Variation of Mean Absolute Relative Differences of Continuous Glucose Monitoring Systems Throughout the Day

Stefan Pleus, MSc[1] [iD]**, Andreas Stuhr, MD[2], Manuela Link, MD[1],
Cornelia Haug, MD[1], and Guido Freckmann, MD[1]** [iD]

Abstract

Background: There is an increasing use of continuous glucose monitoring (CGM) by people with diabetes. Measurement performance is often characterized by the mean absolute relative difference (MARD). However, MARD is influenced by a number of factors and little is known about whether MARD is stable throughout the day.

Material and Methods: A total of 24 participants with type 1 diabetes were enrolled in the study. The study was performed for seven in-patient days. Participants wore two CGM systems in parallel and performed additional frequent blood glucose (BG) measurements. On two days, glucose excursions were induced.

MARD was calculated between pairs of CGM and BG values, with BG values serving as reference values. ARD values calculated from CGM-BG pairs were grouped by hour of the day. Results were analyzed separately for glucose excursion days and for regular days.

Results: Total MARDs for the complete study duration were 12.5% ± 3.6% and 13.2% ± 2.4% ($n=24$). Throughout the day marked variability of MARD was observed (8.0% ± 1.3%-16.3% ± 2.9% (G5); 9.1% ± 1.4%-16.3% ± 5.3% (FL), up to $n=157$ each). Low(est) MARD values were observed before breakfast and dinner, when subjects were in or near a fasting state. Especially after breakfast and lunch, MARD values were higher than average.

Conclusions: Analytical performance of the two CGM systems, assessed by MARD, was found to vary markedly throughout the day. Activities of daily life likely triggered these variations. An increasing number of CGM users base therapeutic decisions on CGM values, and they should be aware of these variations of performance throughout the day.

Keywords

mean absolute relative difference, circadian rhythm, continuous glucose monitoring, type 1 diabetes

Introduction

People with diabetes, especially those on any kind of insulin therapy, have to regularly check their glucose concentrations. For a long time, they used to (and most still do) rely on blood glucose (BG) values, obtained with systems for self-monitoring of BG. In recent years more and more people transitioned to using continuous glucose monitoring (CGM) systems. Some CGM systems are intended to replace BG measurements for therapeutic decisions in certain situations whereas others have to be used adjunct to self-monitoring of BG. With the increasing use of CGM by people with diabetes, there has been a shift towards a more comprehensive use of CGM data in diabetes management.

The analytical performance of CGM systems is known to affect key metrics that are recommended to be used for diabetes management.[1-3] However, internationally recognized standards to characterize the accuracy with which CGM systems measure glucose concentrations in interstitial fluid do currently not exist. Mean absolute relative differences (MARD) are often used to characterize the measurement performance of CGM systems.[4] The MARD is the average of the absolute relative differences between the measurements of the CGM system in the interstitial fluid and corresponding glucose values measured by a comparison method, which is

[1]Institut für Diabetes-Technologie, Forschungs- und Entwicklungsgesellschaft mbH an der Universität Ulm, Ulm, Germany
[2]Ascensia Diabetes Care, Parsippany, NJ, USA

Corresponding Author:
Stefan Pleus, MSc, Institut für Diabetes-Technologie, Forschungs- und Entwicklungsgesellschaft mbH an der Universität Ulm, Lise-Meitner-Straße 8/2, Ulm D-8908, Germany.
Email: stefan.pleus@idt-ulm.de

usually a lab analyzer or a BG meter.[5] However, the procedures to obtain data for MARD are not standardized and no internationally recognized acceptance criteria exist.

MARD is known to be affected by various factors,[6,7] including physiological differences between blood and interstitial fluid, which are predominantly depending on rates of glucose changes.[8] During steady-state conditions, glucose levels in blood and those provided by CGM, that is, interstitial fluid levels calibrated to BG levels, are almost similar; however, patients with diabetes usually spent most of the day outside of steady-state conditions. Consequently, MARD values may vary throughout the day. However, only one overall MARD is provided to describe accuracy of a CGM system throughout the whole day.

There is a large amount of literature investigating CGM system performance,[9-13] and some publications also stratify by day of sensor lifetime or glucose ranges.[12,13] In some instances, system performance is stratified by time since last calibration,[14] without detailed information about the circumstances in which the underlying data were recorded. It remains unclear how this affects potential variations in CGM system performance throughout the day.

In a study performed during the development of a new CGM system, subjects also wore two other, commercially available CGM systems. These systems' data were used to assess variations of CGM system performance throughout the day based on MARD values.

Materials and Methods

This study was conducted between September 2018 and February 2019 at the Institut für Diabetes-Technologie Forschungs- und Entwicklungsgesellschaft mbH an der Universität Ulm (IfDT), Ulm, Germany, in compliance with the national regulations and provisions and under consideration of the Declaration of Helsinki (revised edition, Fortaleza 2013). The study protocol was approved by the responsible Ethics Committee and the German competent authority [Federal Institute for Drugs and Medical Devices (BfArM)]. The study was registered in the German Clinical Trial Register (Deutsches Register Klinischer Studien, DRKS) with the number DRKS00023538.

Study Participants

A total of 24 participants with type 1 diabetes were enrolled in the study. Subjects had to be 18 to 70 years old to be eligible for enrolment. Potential participants must not have had a severe hypoglycemia in the three months prior to enrolment or suffer from hypoglycemia unawareness, and they had to refrain from intake of paracetamol/acetaminophen. Further exclusion criteria were pregnancy or lactation period, known severe tape reaction/allergies, body-mass index <20 kg/m² or serious chronic or acute disease apart from diabetes that might pose an undue additional risk to the subject.

Continuous Glucose Monitoring Systems

In this study, FreeStyle Libre (FL; Abbott Diabetes Care, Witney, UK) and Dexcom G5 (G5; Dexcom Inc., San Diego, CA) were used by all subjects in parallel. Following manufacturers' labeling, FL was worn on the upper arm and G5 was worn on the abdomen. G5 was calibrated every 12 hours at approx. 07:00 and 19:00 using values obtained with a study-specific CONTOUR®NEXT ONE blood glucose monitoring system (Ascensia Diabetes Care Holdings AG, Basel, Switzerland).

Study Procedures

All potential subjects signed informed consent forms during the screening visit, which took place up to 6 weeks before start of the experimental phase, before any study procedures were performed.

Participants arrived at the study site and stayed for eight calendar days (approximately 168 hours) at the clinical site. On the first study day study staff applied an FL sensor on the back of an upper arm and a G5 sensor on the abdomen of each subject. Although the two investigated CGM systems were indicated by their respective manufacturer to potentially replace BG monitoring for therapeutic decisions, all participants were provided with the study-specific BG monitoring system. This BG monitoring system was used for therapeutic decisions and to calibrate G5 twice daily according to the manufacturer's instructions. BG measurements were performed in duplicate hourly between getting up (approx. 07:00) and going to bed (approx. 22:00), with an additional BG measurement at night at approx. 03:00. Three major meals were scheduled, and additional snacks between the meals were allowed; all meals and time of intake were documented. During the next seven days, participants followed their usual diabetes regimen except for induced glucose excursions after breakfast on two days (second and seventh day). These glucose excursions were induced by providing high-carbohydrate breakfast meals and delaying and increasing the corresponding insulin dose, which lead to early post-prandial hyperglycemia and late post-prandial hypoglycemia. BG was monitored more frequently during these periods (every 15 minutes) for safety reasons. On the eighth study day, after approximately 168 hours of CGM sensor wear time, CGM sensors were removed by study staff. Outside of induced glucose excursions, participants were allowed to leave the site for a few hours at a time, for example, to go on walks.

Data Analysis

Mean absolute relative difference (MARD) was calculated between pairs of CGM values and BG values, with BG values serving as reference values. BG values were excluded if the second value in a duplicate differed by more than ±10 mg/dL

or more than $\pm 10\%$ from the first one, whichever was larger. Otherwise, the first BG value was used in the analysis. CGM values were paired depending on the CGM system to simulate the end-user experience. For G5, the most recently recorded value within the last 5 minutes before a BG measurement was used. For FL, the nearest scanned value within ± 2.5 minutes of the BG measurement was used.

Total MARD for the complete study duration was calculated as average across subject-specific MARD values ($n = 24$).

For hourly MARD, ARD values calculated from CGM-BG pairs were grouped by hour of the day based on the measurements' timestamps, if the BG measurements were performed within ± 10 minutes of a full hour, or excluded otherwise.

In order to provide more reliable results, MARD across the individual ARD values was only calculated if at least 100 individual ARD values (out of potentially 168 values from 24 participants over 7 days) were available for a specific hour of the day. Results were analyzed separately for glucose excursion days and for days without induced glucose excursion ("regular days") to assess the potential influence of induced glucose excursions with rapid rates of change on MARD.

MARD was also stratified by BG rates of change, by calculated as the difference between two subsequently recorded results divided by the elapsed time in minutes. For BG data, this approach was considered adequate only for the glucose excursions, during which BG measurements were performed approximately every 15 minutes. This stratification was only done for BG rates of change because of deviations between G5 rates of change and FL rates of change (see Supplemental Figure SF1 and Supplemental Tables ST1-ST3).

Similarly to calculation of MARD, CGM-BG pairs were used to calculated mean relative differences (MRD) as well as percentage of results within ± 15 mg/dL or $\pm 15\%$ (for BG concentrations <100 mg/dL or ≥ 100 mg/dL, respectively), within ± 20 mg/dL or $\pm 20\%$ and within ± 30 mg/dL or $\pm 30\%$.

Results

Complete Study

Total MARDs for the complete study duration were $12.5\% \pm 3.6\%$ (FL) and $13.2\% \pm 2.4\%$ (G5) (mean $\pm$ standard deviation, $n = 24$).

Detailed results are provided in Tables 1 to 3 and in Figures 1 to 4. Sufficient numbers of individual ARD values were available for 03:00 and from 07:00 to 22:00.

Before breakfast was started, both CGM systems showed low (ie, better) MARD values of 8.0% (G5) and 9.3% (FL). Both CGM systems' MARD values increased until 11:00 and remained markedly above the pre-breakfast value until dinner at 19:00. After dinner, G5 MARD values tended to remain stable, whereas FL MARD values increased slightly.

Glucose Excursion Days

MARD varied markedly on days on which glucose excursions were induced after breakfast. The lowest MARD values were found at 08:00, when breakfast started, for G5 (6.9%), and at 09:00 for FL (7.6%). Both CGM systems exhibited higher-than-average MARD results until 18:00.

BG concentrations showed that the induced glucose excursions resulted in a pronounced peak at 0900 and comparably low glucose values before lunch (Figure 2).

Between 08:00 and 13:00, when BG measurements were performed every 15 minutes, MARD results tended to be higher at faster BG rates of change (Figure 3).

Regular Days

On days without induced glucose excursions ("regular days"), MARD was lowest at 03:00 for G5 (7.6%) and at breakfast (08:00) for FL (8.9%). G5 showed its second-lowest MARD value before breakfast. Both CGM systems exhibited high MARD values at 11:00 (G5: 15.2%, FL: 18.8%), and MARD values remained comparably high until dinner at 19:00.

Median BG concentrations as well as 10th and 90th percentiles did not vary markedly throughout the day (Figure 4). Although there was a small peak after breakfast, it was far less pronounced than on glucose excursion days.

Discussion

In this study, marked variability of MARD throughout the day was observed. Low(est) MARD values were observed before breakfast (08:00) and dinner (19:00), when subjects were in or near a fasting state. Especially after breakfast and lunch, MARD values were higher (ie, worse) than average.

On glucose excursion days, MARD at 09:00 was comparably low for both systems. This could be influenced by the comparably large BG values at that time: MARD is calculated by dividing the difference between CGM and reference values by the reference value. For a specific difference, higher reference values would lead to smaller relative differences. On the other hand, MARD values at 11:00 were comparably high both on glucose excursion days and on regular days. Rapidly changing glucose concentrations at that time likely contributed to these higher MARD values.[8]

Results for G5 could be affected by calibration. G5's US version of the user guide indicates that there is a decrease in sensor performance throughout the calibration cycle.[14] Interestingly, the German version does not provide similar information.[15] G5 requires calibration once every 12 hours according to manufacturer's instructions, and calibrations were performed at 07:00 and at 19:00. For MARD calculation, the last value before calibration was used. Therefore, the 07:00 MARD value was based on the last values of the old calibration cycle, whereas the 08:00 MARD value was based

Table 1. Mean and 95% Confidence Intervals of Hourly BG, ARD, and RD, As Well As Percentage of Results Within ±15 mg/dL or ±15%, ±20 mg/dL or ±20%, and ±30 mg/dL or ±30% for FreeStyle Libre (FL) and Dexcom G5 (G5) Based on the Complete Study.

Hour of the day	Mean BG [95% CI] [mg/dL]	MARD G5 [95% CI] [%]	MARD FL [95% CI] [%]	MRD G5 [95% CI] [%]	MRD FL [95% CI] [%]	±15 mg/dL/ ±15% G5 [%]	±15 mg/dL/ ±15% FL [%]	±20 mg/dL/ ±20% G5 [%]	±20 mg/dL/ ±20% FL [%]	±30 mg/dL/ ±30% G5 [%]	±30 mg/dL/ ±30% FL [%]	n
3	147.2 [137.3, 157.0]	9.3 [7.6, 10.9]	11.4 [9.3, 13.6]	2.1 [−0.3, 4.4]	−8.9 [−11.5, −6.4]	81.8	78.2	90.9	85.5	96.4	94.5	110
7	155.5 [146.4, 164.7]	12.5 [10.9, 14.1]	9.9 [8.6, 11.1]	−8.8 [−10.9, −6.7]	−7.9 [−9.4, −6.3]	68.2	83.4	80.9	92.4	94.3	97.5	157
8	150.1 [142.2, 158.0]	8.0 [6.7, 9.2]	9.3 [8.0, 10.5]	1.4 [−0.4, 3.2]	−5.2 [−6.9, −3.4]	87.0	83.6	92.5	89.7	98.6	99.3	146
9	210.8 [197.9, 223.8]	10.3 [8.9, 11.7]	9.7 [8.3, 11.1]	−2.2 [−4.3, 0.0]	−4.7 [−6.6, −2.7]	73.4	75.5	88.1	89.5	97.9	98.6	143
10	186.5 [174.4, 198.6]	13.2 [11.5, 14.9]	10.7 [9.2, 12.2]	7.1 [4.5, 9.6]	0.5 [−1.9, 2.9]	68.4	77.4	79.7	87.2	94.7	95.5	133
11	133.5 [124.0, 143.1]	16.3 [13.4, 19.3]	16.3 [11.0, 21.7]	9.5 [5.8, 13.2]	4.4 [−1.5, 10.4]	63.2	69.9	77.9	77.9	89.7	91.9	136
12	121.8 [114.5, 129.0]	12.2 [10.6, 13.8]	11.8 [10.0, 13.6]	2.8 [0.3, 5.4]	−1.3 [−4.0, 1.3]	73.4	76.9	87.4	86.0	96.5	95.8	143
13	120.9 [113.3, 128.5]	11.4 [9.8, 13.0]	12.2 [10.5, 13.9]	0.4 [−2.0, 2.8]	−6.0 [−8.4, −3.6]	80.1	75.0	85.9	87.2	94.9	94.9	156
14	151.2 [142.5, 159.8]	14.2 [12.4, 15.9]	13.4 [11.4, 15.4]	−3.4 [−6.2, −0.6]	−6.3 [−9.0, −3.5]	64.6	71.4	77.6	81.0	92.5	89.1	147
15	149.3 [139.9, 158.7]	15.8 [13.2, 18.4]	12.9 [11.1, 14.8]	8.1 [4.7, 11.5]	1.0 [−1.8, 3.8]	68.1	71.5	77.1	81.9	88.9	95.1	144
16	136.3 [127.7, 144.9]	15.3 [13.3, 17.4]	13.7 [11.6, 15.8]	3.5 [0.3, 6.7]	−2.5 [−5.6, 0.5]	59.0	70.1	75.7	81.3	93.1	92.4	144
17	143.7 [136.2, 151.2]	14.7 [12.8, 16.7]	14.3 [12.1, 16.5]	−0.6 [−3.7, 2.5]	−5.2 [−8.3, −2.1]	62.6	69.4	74.8	79.6	91.8	90.5	147
18	150.2 [142.3, 158.0]	14.2 [11.7, 16.7]	13.9 [11.6, 16.1]	5.8 [2.5, 9.0]	0.5 [−2.7, 3.6]	69.2	66.4	80.8	80.1	90.4	93.2	146
19	152.9 [144.9, 160.9]	10.5 [9.0, 12.0]	9.1 [7.8, 10.5]	2.1 [−0.1, 4.4]	−4.6 [−6.5, −2.8]	78.4	85.8	87.8	91.9	94.6	97.3	148
20	155.1 [146.6, 163.5]	10.5 [9.2, 11.9]	11.0 [9.5, 12.4]	−1.5 [−3.7, 0.6]	−4.3 [−6.5, −2.2]	76.5	77.2	91.3	87.2	99.3	96.6	149
21	154.0 [145.1, 162.9]	9.7 [8.1, 11.2]	12.0 [10.2, 13.9]	0.6 [−1.6, 2.8]	−5.5 [−8.0, −2.9]	81.1	75.5	89.5	82.5	95.8	93.0	143
22	167.5 [158.4, 176.7]	9.5 [8.0, 11.0]	11.4 [9.3, 13.6]	−2.0 [−4.1, 0.2]	−4.2 [−7.0, −1.4]	77.9	77.1	90.1	84.7	97.7	93.9	131

Abbreviations: ARD, absolute relative differences; BG, blood glucose; CI, confidence interval; FL, FreeStyle Libre; MARD, mean absolute relative difference; MRD, mean relative difference RD, relative differences.

Table 2. Mean and 95% Confidence Intervals of Hourly BG, ARD, and RD, As Well As Percentage of Results Within ±15 mg/dL or ±15%, ±20 mg/dL or ±20%, and ±30 mg/dL or ±30% for FreeStyle Libre (FL) and Dexcom G5 (G5) on Days with Induced Glucose Excursions. Glucose Excursions were Induced After Breakfast by Providing a Carbohydrate-Rich Breakfast and Delaying and Increasing the Corresponding Insulin Dose.

Hour of the day	Mean BG [95% CI] [mg/dL]	MARD G5 [95% CI] [%]	MARD FL [95% CI] [%]	MRD G5 [95% CI] [%]	MRD FL [95% CI] [%]	±15 mg/dL ±15% G5 [%]	±15 mg/dL ±15% FL [%]	±20 mg/dL ±20% G5 [%]	±20 mg/dL ±20% FL [%]	±30 mg/dL ±30% G5 [%]	±30 mg/dL ±30% FL [%]	n
3	140.4 [123.3, 157.5]	13.0 [8.9, 17.1]	14.0 [8.9, 19.1]	3.7 [−2.2, 9.6]	−11.6 [−17.3, −5.9]	67.6	70.6	76.5	73.5	91.2	85.3	34
7	159.2 [140.4, 178.1]	12.3 [9.3, 15.4]	11.3 [8.9, 13.8]	−5.9 [−10.3, −1.4]	−8.6 [−11.9, −5.4]	66.7	75.6	80.0	88.9	93.3	95.6	45
8	159.0 [142.3, 175.6]	6.9 [4.8, 9.0]	10.2 [8.0, 12.5]	3.2 [0.3, 6.0]	−6.0 [−9.4, −2.6]	87.8	80.5	92.7	87.8	100.0	100.0	41
9	297.2 [278.6, 315.7]	8.5 [6.2, 10.8]	7.6 [5.8, 9.5]	1.2 [−2.3, 4.6]	−3.2 [−6.0, −0.4]	81.4	86.0	93.0	95.3	97.7	100.0	43
10	214.2 [193.7, 234.8]	15.7 [12.5, 18.8]	10.3 [8.3, 12.3]	14.4 [10.9, 18.0]	3.6 [0.2, 7.0]	55.1	77.6	67.3	89.8	93.9	98.0	49
11	120.6 [105.4, 135.8]	18.1 [14.7, 21.6]	12.5 [9.5, 15.6]	14.8 [10.3, 19.3]	3.6 [−0.9, 8.1]	54.7	77.4	69.8	83.0	88.7	94.3	53
12	101.3 [92.0, 110.6]	14.8 [12.1, 17.6]	15.1 [11.7, 18.6]	6.1 [1.5, 10.6]	−0.3 [−5.6, 5.0]	69.6	71.4	83.9	82.1	94.6	92.9	56
13	95.7 [85.6, 105.7]	12.7 [9.9, 15.4]	12.4 [9.3, 15.5]	5.1 [0.8, 9.5]	−5.8 [−10.3, −1.4]	80.0	77.8	86.7	93.3	95.6	97.8	45
14	102.7 [93.1, 112.4]	15.5 [12.0, 18.9]	12.9 [9.3, 16.5]	2.1 [−3.7, 7.9]	−5.4 [−10.5, −0.3]	69.8	76.7	79.1	83.7	88.4	93.0	43
15	153.7 [141.6, 165.8]	12.2 [9.5, 15.0]	11.2 [8.6, 13.8]	4.7 [0.4, 9.0]	−2.3 [−6.4, 1.9]	73.9	73.9	82.6	78.3	93.5	97.8	46
16	138.0 [121.2, 154.8]	12.6 [8.9, 16.3]	11.3 [8.1, 14.5]	9.5 [5.0, 14.0]	−1.9 [−6.6, 2.9]	66.7	76.9	82.1	87.2	94.9	97.4	39
17	148.1 [133.5, 162.7]	12.4 [9.5, 15.3]	12.9 [9.3, 16.5]	2.8 [−1.9, 7.6]	−6.5 [−11.5, −1.5]	73.2	73.2	82.9	80.5	95.1	90.2	41
18	164.8 [152.7, 176.9]	8.9 [7.1, 10.8]	9.5 [6.8, 12.3]	4.7 [1.7, 7.6]	−5.0 [−8.6, −1.3]	86.4	81.8	95.5	93.2	97.7	95.5	44
19	171.1 [157.8, 184.3]	8.7 [6.2, 11.1]	8.9 [6.3, 11.5]	6.3 [3.3, 9.3]	−4.2 [−7.7, −0.6]	79.1	90.7	90.7	90.7	97.7	95.3	43
20	162.4 [149.5, 175.3]	8.9 [7.1, 10.8]	9.7 [7.6, 11.9]	1.8 [−1.4, 5.0]	−3.5 [−7.0, 0.0]	86.0	79.1	93.0	93.0	100.0	97.7	43
21	156.5 [139.4, 173.6]	8.5 [6.1, 10.8]	12.8 [8.3, 17.4]	−0.7 [−4.3, 2.8]	−9.2 [−14.6, −3.8]	84.2	76.3	89.5	81.6	97.4	86.8	38
22	168.7 [153.1, 184.2]	8.7 [6.3, 11.1]	9.7 [5.5, 14.0]	−3.9 [−7.4, −0.3]	−5.4 [−10.4, −0.4]	78.4	83.8	83.8	91.9	100.0	94.6	37

Abbreviations: ARD, absolute relative differences; BG, blood glucose; CI, confidence interval; FL, FreeStyle Libre; MARD, mean absolute relative difference; MRD, mean relative difference RD, relative differences.

Table 3. Mean and 95% Confidence Intervals of Hourly BG, ARD, and RD, As Well As Percentage of Results Within ±15 mg/dL or ±15%, ±20 mg/dL or ±20%, and ±30 mg/dL or ±30% for FreeStyle Libre (FL) and Dexcom G5 (G5) Based on Days Without Induced Glucose Excursions.
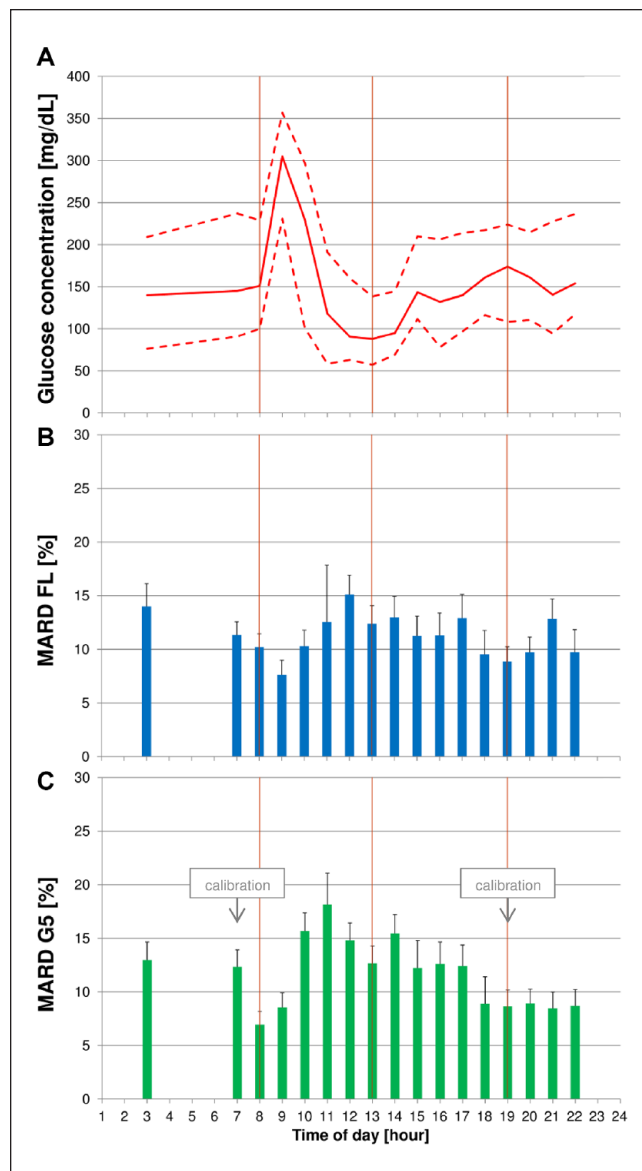
Hour of the day	Mean BG [95% CI] [mg/dL]	MARD G5 [95% CI] [%]	MARD FL [95% CI] [%]	MRD G5 [95% CI] [%]	MRD FL [95% CI] [%]	±15 mg/dL/ ±15% G5 [%]	±15 mg/dL/ ±15% FL [%]	±20 mg/dL/ ±20% G5 [%]	±20 mg/dL/ ±20% FL [%]	±30 mg/dL/ ±30% G5 [%]	±30 mg/dL/ ±30% FL [%]	n
3	150.2 [138.2, 162.2]	7.6 [6.1, 9.1]	10.3 [8.2, 12.4]	1.3 [−0.9, 3.6]	−7.7 [−10.3, −5.2]	88.2	81.6	97.4	90.8	98.7	98.7	76
7	154.0 [143.6, 164.5]	12.5 [10.7, 14.4]	9.3 [7.9, 10.7]	−10.0 [−12.3, −7.7]	−7.6 [−9.3, −5.8]	68.8	86.6	81.3	93.8	94.6	98.2	112
8	146.6 [137.9, 155.4]	8.4 [6.9, 9.9]	8.9 [7.4, 10.3]	0.7 [−1.5, 3.0]	−4.8 [−6.9, −2.8]	86.7	84.8	92.4	90.5	98.1	99.0	105
9	173.7 [163.5, 183.8]	11.1 [9.4, 12.7]	10.6 [8.8, 12.3]	−3.6 [−6.3, −1.0]	−5.3 [−7.8, −2.8]	70.0	71.0	86.0	87.0	98.0	98.0	100
10	170.3 [156.5, 184.2]	11.8 [9.9, 13.7]	11.0 [8.9, 13.1]	2.8 [−0.3, 5.9]	−1.3 [−4.5, 1.8]	76.2	77.4	86.9	85.7	95.2	94.0	84
11	141.8 [129.8, 153.7]	15.2 [10.9, 19.4]	18.8 [10.3, 27.2]	6.1 [0.9, 11.3]	4.9 [−4.4, 14.3]	68.7	65.1	83.1	74.7	90.4	90.4	83
12	134.9 [125.5, 144.3]	10.6 [8.7, 12.5]	9.7 [7.9, 11.5]	0.8 [−2.2, 3.7]	−2.0 [−4.7, 0.7]	75.9	80.5	89.7	88.5	97.7	97.7	87
13	131.1 [121.8, 140.4]	10.9 [8.9, 12.9]	12.2 [10.1, 14.2]	−1.5 [−4.4, 1.3]	−6.1 [−8.9, −3.3]	80.2	73.9	85.6	84.7	94.6	93.7	111
14	171.2 [162.0, 180.4]	13.7 [11.6, 15.7]	13.6 [11.1, 16.0]	−5.7 [−8.8, −2.5]	−6.6 [−10.0, −3.3]	62.5	69.2	76.9	79.8	94.2	87.5	104
15	147.2 [134.6, 159.9]	17.5 [13.9, 21.0]	13.7 [11.3, 16.2]	9.7 [5.1, 14.3]	2.6 [−1.1, 6.2]	65.3	70.4	74.5	83.7	86.7	93.9	98
16	135.7 [125.7, 145.8]	16.3 [13.9, 18.8]	14.6 [12.0, 17.2]	1.3 [−2.7, 5.3]	−2.8 [−6.6, 1.0]	56.2	67.6	73.3	79.0	92.4	90.5	105
17	142.0 [133.3, 150.8]	15.6 [13.2, 18.1]	14.8 [12.1, 17.5]	−1.9 [−5.8, 1.9]	−4.7 [−8.6, −0.9]	58.5	67.9	71.7	79.2	90.6	90.6	106
18	143.9 [134.1, 153.6]	16.5 [13.1, 19.8]	15.8 [12.8, 18.7]	6.3 [1.8, 10.8]	2.8 [−1.4, 7.0]	61.8	59.8	74.5	74.5	87.3	92.2	102
19	145.5 [135.9, 155.1]	11.2 [9.4, 13.1]	9.2 [7.6, 10.9]	0.4 [−2.4, 3.3]	−4.9 [−7.1, −2.6]	78.1	83.8	86.7	92.4	93.3	98.1	105
20	152.1 [141.5, 162.7]	11.2 [9.5, 12.9]	11.5 [9.7, 13.2]	−2.9 [−5.6, −0.2]	−4.6 [−7.3, −2.0]	72.6	76.4	90.6	84.9	99.1	96.2	106
21	153.1 [142.7, 163.6]	10.1 [8.2, 12.0]	11.8 [9.8, 13.7]	1.0 [−1.7, 3.7]	−4.1 [−7.0, −1.2]	80.0	75.2	89.5	82.9	95.2	95.2	105
22	167.1 [155.8, 178.3]	9.8 [7.9, 11.7]	12.1 [9.7, 14.5]	−1.2 [−3.9, 1.5]	−3.8 [−7.1, −0.4]	77.7	74.5	92.6	81.9	96.8	93.6	94

Abbreviations: ARD, absolute relative differences; BG, blood glucose; CI, confidence interval, FL, FreeStyle Libre; MARD, mean absolute relative difference; MRD, mean relative difference RD, relative differences.
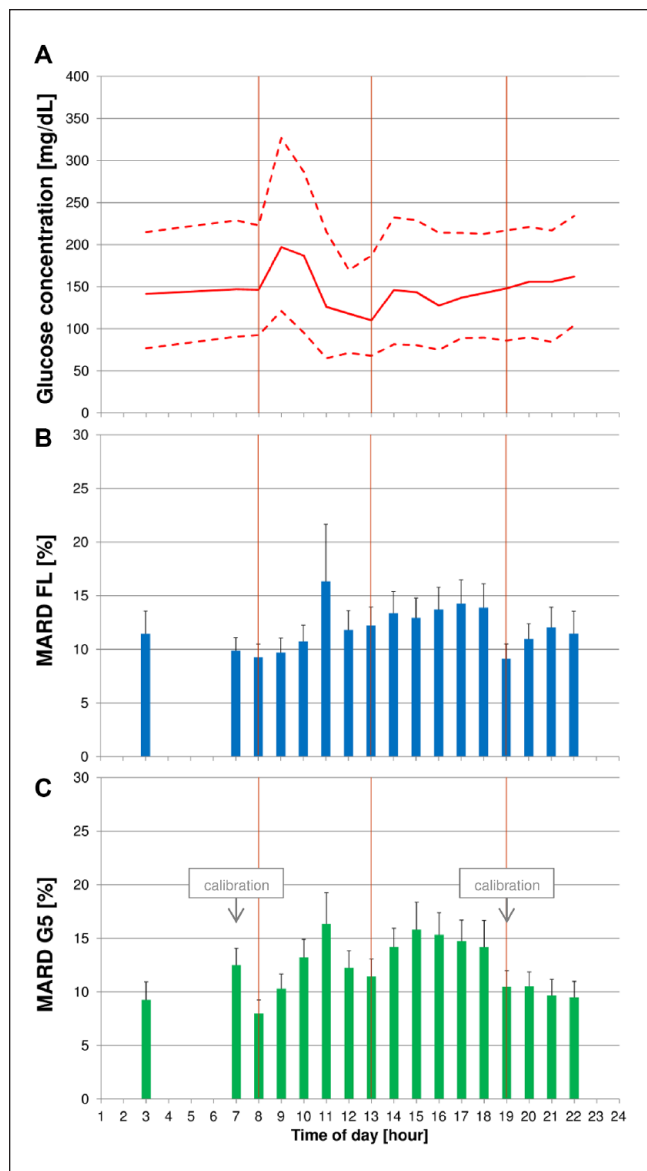
Figure 1. (A) Median blood glucose (BG) concentrations (continuous line) with 10th and 90th percentiles (dashed line) of the complete study. (B + C) Mean absolute relative differences (MARD) with upper bounds of 95% confidence intervals for FreeStyle Libre (FL) (B) and Dexcom G5 (G5) (C). Vertical lines indicate scheduled intake times for breakfast, lunch, and dinner.

Figure 2. (A) Median blood glucose (BG) concentrations (continuous line) with 10th and 90th percentiles (dashed line) on days with induced glucose excursions. (B + C) Mean absolute relative differences (MARD) with upper bounds of 95% confidence intervals for FreeStyle Libre (FL) (B) and Dexcom G5 (G5) (C). Glucose excursions were induced after breakfast by providing a carbohydrate-rich breakfast and delaying and increasing the corresponding insulin dose. Vertical lines indicate scheduled intake times for breakfast, lunch, and dinner.

on values obtained approximately one hour after calibration. The fact that G5 values and BG values did not diverge too much can likely be attributed to the absence of carbohydrate intake and strenuous physical activity between calibration and breakfast. This effect could not be reproduced at 20:00, because the second daily calibration coincided with dinner. However, the factory-calibrated FL also showed low MARD values before breakfast, and both systems exhibited higher-than-average MARD results in the afternoon. Therefore, calibration likely is not the only influence.

MARD values between lunch and dinner were found to be higher on regular days (Table 3, Figure 4) than on glucose excursion days with partially restricted daily life activities (Table 2, Figure 2). An explanation for these differences could be that participants were physically more active, especially in the afternoon on days without glucose excursion, which resulted in higher rates of glucose changes. Results for
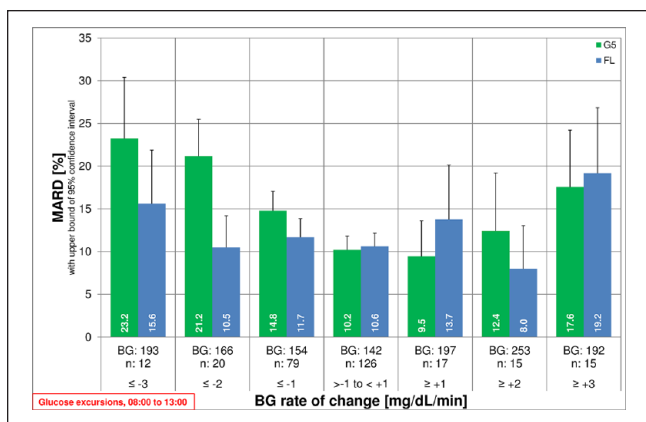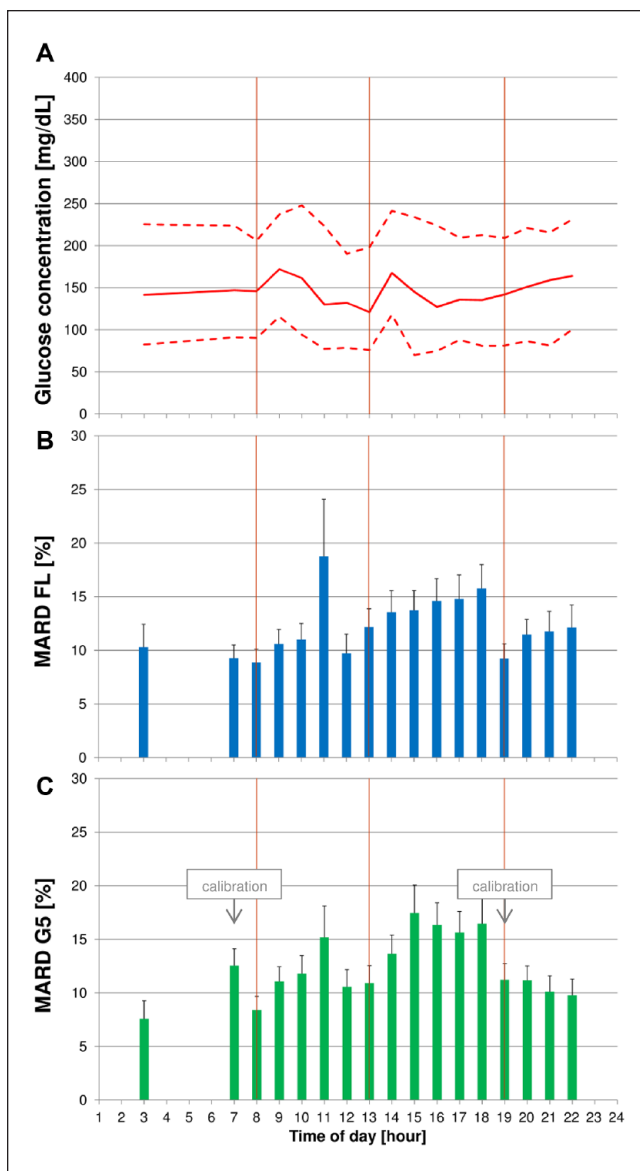
Figure 3. Mean absolute relative differences (MARD) with upper bound of two-sided 95% confidence interval for G5 and FL over blood glucose (BG) rate of change. X-axis label includes mean BG in mg/dL and number of results in the respective rate-of-change category.

percentages of results within specific limits for differences between CGM and BG values show a similar picture.

Physiologic differences of glucose levels between blood and interstitial fluid are more pronounced at higher rates of glucose change.[16] MARD calculations are based on concentrations from both compartments. Thus, MARD values can be higher (ie, worse) at higher rates of change than during steady-state.[4,8] Consequently, the overall quality of the CGM sensor itself is not the only factor influencing the MARD value; the physiologic conditions during the performance assessment also play an important role.[4,6,17] A study setting avoiding large glycemic variations (eg, with an appropriate choice of meals or (almost) exclusive enrolment of study participants with type 2 diabetes) may lead to a lower MARD than a setting in which these variations do occur. This potential impact could be a reason for the different MARD values for the same CGM system reported in a number of studies.[11,13,18]

Consequently, for the assessment of the analytical performance of CGM systems, the calculation of multiple MARD values for different time periods may be more representative than one overall MARD.[4] This is especially clinically important as more and more patients make treatment decisions solely based on CGM information—trusting its promoted overall performance—at any time of day.

A potential limitation of this study is the comparably small number of participants. In light of the relatively large standard deviations, the number of data points available per hour resulted in comparably large error margins: 95% confidence interval half-widths are, on average, 2% to 3%. As this may affect the size of differences that can be reliably detected, studies that explicitly try to assess MARD variations throughout the day should be planned with a larger numbers of data points in mind. While participants followed their usual diabetes regimen outside of induced glucose excursions, some

Figure 4. (A) Median blood glucose (BG) concentrations (continuous line) with 10th and 90th percentiles (dashed line) on regular days. (B + C) Mean absolute relative differences (MARD) with upper bounds of 95% confidence intervals for FreeStyle Libre (FL) (B) and Dexcom G5 (G5) (C). Vertical lines indicate scheduled intake times for breakfast, lunch, and dinner.

aspects of daily-life activities were structured to a certain degree, like meal times and times of getting up or going to bed. Although the breakfast on days with induced glucose excursions was rich with carbohydrates, it was designed as a meal that people with type 1 diabetes would eat, at least those who like sweet breakfast. However, the corresponding therapeutic decision, that is, the insulin dosing for meal compensation, was altered by increasing and delaying the insulin bolus. Having different subgroups that perform these activities according to different schedules would likely have led to MARD values exhibiting a different pattern, thus allowing

for a more detailed analysis of individual contributing factors. Still, this study was able to shed some light on a situation that was previously not explored in detail. BG values at night were very limited, so that a sufficient number of glucose values could only be obtained at 03:00. The study, therefore, does not allow to discern if the CGM systems performed differently at night than during the day, for example, due to pressure-induced sensor artifacts. The FL system used in the study was indicated for up to 14 days of use, so that potential changes in sensor sensitivity over time may only be addressed incompletely.

Nevertheless, considerable variation was found in MARD values throughout the days with both CGM systems. As MARD is seen as parameter for the analytical performance of a CGM system, and as such an indicator for the safety of clinical decision making, this may very well have implications on how CGM data should be used in the management of diabetes by patients. International consensus statements,[1,19] for example, highlight the need to assess CGM data also on graphical presentations like the ambulatory glucose profile (AGP). If, however, the analytical performance is systematically worse at some times of the day, as indicated by high MARD values, this suggests that CGM data at these times of day are less reliable and less safe to make clinical decisions, like dosing insulin. This study also suggests that MARD might not be the best parameter for assessments of overall CGM performance.

Recent analyses of data from the T1D Exchange database indicate that CGM does lead, on average, to improved HbA1c levels.[20] However, HbA1c levels in the registry as a whole did not improve between a 2010 to 2012 observation window and a 2016 to 2018 observation window even with increased use of CGM.[21] Knowledge of CGM's limitations may enable users to maximize the benefit provided by CGM.

Conclusions

Analytical performance of the two CGM systems, assessed by MARD, was found to vary markedly throughout the day. Activities of daily life likely impacted the timing of these variations in the study. As more and more people with diabetes wear CGM systems whose values can be used for immediate therapeutic decisions, CGM users (patients and caregivers alike) should be aware of these performance variations when managing their diabetes.

Such variations should also be kept in mind when performance studies are interpreted because study design could affect performance outcomes like MARD. Therefore, standardized assessment procedures should be defined to allow independent appraisal and comparability of MARDs. Definition of acceptance criteria both overall and as shown in this analysis throughout the day could improve the safe and effective use of CGM.

Abbreviations

BG, blood glucose (concentration); CGM, continuous glucose monitoring; FL, FreeStyle Libre; G5, Dexcom G5; MARD, mean absolute relative difference.

ORCID iDs

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

Guido Freckmann https://orcid.org/0000-0002-0406-9529

Supplemental Material

Supplemental material for this article is available online.

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;42:1593-1603.
2. Michalak A, Pagacz K, Mlynarski W, et al. Discrepancies between methods of continuous glucose monitoring in key metrics of glucose control in children with type 1 diabetes. *Pediatr Diabetes*. 2019;20:604-612.
3. 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*. Published online June 8, 2020. doi:10.1177/1932296820931825
4. Heinemann L, Schoemaker M, Schmelzeisen-Redecker G, et al. Benefits and limitations of MARD as a performance parameter for continuous glucose monitoring in the interstitial space. *J Diabetes Sci Technol*. 2020;14:135-150.
5. Clinical and Laboratory Standards Institute. POCT05-A - performance metrics for continuous interstitial glucose monitoring. 2008. Clinical and Laboratory Standards Institute. ISBN 1 56238 685 9.

6. 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.

7. 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.

8. 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.

9. 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.

10. 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.

11. Bonora B, Maran A, Ciciliot S, et al. Head-to-head comparison between flash and continuous glucose monitoring systems in outpatients with type 1 diabetes. *J Endocrinol Invest*. 2016;39:1391-1399.

12. Freckmann G, Link M, Pleus S, et al. Measurement performance of two continuous tissue glucose monitoring systems intended for replacement of blood glucose monitoring. *Diabetes Technol Ther*. 2018;20:541-549.

13. 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.

14. Dexcom. User guide: for Dexcom G5 Mobile Continuous Glucose Monitoring (CGM) system. Accessed November 30, 2020. https://s3-us-west-2.amazonaws.com/dexcompdf/G5-Mobile-Users-Guide-Touchscreen-Receiver.pdf

15. Dexcom. Dexcom G5 mobile - system zur kontinuierlichen Glukoseüberwachung - Bedienungsanleitung. Accessed November 30, 2020. https://s3-us-west-2.amazonaws.com/dexcompdf/Downloads+and+Guides+Updates/LBL013328+G5+Mobile+UG+OUS+DE+mgdL.pdf

16. 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.

17. Schrangl P, Reiterer F, Heinemann L, et al. Limits to the evaluation of the accuracy of continuous glucose monitoring systems by clinical trials. *Biosensors (Basel)*. 2018;8:50.

18. Aberer F, Hajnsek M, Rumpler M, et al. Evaluation of subcutaneous glucose monitoring systems under routine environmental conditions in patients with type 1 diabetes. *Diabetes Obes Metab*. 2017;19:1051-1055.

19. Danne T, Nimri R, Battelino T, et al. International consensus on use of continuous glucose monitoring. *Diabetes Care*. 2017;40:1631-1640.

20. Miller KM, Beck RW, Foster NC, et al. HbA1c levels in type 1 diabetes from early childhood to older adults: a deeper dive into the influence of technology and socioeconomic status on HbA1c in the T1D Exchange clinic registry findings. *Diabetes Technol Ther*. 2020;22:645-650.

21. Foster NC, Beck RW, Miller KM, et al. State of type 1 diabetes management and outcomes from the T1D Exchange in 2016-2018. *Diabetes Technol Ther*. 2019;21:66-72.