results are imprecise or biased or indeed both (as is apparent on examining plots B, C, and D of Figure 1, which describes four scenarios of modelled data based on BGM data). In Figure 1B, the MARD result of 10% is primarily driven by high bias, whereas in Figure 1C, bias is considerably lower and imprecision is substantially higher. Thus, potentially important information about the location and dispersion of data points is lost. In a real-life situation, if a MARD result of 10% were caused only by bias, with perfect precision (ie, all glucose monitoring system results are shown either 10% higher or 10% lower than the comparison method result), users may adapt their diabetes management (or calculate the "true" glucose concentration), once the bias is known. If the same MARD result of 10% were caused by imprecision with zero mean bias (ie, individual results vary considerably but the average glucose concentration is the same as the average comparison method result), users would not know whether their result happens to be higher or lower than the true glucose concentrations.

As with MARD, application of accuracy limits like those of ISO 15197:2013 does not allow differentiation between imprecision and bias (see ISO 15197:2013 percentage values in Figure 1, plots A, B, and D). However, when applying the accuracy limit approach, a more quantitative assessment of data distribution can be obtained by reporting percentages of results within a range of different accuracy limits (eg, ±5%, ±10% or other). It may therefore be concluded that reducing measurement accuracy to a single parameter (like MARD or accuracy limits of ISO 15197:2013) is inappropriate, especially when looking to compare results between different studies.

Bias analysis, on the other hand, can provide high-level estimators of both the location and dispersion of data points. The suitability of bias analysis depends on the data, and it may not necessarily adequately reflect how individual data points are located or spread, as may be appreciated with reference to Figure 2. Typically, location and dispersion
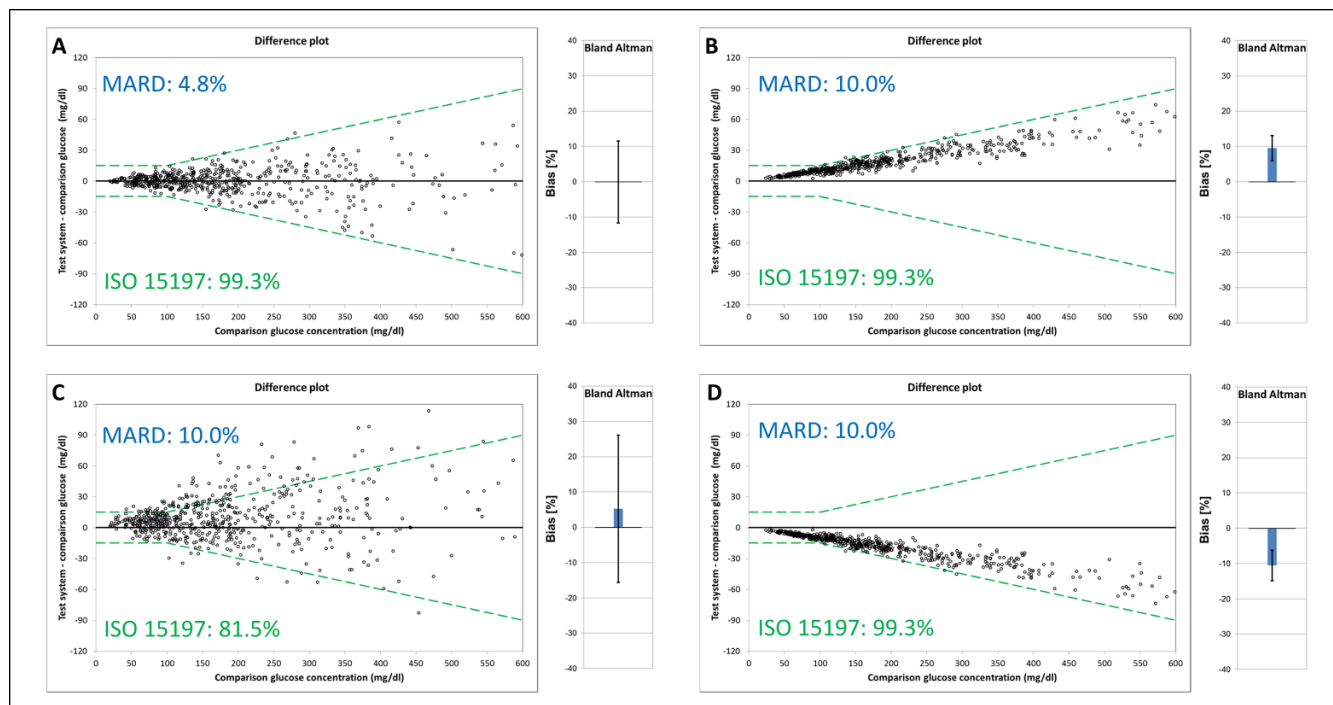
**Figure 1.** Modeled data (n = 600) for a glucose monitoring system with (A) MARD 4.8%, 99.3% of results within accuracy limits of ISO 15197, and no bias, medium precision; (B) MARD 10.0%, 99.3% of results within accuracy limits of ISO 15197, and medium bias, high precision; (C) 10.0% MARD, 81.5% of results within accuracy limits of ISO 15197, and positive bias, low precision; (D) MARD 10.0%, 99.3% of results within accuracy limits of ISO 15197, and negative bias, high precision.

estimators are provided for either differences (expressed in the same unit as the measurement results) or relative differences (expressed in percent), and it is not standard practice to combine the two in bias analysis.

EG analysis differs from the other assessments described, since it focuses on the clinical relevance of differences between CGM and BGM results and the comparison method. While only the continuous glucose EG acknowledges the importance of rate-of-change accuracy, its point accuracy grid is based upon the Clarke EG,[28] with the difference that the continuous glucose EG risk zone boundaries are slightly adjusted based on rate of change to account for time-based concentration differences arising from glucose migration from blood to ISF. The Clarke EG may seem outdated in light of the more modern consensus and surveillance EGs.

Care should be taken regarding prerequisites for adequate accuracy assessments. Consideration should be made as to manufacturer's labeling, such as with respect to interfering substances or intended use. In addition, the instrumentation used, and methodology applied for obtaining comparative results, plays a major role in defining device performance and therefore should be shown to be of sufficiently high quality, because in such comparative assessments, inaccuracies of the comparison and test systems cannot be differentiated. This high quality does not only concern possible bias between methods, such as between laboratory analyzers (Twomey and colleagues reported 8%

bias between glucose oxidase and hexokinase methods[29]) and possible bias within the same type of analyzer (as observed by Bailey and colleagues when using YSI 2300 STAT Plus analyzers at different study sites[30]), but also imprecision of the comparison method itself. Apart from analytical performance of the comparison method, pre- and postanalytical errors should be minimized. While laboratory analyzers may provide higher analytical quality than most BGM systems under optimal conditions, they might be affected by preanalytical errors that can more easily be avoided with BGM systems, for example, inadequate sample handling leading to hemolysis or glycolysis. High-quality BGM systems might therefore provide comparable accuracy depending on the specific conditions.

If a manually calibrated CGM system is used, additional influencing factors should be considered, especially compartmental differences and possible measurement bias between the instruments used for calibration and comparison measurements. If BGM or CGM systems are calibrated with capillary blood samples, using samples other than capillary blood for comparison measurements may introduce additional bias.[31-33] Because laboratory analyzers typically require larger blood volumes than BGM systems, obtaining capillary blood samples with sufficient volume may be an issue. If a measurement bias exists between instruments used for calibration and for comparison measurements, this bias will affect all comparison measurements.
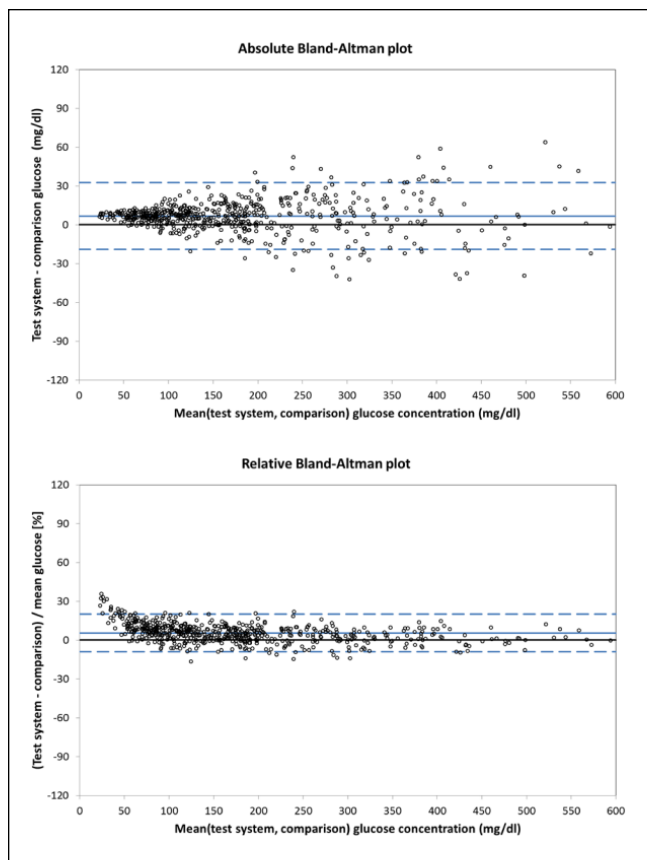
**Figure 2.** Modeled data (n = 600) for a glucose monitoring system showing constant bias and glucose-concentration-dependent variability.

Any issues affecting accuracy of CGM systems and assessments of accuracy, including those mentioned above, should be identified as part of on-going efforts toward standardizing accuracy assessments of CGM systems.

## Real-World Implications

Beyond directly comparing accuracy of different systems (BGM vs CGM; BGM vs BGM; CGM vs CGM), the parameters described above also apply to comparisons between different lots (or batches) of a BGM or CGM system, which are routinely released into the retail market. In terms of providing evidence that end users can have confidence in the product when transitioning from one lot to another in the field, ISO 15197:2013, for example, explicitly states that three different strip lots must be used in system accuracy.

For CGM systems, however, public reports describing the individual (or collective) performance of different sensor lots released to the market are absent. Furthermore, there is no equivalent standard to ISO 15197:2013 mandating that multiple different sensor lots are tested as part of the pre-market assessment. Even sensor-to-sensor comparisons within the same subject[34] are not consistently reported, and

information about the sensor lots used is often absent. While there are reports indicating that the application site of the CGM system, such as abdomen versus arm, may influence accuracy,[35] differences have also been reported for CGM systems applied in close proximity to each other.[36] When considering the typically limited number of subjects within an assessment, coupled to sensor-to-sensor variability and the effect of other factors such as intersubject variability or application-site differences, it is not surprising that it is often difficult to draw definitive conclusions regarding CGM device performance.

CGM systems provide more context than BGM systems in terms of glycemic control. As already mentioned, they provide the current ISF glucose value, an indication of the glucose trend at that time (eg, going up or down) and importantly an estimate of the rate of glucose change in terms of the trend arrow. Some rtCGM systems (but not iscCGM systems such as FreeStyle Libre) provide low or high alerts as soon as certain glucose concentration thresholds (eg, hypoglycemia and hyperglycemia) have been crossed or if they will be crossed soon.[11] In iscCGM systems such as FreeStyle Libre, alerts are shown only upon scanning, so that hypo- or hyperglycemic events may be missed if the user does not happen to scan at the appropriate time. The value of this additional insight and information for the patient is directly linked to the measurement accuracy of the CGM system, because this information is derived from previously recorded glucose data, and therefore such functionality can sometimes be flawed.[37-39] However, used appropriately, it can empower patients to make more informed therapeutic decisions.[40]

Clearly, CGM system point accuracy is not yet at the same level that is required for BGM systems,[24] so the question arises as to what level of point accuracy is required under which conditions, such as rates of change for CGM systems labeled for nonadjunctive use, to allow appropriate diabetes therapy decisions to be made.[41] Despite multiple circumstances and conditions under which guidance or warning instructions in the labeling of CGMs instruct users to verify readings using a BGM system, it is clear from clinical trials that for the most part the majority of users do not verify the accuracy of applicable sensor readings.[42-45] To their credit, some sensor companies have introduced helpful on-screen BG checking alerts that automatically appear to advise the user to verify their sensor value based on the prevailing conditions, and it will be interesting to see future data that demonstrate compliance to such advisory safety features.

It has been argued that CGM systems may be allowed to be less accurate than BGM systems, because adequate diabetes therapy is still possible as long as this additional information (such as glucose trend arrows) is provided.[41] However, additional information might also be harmful if it is inappropriately applied to therapeutic decisions. Early postprandial hyperglycemia is common in people with diabetes, because currently available subcutaneously injected insulin analogues do not act sufficiently rapidly to effectively blunt

increases in glucose concentration.[46] If patients were delivering corrective insulin doses too early after meals, that is, based on these expected hyperglycemic results, insulin-induced hypoglycemia would likely follow. Thus, patients adopting CGM systems in their diabetes regimen should receive specific education and training if they are supposed to effectively use the additional information provided by CGM systems.

It may be insufficient to apply BGM system accuracy requirements to CGM systems not only because they provide additional helpful information, but also because of the time delay of glucose changes between different compartments. This time delay may result in differences between venous, interstitial and capillary glucose concentrations, particularly at times of rapidly changing glucose.[31] This may possibly lead to undesirable physiological glucose states, for example, if recognition of a hypoglycemic state is delayed or if treatment of a hypoglycemic state is not recognized in a timely manner, consequently potentially leading to overcorrective actions. The physiologic time delay between the interstitial and capillary compartments differs between individuals, but has been estimated to be around 6 minutes, with a conservative estimation of approximately 10 minutes.[47] In addition to this physiologic time delay, a physical (or technical) time delay is introduced via the sensor, for example by the time taken for glucose to diffuse through a biocompatible membrane and the signal de-noising algorithms.[48,49] It is expected that shorter time delays reported for CGM systems are most likely achieved through the use of algorithms aimed at compensating for this time delay.[50] Regarding nonadjunctive use, compensation of this time delay may be helpful: if BGM is intended to be replaced by CGM, diabetes therapy parameters may otherwise have to be adjusted to compensate for any differences between the capillary and interstitial compartments.

Due to these physiological differences between glucose concentrations in different compartments, using ISF samples for comparative assessments would likely lead to the perception of better performance results than using venous of capillary samples. Currently, however, there is no established reference method for ISF glucose concentrations. It might also be argued that, at least for those CGM systems that are intended to replace BGM systems, that it is reasonable to provide capillary-like glucose values.

Although there are detailed requirements regarding accuracy of BGM systems,[9] similar requirements were missing for CGM systems until the recent introduction of the FDA's iCGM requirements. It seems inconsistent of regulatory authorities and expert groups to require minimum BGM system accuracy and, at the same time, not require minimum CGM system accuracy, especially regarding nonadjunctive use of CGM systems, independent of whether the system is an "integrated" CGM system or not. The impact of nonadjunctive use on long-term diabetes complications is yet unknown.

## Conclusion

For a variety of reasons, the accuracy of BGM and CGM systems have not been easy to compare. Despite measuring glucose using similar enzyme-based reagents in the episodic strip or CGM sensor, the two types of systems measure from different compartments (blood vs ISF) and each is exposed to differing glucose concentrations that require specific algorithmic compensations and/or real-time calibrations to improve accuracy. Metrics to describe accuracy (eg, ISO or MAD, MARD or precision absolute relative difference [PARD]) evolved to account for differences in the scale of the BGM and CGM datasets (and comparison values recorded) making comparisons more challenging.

Given their typically superior point accuracy, BGM devices are routinely used for insulin dosing decisions and in recent years improvements in CGM accuracy have enabled certain CGM systems to label for nonadjunctive use when informed by additional CGM functionality (eg, CGM sensor reading, rate and trend of glucose).

Despite labeling for nonadjunctive use, recommendations for minimum CGM accuracy performance are still at the formative stage. In contrast, acceptance criteria and guidelines (eg, ISO/FDA) for BGM systems have evolved with a long history and are now internationally accepted.

Encouragingly with the advent of a new FDA CGM categorization (integrated CGM, abbreviated to iCGM) and the special controls on performance inherent in this new class, CGM systems will hopefully in the future be held to a minimum standard regarding accuracy around the world.

### Abbreviations

BGM, blood glucose monitoring; CGM, continuous glucose monitoring; EG, error grid; FDA, Food and Drug Administration; iCGM, integrated CGM; iscCGM, intermittently scanned CGM; ISF, interstitial fluid; ISO, International Organization for Standardization; MAD, mean absolute difference; MARD, mean absolute relative difference; PARD, precision absolute relative difference; rtCGM, real-time CGM; SMBG, self-monitoring of blood glucose; YSI, Yellow Springs Instrument.

### Declaration of Conflicting Interests

## ORCID iD

Stefan Pleus [iD] https://orcid.org/0000-0003-4629-7754

## References

1. American Diabetes Association. Standards of medical care in diabetes—2018. *Diabetes Care*. 2018;41(suppl 1):s1-s156.
2. Food and Drug Administration. PMA P120005/S041: FDA summary of safety and effectiveness data. https://www.access-data.fda.gov/cdrh_docs/pdf12/P120005S041b.pdf. Accessed March 27, 2017.
3. Food and Drug Administration. Premarket approval (PMA): FreeStyle Libre flash glucose monitoring system (P160030). https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P160030. Accessed November 13, 2017.
4. Hirsch IB, Bode BW, Childs BP, et al. Self-monitoring of blood glucose (SMBG) in insulin- and non-insulin-using adults with diabetes: consensus recommendations for improving SMBG accuracy, utilization, and research. *Diabetes Technol Ther*. 2008;10:419-439.
5. Miller KM, Beck RW, Bergenstal RM, et al. Evidence of a strong association between frequency of self-monitoring of blood glucose and hemoglobin A1C levels in T1D exchange clinic registry participants. *Diabetes Care*. 2013;36:2009-2014.
6. National Institute for Health and Care Excellence. *Type 1 diabetes in adults: diagnosis and management*. NICE guideline NG17. 2015.
7. Food and Drug Administration. Blood glucose monitoring test systems for prescription point-of-care use—guidance for industry and food and drug administration staff. http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM380325.pdf. Accessed October 10, 2016.
8. Food and Drug Administration. Self-monitoring blood glucose test systems for over-the-counter use—guidance for industry and food and drug administration staff. http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM380327.pdf. Accessed October 10, 2016.
9. International Organization for Standardization. *In vitro diagnostic test systems—Requirements for blood-glucose monitoring systems for self-testing in managing diabetes mellitus*. ISO 15197:2013. 2013.
10. Clinical and Laboratory Standards Institute. *POCT05-A—Performance metrics for continuous interstitial glucose monitoring*. 2008.
11. 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.
12. Laffel L. Improved accuracy of continuous glucose monitoring systems in pediatric patients with diabetes mellitus: results from two studies. *Diabetes Technol Ther*. 2016;18(suppl 2):s223-s233.
13. Freckmann G, Link M, Pleus S, Westhoff A, Kamecke U, Haug C. Measurement performance of two continuous tissue glucose monitoring systems intended for replacement of blood glucose monitoring. *Diabetes Technol Ther*. 2018;20:541-549.
14. Bailey T, Bode BW, Christiansen MP, Klaff LJ, Alva S. The performance and usability of a factory-calibrated flash glucose monitoring system. *Diabetes Technol Ther*. 2015;17:787-794.
15. Wadwa RP, Laffel LM, Shah VN, Garg SK. Accuracy of a factory-calibrated, real-time continuous glucose monitoring system during 10 days of use in youth and adults with diabetes. *Diabetes Technol Ther*. 2018;20:395-402.
16. Setford S, Grady M, Phillips S, et al. Seven-year surveillance of the clinical performance of a blood glucose test strip product. *J Diabetes Sci Technol*. 2017;11:1155-1162.
17. Freckmann G, Pleus S, Link M, et al. Accuracy evaluation of four blood glucose monitoring systems in unaltered blood samples in the low glycemic range and blood samples in the concentration range defined by ISO 15197. *Diabetes Technol Ther*. 2015;17:625-634.
18. Bedini JL, Wallace JF, Pardo S, Petruschke T. Performance evaluation of three blood glucose monitoring systems using ISO 15197: 2013 accuracy criteria, consensus and surveillance error grid analyses, and insulin dosing error modeling in a hospital setting. *J Diabetes Sci Technol*. 2015;10:85-92.
19. Nakamura K, Balo A. The accuracy and efficacy of the Dexcom G4 Platinum continuous glucose monitoring system. *J Diabetes Sci Technol*. 2015;9:1021-1026.
20. 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.
21. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307-310.
22. Klonoff DC, Lias C, Vigersky R, et al.; Error Grid Panel. The surveillance error grid. *J Diabetes Sci Technol*. 2014;8:658-672.
23. 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.
24. Pardo S, Simmons DA. The quantitative relationship between ISO 15197 accuracy criteria and mean absolute relative difference (MARD) in the evaluation of analytical performance of self-monitoring of blood glucose (SMBG) Systems. *J Diabetes Sci Technol*. 2016;10:1182-1187.
25. Breton MD, Kovatchev BP. Impact of blood glucose self-monitoring errors on glucose variability, risk for hypoglycemia, and average glucose control in type 1 diabetes: an in silico study. *J Diabetes Sci Technol*. 2010;4:562-570.
26. 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.
27. Reiterer F, Polterauer P, Schoemaker M, et al. Significance and reliability of MARD for the accuracy of CGM systems. *J Diabetes Sci Technol*. 2017;11:59-67.

28. Clarke WL, Cox D, Gonder-Frederick LA, Carter W, Pohl SL. Evaluating clinical accuracy of systems for self-monitoring of blood glucose. *Diabetes Care*. 1987;10:622-628.

29. Twomey PJ. Plasma glucose measurement with the Yellow Springs Glucose 2300 STAT and the Olympus AU640. *J Clin Pathol*. 2004;57:752-754.

30. 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;10:872-875.

31. Thennadil SN, Rennert JL, Wenzel BJ, Hazen KH, Ruchti TL, Block MB. Comparison of glucose concentration in interstitial fluid, and capillary and venous blood during rapid changes in blood glucose levels. *Diabetes Technol Ther*. 2001;3:357-365.

32. Kropff J, van Steen SC, deGraaff P, Chan MW, van Amstel RBE, DeVries JH. Venous, arterialized-venous, or capillary glucose reference measurements for the accuracy assessment of a continuous glucose monitoring system. *Diabetes Technol Ther*. 2017;19:609-617.

33. Macleod K, Katz LB, Cameron H. Capillary and venous blood glucose accuracy in blood glucose meters versus reference standards: the impact of study design on accuracy evaluations. *J Diabetes Sci Technol*. 2018. doi:10.1177/1932296818790228

34. Zisser HC, Bailey TS, Schwartz S, Ratner RE, Wise J. Accuracy of the SEVEN continuous glucose monitoring system: comparison with frequently sampled venous glucose measurements. *J Diabetes Sci Technol*. 2009;3:1146-1154.

35. Charleer S, Mathieu C, Nobels F, Gillard P. Accuracy and precision of flash glucose monitoring sensors inserted into the abdomen and upper thigh compared with the upper arm. *Diabetes Obes Metab*. 2018;20:1503-1507.

36. Freckmann G, Pleus S, Link M, Zschornack E, Klotzer HM, Haug C. Performance evaluation of three continuous glucose monitoring systems: comparison of six sensors per subject in parallel. *J Diabetes Sci Technol*. 2013;7:842-853.

37. Zijlstra E, Heise T, Nosek L, Heinemann L, Heckermann S. Continuous glucose monitoring: quality of hypoglycaemia detection. *Diabetes Obes Metab*. 2013;15:130-135.

38. Bailey TS. Clinical implications of accuracy measurements of continuous glucose sensors. *Diabetes Technol Ther*. 2017; 19(suppl 2):s51-s54.

39. Freckmann G, Link M, Westhoff A, Kamecke U, Pleus S, Haug C. Prediction quality of glucose trend indicators in two continuous tissue glucose monitoring systems. *Diabetes Technol Ther*. 2018;20:550-556.

40. New JP, Ajjan R, Pfeiffer AF, Freckmann G. Continuous glucose monitoring in people with diabetes: the randomized controlled Glucose Level Awareness in Diabetes Study (GLADIS). *Diabet Med*. 2015;32:609-617.

41. Hermanns N, Ehrmann D, Kulzer B. How much accuracy of interstitial glucose measurement is enough? Is there a need for new evidence? *J Diabetes Sci Technol*. 2017;11:296-298.

42. Aleppo G, Ruedy KJ, Riddlesworth TD, et al.; REPLACE-BG Study Group. REPLACE-BG: a randomized trial comparing continuous glucose monitoring with and without routine blood glucose monitoring in adults with well-controlled type 1 diabetes. *Diabetes Care*. 2017;40:538-545.

43. Beck RW, Riddlesworth TD, Ruedy K, et al.; DIAMOND Study Group. Continuous glucose monitoring versus usual care in patients with type 2 diabetes receiving multiple daily insulin injections: a randomized trial. *Ann Intern Med*. 2017;167:365-374.

44. Bolinder J, Antuna R, Geelhoed-Duijvestijn P, Kroger J, Weitgasser R. Novel glucose-sensing technology and hypoglycaemia in type 1 diabetes: a multicentre, non-masked, randomised controlled trial. *Lancet*. 2016;388:2254-2263.

45. Haak T, Hanaire H, Ajjan R, Hermanns N, Riveline JP, Rayman G. 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.

46. Basu A, Pieber TR, Hansen AK, et al. Greater early postprandial suppression of endogenous glucose production and higher initial glucose disappearance is achieved with fast-acting insulin aspart compared with insulin aspart. *Diabetes Obes Metab*. 2018;20:1615-1622.

47. Basu A, Dube S, Veettil S, et al. Time lag of glucose from intravascular to interstitial compartment in type 1 diabetes. *J Diabetes Sci Technol*. 2015;9:63-68.

48. 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.

49. Schmelzeisen-Redeker G, Staib A, Strasser M, Muller U, Schoemaker M. Overview of a novel sensor for continuous glucose monitoring. *J Diabetes Sci Technol*. 2013;7:808-814.

50. Schmelzeisen-Redeker G, Schoemaker M, Kirchsteiger H, Freckmann G, Heinemann L, Del Re L. Time delay of CGM sensors: relevance, causes, and countermeasures. *J Diabetes Sci Technol*. 2015;9:1006-1015.