

## Performance Evaluation of Three Continuous Glucose Monitoring Systems: Comparison of Six Sensors per Subject in Parallel

Guido Freckmann, M.D.,<sup>1</sup> Stefan Pleus, M.S.,<sup>1</sup> Manuela Link, M.E.,<sup>1</sup> Eva Zschornack, M.D.,<sup>1</sup>  
Hans-Martin Klötzer, Ph.D.,<sup>2</sup> and Cornelia Haug, M.D.<sup>1</sup>

### Abstract

#### Background:

This study is aimed at comparing the performance of three continuous glucose monitoring (CGM) systems following the Clinical and Laboratory Standards Institute's POCT05-A guideline, which provides recommendations for performance evaluation of CGM systems.

#### Methods:

A total of 12 subjects with type 1 diabetes were enrolled in this study. Each subject wore six CGM systems in parallel, two sensors of each CGM system [FreeStyle Navigator™ (Navigator), MiniMed Guardian® REAL-Time with Enlite sensor (Guardian), DexCom™ Seven® Plus 3rd generation (Seven Plus)]. Each sensor was used for the lifetime specified by the manufacturer. To follow POCT05-A recommendations, glucose excursions were induced on two separate occasions, and venous and capillary blood glucose (BG) concentrations were obtained every 15 min for five consecutive hours. Capillary BG concentrations were measured at least once per hour during the day and once at night. Parameters investigated were CGM-to-BG differences [mean absolute relative difference (MARD)] and sensor-to-sensor differences [precision absolute relative difference (PARD)].

#### Results:

Compared with capillary BG reference readings, the Navigator showed the lowest MARD, with 12.1% overall and 24.6% in the hypoglycemic range; for the Guardian and the Seven Plus, MARD was 16.2%/34.9% and 16.3%/32.7%, respectively. PARD also was lowest for the Navigator (9.6%/9.8%), followed by the Seven Plus (16.7%/25.5%) and the Guardian (18.1%/20.2%). During induced glucose excursions, MARD between CGM and BG was, again, lowest for the Navigator (14.3%), followed by the Seven Plus (15.8%) and the Guardian (19.2%).

#### Conclusions:

In this study, two sensors of each of the three CGM systems were compared in a setting following POCT05-A recommendations. The Navigator CGM system achieved more accurate results than the Guardian or the Seven Plus with respect to MARD and PARD. Performance in the hypoglycemic range was markedly worse for all CGM systems when compared with BG results.

*J Diabetes Sci Technol* 2013;7(4):842–853

**Author Affiliations:** <sup>1</sup>Institut für Diabetes-Technologie Forschungs- und Entwicklungsgesellschaft mbH an der Universität Ulm, Ulm, Germany; and <sup>2</sup>HMKQ Quantitative Analytik - Chemometrie, Weinheim, Germany

**Abbreviations:** (BG) blood glucose, (CGM) continuous glucose monitoring, (EGA) error grid analysis, (MARD) mean absolute relative difference, (PARD) precision absolute relative difference, (SMBG) self-monitoring of blood glucose

**Keywords:** continuous glucose monitoring, glucose sensors, head-to-head comparison, hypoglycemia, type 1 diabetes mellitus

**Corresponding Author:** Stefan Pleus, M.S., Institut für Diabetes-Technologie Forschungs- und Entwicklungsgesellschaft mbH an der Universität Ulm, Helmholtzstrasse 20, 89081 Ulm, Germany; email address [stefan.pleus@uni-ulm.de](mailto:stefan.pleus@uni-ulm.de)

## Background

**R**elevant factors in the selection of a continuous glucose monitoring (CGM) system include accuracy of the glucose measurements and reliability of the CGM system. In 2008, the Clinical and Laboratory Standards Institute published POCT05-A, which provides recommendations for study design and parameters of interest in the performance evaluation of CGM systems.<sup>1</sup>

Multiple studies investigating CGM systems were published over the years;<sup>2–19</sup> only one claimed to use procedures recommended in POCT05-A.<sup>13</sup> A comparison of these studies, however, is difficult due to variations in subject groups<sup>11</sup> and study design, including use of different devices for calibration of CGM systems and for reference measurements, either overall or for different study phases.<sup>2–4,6–10,12,13,16,17</sup> Given that comparability of different CGM systems is heavily influenced by these factors, it is thus improved when the different CGM systems are worn simultaneously by the same subjects under identical conditions (e.g., in head-to-head comparisons). There are studies in which different CGM systems were compared under such conditions,<sup>2,7,12</sup> but in these cases, only one sensor per CGM system was worn by the subjects.

The objective of this study was to investigate the performance of three needle-type CGM systems with two sensors each in parallel in a setting designed to represent daily life conditions and also to incorporate as many of the recommendations of POCT05-A as possible.

## Research Design and Methods

### *Participants*

For this study, 12 subjects were screened and enrolled. All 12 participants (six female, six male) had type 1 diabetes; 11 were treated with continuous subcutaneous insulin infusion, and 1 was treated with intensified insulin therapy using a pen. The mean age  $\pm$  standard deviation of the subjects was  $48 \pm 7$  years (range from 40 to 62 years); the duration of diabetes was  $24 \pm 10$  years (11 to 38 years); body mass index was  $26.5 \pm 4.2$  kg/m<sup>2</sup> (21.6 to 35.5 kg/m<sup>2</sup>), and hemoglobin A1c was  $8.2\% \pm 1.4\%$  (6.5% to 11.9%).

### *Continuous Glucose Monitoring Systems*

This study compared three needle-type CGM systems: FreeStyle Navigator™ by Abbott Diabetes Care, Alameda, CA (Navigator); Guardian® REAL-Time by Medtronic MiniMed, Northridge, CA (Guardian); and DexCom™ Seven® Plus by DexCom, San Diego, CA (Seven Plus). The Guardian was used with Enlite sensors, which were introduced to the European market in 2011; the Seven Plus was used with third-generation sensors. All sensors and disposables were bought on the German market by the investigators. No selection of sensor batches was performed.

### *Study Design*

This open-label, prospective, single-arm study was conducted between October 2011 and December 2011 at the Institute for Diabetes-Technology GmbH in Ulm, Germany, in compliance with the German Medical Devices Act and the Good Clinical Practice provisions of the Declaration of Helsinki. The ethics committee and the competent authority responsible approved the study protocol. Written informed consents were provided by all subjects before beginning study procedures. Each subject visited the study site once for nine consecutive days and wore six sensors in parallel, two sensors of each system. The CGM systems were used according to manufacturers' labeling, especially with regard to sensor location (to be discussed) on the body, calibration intervals, and sensor lifetime (5 days for the Navigator, 6 days for the Guardian, and 7 days for the Seven Plus).

Subjects arrived at the study site on day 0 and remained overnight for 8 nights. All CGM sensors were inserted by a physician into the abdomen between 6:00 AM and 7:00 AM of day 1 to achieve a similar start time for the sensors and ensure that calibration was performed during a period of a low rate of glucose change. One sensor per system was worn on the right side of the abdomen, the other one on the left side. Both sensors of one type of system were worn

either on the upper, middle, or lower part of the abdomen; sensor location was sequentially changed in every four subjects. Sensors were removed by a physician after the CGM system's sensor removal alert. One sensor per system was labeled "A"; the other sensor was labeled "B" to provide unique identification for each sensor.

Throughout the study, capillary blood glucose (BG) was measured at least once per hour between 6:00 AM and 11:00 PM and again at 3:00 AM with the Navigators' built-in BG monitoring meters (Navigator-BG). The built-in meter was used because the Navigator did not allow input of external BG readings for calibration. Subjects performed measurements with each of two Navigator-BG meters (A and B) in parallel. In case the Navigator-BG B result deviated by more than 10% (10 mg/dl) from the Navigator-BG A result [at Navigator-BG A results above (below) 100 mg/dl], measurements with both systems were repeated. All Navigator-BG meters were tested with control solution before they were handed out to the subjects and after they were returned to the study staff to ensure proper functionality. All devices used showed valid control measurements.

Calibration of CGM systems was performed according to manufacturer labeling: after approximately 1, 2, 10, 24, and 72 h for the Navigator; after approximately 2 and 8 h and then every 12 h for the Guardian; and after 2 h and then every 12 h for the Seven Plus. Guardian and Seven Plus calibrations were synchronized and performed at 7:00 AM and 7:00 PM, and the 24 and 72 h calibrations of the Navigator were performed at 7:00 AM. Synchronization of Guardian and Seven Plus calibrations required additional, out-of-turn calibrations on day 1 and again on day 4, after the Guardian was restarted according to manufacturer labeling. This restart was necessary because the Guardian's software automatically ended the sensor experiment after 3 days. Manufacturer labeling stated that the sensor should be restarted once so the sensor could be used for 6 days. For calibration of CGM systems, capillary BG measurement results obtained with the Navigator-BG A and B systems were used. Continuous glucose monitoring sensors labeled "A" (or "B") were calibrated with Navigator-BG A (or B) results. The initial Seven Plus calibration, which requires input of two values, for Seven Plus A was performed entering the BG result of the Navigator-BG A, followed by the Navigator-BG B result, and vice versa for Seven Plus B.

For evaluation of CGM system performance following POCT05-A recommendations, glucose excursions were induced on days 2 and 5 by serving fast-absorbing meals (approximately 80% carbohydrates, 20% of daily caloric need) at 8:00 AM and 11:00 AM, respectively. To provide considerable postprandial glycemic excursions, the corresponding insulin doses, which were calculated by a physician based on subjects' individual factors, were delayed by 15 min and increased by approximately 15% at the attending physician's discretion. The times of meal bolus delivery were determined considering the subjects' individual injection-meal interval. For five consecutive hours after these meals, capillary BG was measured every 15 min. Venous blood was sampled in parallel to each capillary measurement. Capillary BG concentrations at morning and evening CGM calibration times and venous BG concentrations were measured in plasma with a laboratory reference [YSI 2300 STAT Plus (YSI), Yellow Springs, OH].

Subjects were generally allowed to move around the study site. Except for the standardized breakfast meals before glucose excursions, subjects could compose their meals without restrictions.

The following three periods of time were defined:

1. **Complete experiments** (5 days for the Navigator, 6 days for the Guardian, and 7 days for the Seven Plus) to provide an evaluation of each sensor's full lifetime. In the analysis, the "days" were aligned to the sensor lifetimes, thus each day began at 7:00 AM of one calendar day and ended at 7:00 AM the following calendar day, with exception of day 1, which began with the recording of the first sensor reading.
2. **Core phase** (7:00 AM day 2 to 7:00 AM day 5) to provide maximum comparability for all CGM systems (compensation of different run-in times as well as different initial calibrations and other issues at sensor start on day 1; identical glucose fluctuations, thus true parallel use of all sensors). For this data set, clinical analysis was performed using Clarke error grid analysis (EGA).
3. **Induced glucose excursions** for evaluation, in compliance with POCT05-A recommendations.

Figure 1 shows data of a representative experiment (i.e., data from one subject).

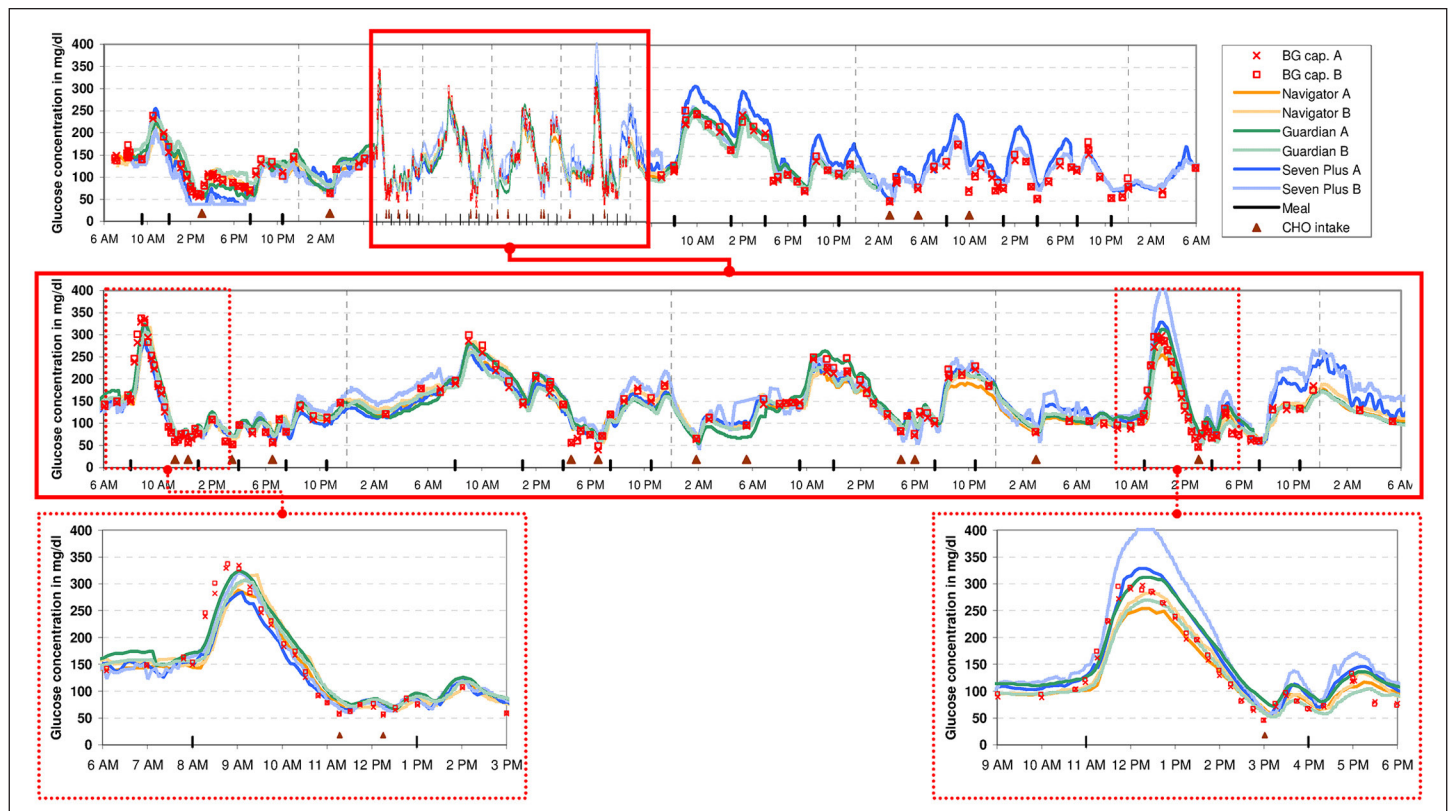


Figure 1. Study procedures as displayed by a representative experiment. The red, solid rectangle marks the core phase. The induced glucose excursions are highlighted and enlarged in red, dotted rectangles. cap, capillary; CHO, carbohydrate.

### Data Analysis

The sensor readings, which were recorded with rates of one reading per 5 min (Guardian, Seven Plus) or one reading per 10 min (Navigator), were interpolated linearly after data exclusion if two adjacent sensor readings were recorded within 6.25 min for the Guardian and the Seven Plus and 12.5 min for the Navigator.

Valid capillary BG readings from Navigator-BG double measurements (one value each from Navigator-BG A and Navigator-BG B; discussed earlier) were averaged and used for pairing to interpolated CGM readings of the Navigator A/B, Guardian A/B, and Seven Plus A/B, respectively. All six sensors of each subject were analyzed separately.

For the complete experiments and for the core phase, only one capillary BG reading per hour during induced glucose excursions was included in the analysis to avoid over-representation of these excursions and to provide similar numbers of BG readings on each day.

Continuous glucose monitoring system performance was evaluated numerically and clinically. Numerical evaluation was performed for all subanalyses using mean absolute relative difference (MARD) and precision absolute relative difference (PARD). These parameters were calculated as averages of all of the experiment's MARD and PARD results ( $n = 24$  and  $n = 12$ , respectively) and as aggregated MARD and PARD of all individual absolute relative differences.

Mean absolute relative difference is the average of the absolute differences between paired capillary BG and interpolated CGM readings and is expressed as a percentage of the corresponding capillary BG readings: absolute relative deviation =  $|(\text{CGM} - \text{reference})| / \text{reference}$ .

Precision absolute relative difference was calculated in a similar fashion, but instead of sensor-to-BG differences, sensor-to-sensor differences were calculated, with the mean of the interpolated and the noninterpolated CGM readings replacing the capillary BG readings.<sup>8,9</sup>

In addition to these parameters, difference plots are provided to show the difference between single CGM readings and paired capillary BG readings. These plots also show the system accuracy limits of ISO 15197:2003,<sup>20</sup> an international standard for system accuracy evaluation of systems for self-monitoring of blood glucose (SMBG). Continuous glucose monitor sensors do not have to fulfill the criteria of ISO 15197; this analysis was performed to obtain results that could be compared with SMBG system accuracy results.

The statistical head-to-head comparisons between CGM systems were performed using the core phase data set because all sensors had identical glucose traces to follow. Calculations were performed in MATLAB R2012a v7.14 (The MathWorks Inc., Natick, MA). Statistical significance was tested using Kruskal–Wallis tests and Wilcoxon rank sum tests at an unmodified level of significance of  $\alpha = 0.05$ . Statistical analyses were performed using SAS<sup>®</sup> 9.2 (SAS Institute, Cary, NC).

## Results

Results for MARD and PARD for all subanalyses are displayed in **Tables 1** and **2**.

**Table 1.**

**Average Mean Absolute Relative Difference and Precision Absolute Relative Difference for the Continuous Glucose Monitoring Systems FreeStyle Navigator, DexCom Seven Plus, and Guardian REAL-Time<sup>a</sup>**

		MARD (%; n = 24)			PARD (%; n = 12)		
		Navigator	Guardian	Seven Plus	Navigator	Guardian	Seven Plus
Complete experiment (days 1–7)	<b>Overall</b>	<b>12.4 ± 3.6</b>	<b>16.4 ± 6.9</b>	<b>16.7 ± 3.8</b>	<b>10.1 ± 4.1</b>	<b>18.1 ± 6.5</b>	<b>15.4 ± 4.2</b>
	Day 1	14.1 ± 9.5	15.3 ± 6.2	19.2 ± 7.1	12.2 ± 4.0	16.3 ± 5.8	19.0 ± 7.7
	Day 2	12.8 ± 5.5	16.3 ± 9.0	18.1 ± 12.7	11.6 ± 6.5	16.3 ± 9.7	19.4 ± 14.2
	Day 3	13.2 ± 5.5	17.6 ± 11.7	15.3 ± 7.7	11.7 ± 8.1	21.8 ± 15.3	17.9 ± 10.7
	Day 4	9.6 ± 3.9	14.2 ± 6.8	14.5 ± 5.7	7.0 ± 4.1	16.5 ± 7.5	13.9 ± 5.8
	Day 5	13.3 ± 6.1	16.6 ± 8.6	18.0 ± 10.3	7.9 ± 4.6	17.6 ± 6.9	15.4 ± 8.2
	Day 6	NA	18.0 ± 16.9	13.3 ± 6.0	NA	20.0 ± 11.8	10.8 ± 3.7
	Day 7	NA	NA	18.1 ± 8.3	NA	NA	11.2 ± 5.7
	<70 mg/dl	22.6 ± 7.8	32.2 ± 18.1	38.3 ± 24.3	10.8 ± 4.8	22.3 ± 17.5	24.3 ± 7.5
	70–180 mg/dl	11.9 ± 3.5	15.4 ± 6.9	17.0 ± 4.0	10.0 ± 4.2	17.7 ± 5.7	14.9 ± 4.4
>180 mg/dl	11.0 ± 4.9	15.1 ± 8.1	11.7 ± 4.2	9.9 ± 4.1	18.5 ± 9.0	16.4 ± 5.1	
Core phase (Days 2–5)	<b>Overall</b>	<b>12.1 ± 3.8</b>	<b>16.2 ± 6.8</b>	<b>16.3 ± 4.9</b>	<b>9.6 ± 4.4</b>	<b>18.1 ± 7.4</b>	<b>16.7 ± 5.9</b>
	<70 mg/dl	24.6 ± 8.8	34.9 ± 26.1	32.7 ± 27.9	9.8 ± 5.9	20.2 ± 17.7	25.5 ± 15.0
	70–180 mg/dl	11.6 ± 4.0	15.1 ± 7.1	16.4 ± 5.4	9.5 ± 4.5	18.2 ± 7.2	16.3 ± 6.1
	>180 mg/dl	10.2 ± 4.5	14.6 ± 7.4	12.7 ± 5.2	9.3 ± 3.7	15.5 ± 7.8	17.0 ± 6.4
Induced glucose excursions	<b>Analysis versus capillary BG measured by the Navigator</b>	<b>13.4 ± 4.7</b>	<b>18.1 ± 6.5</b>	<b>17.4 ± 7.3</b>	9.5 ± 4.0	17.6 ± 6.3	15.4 ± 8.5
	Analysis versus venous YSI 2300	14.3 ± 5.4	19.2 ± 7.9	15.8 ± 6.2			
	Analysis versus compensated venous YSI 2300 <sup>b</sup>	13.0 ± 4.9	17.5 ± 6.5	18.7 ± 8.0			

<sup>a</sup> The results displayed are mean and standard deviation across all experiments. NA, not applicable.

<sup>b</sup> A bias of -14% was found for Navigator-BG versus capillary YSI 2300. Correction was implemented by multiplying venous YSI 2300 BG values by 0.86.

**Table 2.**  
**Aggregated Mean Absolute Relative Difference and Precision Absolute Relative Difference for the Continuous Glucose Monitoring Systems FreeStyle Navigator, DexCom Seven Plus, and Guardian REAL-Time<sup>a</sup>**

		MARD (%)			PARD (%)		
		Navigator	Guardian	Seven Plus	Navigator	Guardian	Seven Plus
Complete experiment (Days 1–7)	<b>Overall<sup>b</sup></b>	<b>12.6 ± 12.0</b> <b>n = 3021<sup>b</sup></b>	<b>16.7 ± 19.3</b> <b>n = 3304<sup>b</sup></b>	<b>16.7 ± 17.1</b> <b>n = 3761<sup>b</sup></b>	<b>10.1 ± 8.5</b> <b>n = 8427<sup>b</sup></b>	<b>18.1 ± 18.5</b> <b>n = 20157<sup>b</sup></b>	<b>15.3 ± 15.7</b> <b>n = 22908<sup>b</sup></b>
	Day 1	13.4 ± 12.1 n = 593	15.6 ± 15.2 n = 580	19.3 ± 15.6 n = 592	12.6 ± 9.5 n = 1404	16.3 ± 15.8 n = 3138	18.5 ± 19.1 n = 3019
	Day 2	13.0 ± 11.4 n = 580	16.5 ± 17.5 n = 547	18.2 ± 22.8 n = 530	11.7 ± 8.2 n = 1695	16.2 ± 15.5 n = 3361	19.5 ± 20.6 n = 3306
	Day 3	13.4 ± 15.0 n = 594	18.4 ± 22.1 n = 555	15.3 ± 17.7 n = 533	11.8 ± 9.7 n = 1701	21.8 ± 23.0 n = 3417	18.0 ± 17.3 n = 3341
	Day 4	9.6 ± 8.5 n = 640	14.4 ± 14.5 n = 550	14.5 ± 12.5 n = 583	7.0 ± 6.5 n = 1704	16.5 ± 15.7 n = 3365	13.9 ± 13.1 n = 3371
	Day 5	13.3 ± 11.2 n = 585	16.9 ± 16.4 n = 534	17.1 ± 15.3 n = 518	7.9 ± 7.0 n = 1706	17.5 ± 16.1 n = 3376	15.2 ± 13.4 n = 3337
	Day 6	NA	18.9 ± 27.9 n = 502	13.6 ± 13.3 n = 491	NA	20.3 ± 22.3 n = 3194	10.7 ± 9.3 n = 3290
	Day 7	NA	NA	18.3 ± 19.8 n = 513	NA	NA	11.1 ± 10.5 n = 3067
	<70 mg/dl	23.3 ± 21.3 n = 179	34.3 ± 41.0 n = 193	31.7 ± 31.6 n = 243	13.6 ± 11.4 n = 224	28.2 ± 23.7 n = 552	21.8 ± 18.3 n = 854
	70–180 mg/dl	11.9 ± 11.0 n = 2070	15.2 ± 16.9 n = 2220	16.8 ± 16.0 n = 2582	10.2 ± 8.7 n = 6488	17.5 ± 17.7 n = 15966	14.6 ± 15.5 n = 15995
	>180 mg/dl	11.9 ± 10.3 n = 772	16.7 ± 15.5 n = 891	12.3 ± 11.5 n = 936	9.5 ± 7.3 n = 1715	19.2 ± 20.4 n = 3639	16.2 ± 15.5 n = 6059
Core phase (Days 2–5)	<b>Overall</b>	<b>12.3 ± 11.8</b> <b>n = 2399</b>	<b>16.6 ± 17.9</b> <b>n = 2186</b>	<b>16.2 ± 17.4</b> <b>n = 2164</b>	<b>9.6 ± 8.2</b> <b>n = 6869</b>	<b>18.1 ± 18.1</b> <b>n = 13594</b>	<b>16.6 ± 16.5</b> <b>n = 13440</b>
	<70 mg/dl	26.7 ± 22.6 n = 138	34.9 ± 37.1 n = 129	31.4 ± 35.8 n = 128	13.9 ± 12.3 n = 163	29.4 ± 24.2 n = 383	22.9 ± 19.5 n = 420
	70–180 mg/dl	11.6 ± 10.7 n = 1624	14.9 ± 15.3 n = 1436	16.0 ± 16.0 n = 1424	9.6 ± 8.4 n = 5242	17.8 ± 18.3 n = 10683	15.8 ± 16.1 n = 9006
	>180 mg/dl	10.9 ± 8.8 n = 637	16.6 ± 15.0 n = 621	13.5 ± 12.5 n = 612	8.9 ± 6.6 n = 1464	17.3 ± 15.4 n = 2528	17.8 ± 17.0 n = 4014
Induced glucose excursions	<b>Analysis versus capillary BG measured by the Navigator</b>	<b>13.5 ± 11.1</b> <b>n = 996</b>	<b>18.1 ± 14.6</b> <b>n = 1002</b>	<b>17.4 ± 16.2</b> <b>n = 971</b>	9.5 ± 6.9 n = 742	17.6 ± 16.4 n = 1494	15.4 ± 14.4 n = 1462
	Analysis versus venous YSI 2300	14.3 ± 10.4 n = 952	19.2 ± 14.3 n = 958	15.9 ± 14.1 n = 927			
	Analysis versus compensated venous YSI 2300 <sup>c</sup>	13.1 ± 10.9 n = 952	17.5 ± 14.2 n = 958	18.6 ± 18.5 n = 927			

<sup>a</sup> The number of values (*n*) provided is the number of individual absolute relative differences that were included in the evaluation. Displayed are mean and standard deviation calculated from these individual differences. NA, not applicable.

<sup>b</sup> The number of individual values in the overall evaluation of complete experiments is slightly larger than the sum of numbers of individual values when analyzed day-wise. This deviation is caused by CGM readings recorded after the end of the fifth, sixth, or seventh 24 h period (for Navigator, Guardian, and Seven Plus, respectively) but before sensor removal. These additional values were not evaluated as a separate day, because the results would not have been comparable to the other day-wise results.

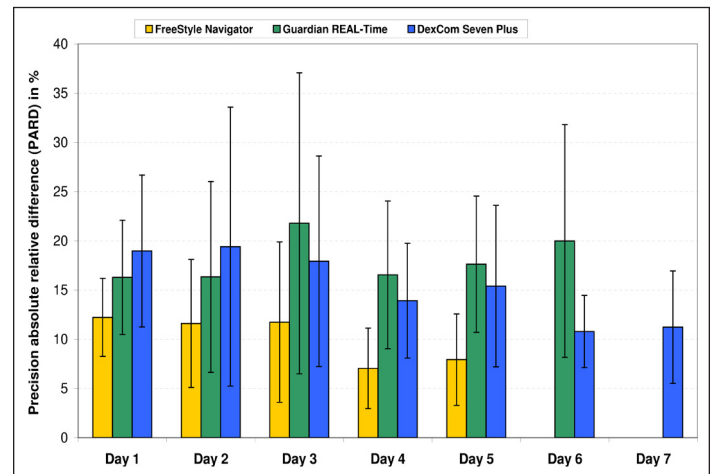
<sup>c</sup> A bias of -14% was found for Navigator-BG versus capillary YSI 2300. Correction was implemented by multiplying venous YSI 2300 BG values by 0.86.

## Complete Experiments

Complete experiments consist of 5 days (Navigator), 6 days (Guardian), or 7 days (Seven Plus) of CGM readings. In total, 1501/1520, 1652/1652, and 1871/1890 capillary BG readings were viable for pairing to interpolated CGM readings of the Navigator A/B, Guardian A/B, and Seven Plus A/B, respectively. The BG readings ranged from 36.5 to 438.5 mg/dl. Data reporting percentage, i.e., the time during which data were recorded divided by the manufacturer specified lifetime, was  $97.3\% \pm 5.2\%$  (mean  $\pm$  standard deviation) for the Navigator,  $98.9\% \pm 0.8\%$  for the Guardian, and  $95.9\% \pm 4.6\%$  for the Seven Plus.

The results of the complete experiments are very similar to those of the core phase. PARD seemed to decrease over the study days for the Navigator and the Seven Plus systems (see **Tables 1** and **2** and **Figure 2**).

Issues with starting the CGM systems occurred nine times: five times with the Navigator, three times with the Seven Plus, and one time with the Guardian. On one occasion, a Navigator insertion set did not contain a sensor. In two cases, there were issues with capillary BG measurements on the first day. In two other cases, the transmitters had to be reconnected to the sensor plates. For the Seven Plus, two sensors had to be replaced immediately after application because of sensor errors as indicated on the receiver display and, one time, the sensor had to be restarted as indicated on the receiver display. For the Guardian, one sensor experiment had to be restarted. All sensors with the exception of one Guardian sensor were used for the whole specified run time. Following the information on the CGM receiver, this Guardian sensor had to be stopped approximately 12 h before the planned end of sensor run time for no apparent reason.



**Figure 2.** Precision absolute relative difference over study days. For the Navigator and the Seven Plus, PARD decreases over time, i.e., precision increases.

The difference plots over the complete experiments show that all CGM systems experienced sensor inaccuracy (see **Figures 3A–C**). However, for the Navigator, the maximum differences between CGM readings and BG measurements are markedly smaller than for the Guardian and the Seven Plus, resulting in 80.2% (63.1% for reference glucose  $<75$  mg/dl, 81.9% for reference glucose  $\geq 75$  mg/dl) of readings within ISO 15197:2003 system accuracy criteria for the Navigator, 70.4% (58.1%, 71.6%) for the Guardian, and 71.7% (55.0%, 73.4%) for the Seven Plus.

## Core Phase

In the time frame of the four core days, on average, approximately 100 capillary BG measurements per subject were suitable for pairing with CGM data. This number differs between sensors because the phases during which no CGM readings were recorded varied from sensor to sensor.

The Navigator system achieved significantly lower MARD (12.1%) compared with the Guardian and the Seven Plus, which performed similarly (MARD 16.2% and 16.3%, respectively; see **Table 1**). A Kruskal–Wallis test showed a significant difference ( $p = .0009$ ) between all three CGM systems. Pairwise Wilcoxon rank sum tests showed no significant difference between the Guardian and the Seven Plus ( $p = .2818$ ), while significant differences were found between the Navigator and the Seven Plus ( $p = .00016$ ) and between the Navigator and the Guardian ( $p = .01$ ). PARD was markedly lower for the Navigator (9.6%) compared with the Guardian (18.1%) and the Seven Plus (16.7%).

In the hypoglycemic range ( $<70$  mg/dl), containing approximately 6% of paired measurement results, the Navigator again showed lower MARD and PARD than the Guardian and the Seven Plus.

With the EGA, 81.5%, 15.5%, 0.0%, 2.9%, and 0.0% of paired CGM BG readings were found within regions A, B, C, D, and E, respectively for the Navigator; 72.0%, 23.2%, 0.5%, 4.1%, and 0.2% for the Guardian; and 74.2%, 22.5%, 0.9%, 2.2%, and 0.2% for the Seven Plus. Regions A and B displayed clinically accurate or acceptable readings.

Figure 4 shows the percentage of CGM values that have a <5%, 10%, 20%, and 30% difference with respect to the corresponding reference value during the core phase and separated by glycemic range, with higher percentages representing a closer agreement between CGM and BG readings.

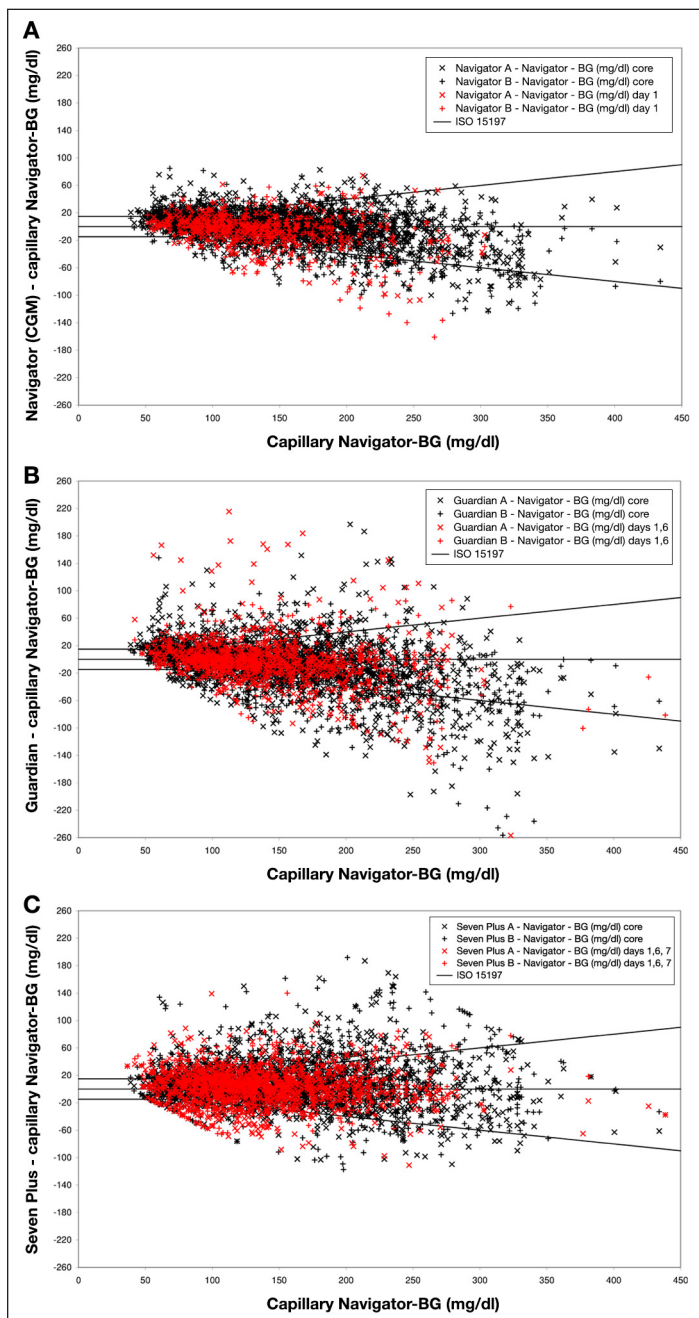


Figure 3. (A) Difference plot with ISO 15197:2003 system accuracy limits ( $\pm 15$  mg/dl for reference BG values  $< 75$  mg/dl and  $\pm 20\%$  for reference BG values  $\geq 75$  mg/dl) for Navigator CGM values compared with Navigator-BG. (B) Difference plot with ISO 15197:2003 system accuracy limits for Guardian CGM values compared with Navigator-BG. (C) Difference plot with ISO 15197:2003 system accuracy limits for Seven Plus CGM values compared with Navigator-BG.

### Induced Glucose Excursions

Capillary BG measurements ( $n = 108$ ) performed at sensor calibration times with the Navigator-BG were found to be lower than capillary BG measurements performed with the YSI 2300 device, with a bias of approximately -14%. Thus, when comparing CGM readings, which are calibrated using capillary Navigator-BG measurements, with YSI 2300 measurements in venous blood, this systematic error would be added to any deviations in BG between capillary and venous blood. For compensation of this bias, all YSI 2300 measurements in venous blood were corrected with the factor 0.86 in an additional evaluation (Tables 1 and 2).

All MARD results, i.e., comparison with capillary BG measurements, uncompensated venous BG measurements, and compensated venous BG measurements, as well as PARD results are displayed in Table 1.

During the induced glucose excursions, rates of glucose change ranged from -3.6 to +7.0 mg/dl/min for capillary BG and from -4.4 to +6.8 mg/dl/min for venous BG. Approximately 30% of rates were below -1 mg/dl/min and approximately 20% were above +1 mg/dl/min (Figure 5).

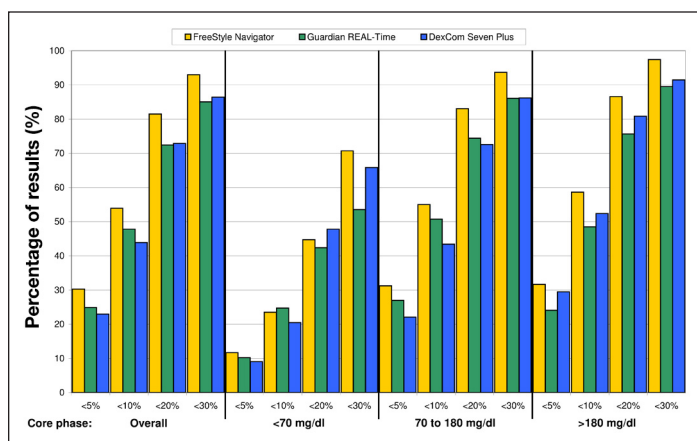
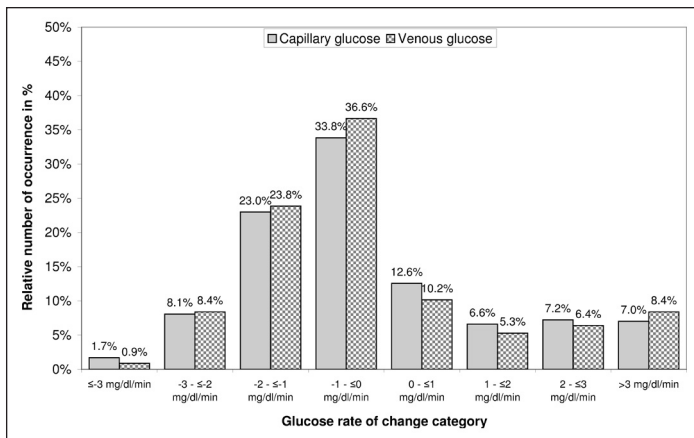


Figure 4. Cumulative percentage of CGM values differing <5%, 10%, 20%, and 30% from the capillary reference values. Results displayed for core phase (days 2 to 5): overall, <70, 70 to 180, and >180 mg/dl.





**Figure 5.** Relative number of occurrence of capillary and venous glucose rates of change (obtained from Navigator-BG and YSI 2300, respectively) during induced glucose excursions. Measurements were performed every 15 min.

## Discussion

In this study, three different CGM systems were worn by 12 subjects, and their performance was assessed in a clinical setting similar to daily life. The MARD and PARD results obtained from complete experiments are in agreement with previous reports about the three CGM systems,<sup>2-10,12,16-19</sup> with the Navigator achieving better performance than the other two systems. Results obtained from core phase evaluation, i.e., during days 2 to 5 when all sensors were used in parallel minus ~24 h run-in time, differ only slightly from results for complete experiments, thus the influence from run-in time and different duration of sensor usage was, in this study, less important than expected. Over the course of the study, MARD varied from day to day. With the Navigator and the Seven Plus, PARD seemed to slightly improve with sensor usage time.

The performance of all systems within the hypoglycemic range (<70 mg/dl) was markedly lower than at higher glucose concentrations, a known fact<sup>3,5,7,12,14,18</sup> that is still dissatisfying for CGM users. While the Guardian achieved a slightly lower MARD than the Seven Plus, its PARD was slightly higher. The Navigator, again, showed a lower MARD and PARD than both the Guardian and the Seven Plus. Similar results were found in the euglycemic and hyperglycemic ranges (70 to 180 and >180 mg/dl, respectively). While PARD for the Navigator was similar in all three glucose ranges, MARD was worse in the hypoglycemic range. This suggests that the decreased performance at low glucose values is independent of individual sensors.

In the core phase, 95% to 97% of paired CGM reference readings were found in clinically accurate and acceptable zones A and B of the EGA. This result is in agreement with other studies incorporating EGA analysis.<sup>3,8,9,12,13,16-19,21,22</sup>

Only 70% to 80% of paired readings were found to be within the system accuracy limits of ISO 15197:2003.<sup>20</sup> According to ISO 15197, 95% of device readings must be within  $\pm 15$  mg/dl or  $\pm 20\%$  of the paired reference readings below or above 75 mg/dl. Thus, CGM systems still have lower accuracy than devices for SMBG; however, when an SMBG device is used for calibration, as in the study presented here, its accuracy (precision and trueness) directly affects the CGM's accuracy.<sup>10</sup>

The study setting was designed as a compromise between prerequisites of POCT05-A, a Clinical and Laboratory Standards Institute guideline providing recommendations for evaluating performance of CGM systems, and parallel use of sensors in an in-house setting similar to daily life using induced excursions at the beginning of sensor use and calibration phase and at a later time in sensor use. Only one study claimed to follow POCT05-A recommendations in design and evaluation.<sup>13</sup> The study presented here is, to the best knowledge of the authors, the first to use two sensors of three different CGM systems in parallel for the respective entire lifetime of the sensor, thus allowing the evaluation of accuracy (with regard to BG measurements) as well as sensor-to-sensor precision between different sensors of one type of CGM system.<sup>23</sup> In other studies in which more than one sensor was used per subject, only one type of CGM system was used,<sup>3,4,6,8-11,15</sup> and in studies that compared multiple CGM systems, only one sensor per subject was used.<sup>2,7,12,19</sup> In this study, both aspects were combined, thus allowing an evaluation in which groups of six sensors (i.e., two sensors of each of three CGM systems in one individual subject) follow identical glycemc excursions. The choice of participants may have a marked influence on the analyzed parameter,<sup>11</sup> this setting, however, allows very high comparability between the CGM systems for both MARD and PARD because data are generated for all systems and all sensors at the same time in the same subject. To further improve comparability between the single experiments in one subject, the core phase was chosen as the primary evaluation because possible influences from

different run-in times and different glucose fluctuations were minimized. Most analyses, unless specified otherwise, were performed with capillary Navigator-BG readings as reference. Some studies in which CGM performance was evaluated also used the same devices for calibration and for reference readings,<sup>5,11,14,18</sup> while in some studies, the same devices were used only in parts of the study or depending on the parameter analyzed.<sup>4,6,9,10,13,16,17</sup> There also are studies in which different devices were used during the entire study.<sup>2,3,7,8,12</sup> With the calibration BG readings being obtained from the same system as the reference readings, the influence of systematic differences in trueness between the reference device and calibration device were minimized.

One limitation of this study was that the two sensors of each type of CGM system were calibrated using different BG values, which could lead to higher sensor-to-sensor differences (PARD) in comparison with calibration using identical values. However, with the Navigator system not allowing input of externally obtained BG values, it increases the comparability of the three systems in this study as compared with calibrating the other systems with identical values.

The Seven Plus and the Navigator CGM systems, which were current CGM systems at the time the study was performed, are no longer on the market. The Seven Plus's successor, the DexCom G4 Platinum (DexCom, San Diego, CA) is reported to be more accurate,<sup>24</sup> and the Navigator was replaced by the FreeStyle Navigator II (Abbott Diabetes Care, Alameda, CA). However, it is unclear if its performance has changed. It should also be noted that the Enlite sensors are also used with the Paradigm Veo Pump (Medtronic MiniMed, Northridge, CA), and that the Veo's calibration algorithm is reported to be different from that of the Guardian REAL-Time.<sup>25</sup> However, the Guardian is a stand-alone CGM system, i.e., it can be used by people with diabetes who do not wear an insulin pump. The Guardian is also used with Enlite sensors; at least in Germany, Enlite sensors are sold to be used with the Guardian.

During the course of this study, a considerable negative systematic glucose measurement error ("bias") of the FreeStyle test strips used for capillary BG measurements was observed when calibration measurements were compared with capillary BG measurements performed with YSI 2300 STAT Plus. As stated in Research Design and Methods, all Navigator-BG meters were tested with control solution before and after use. Considering that this bias would likely affect the analyses versus venous BG measurements, an additional set of analyses was performed with venous BG measurements corrected by a fixed factor to estimate its effect on MARD. This approach may be an oversimplification because the bias may not be identical across the whole range of BG measurements. The correction of venous BG readings was found to have only a small effect on the systems' MARD (see **Table 1**), and interestingly, the MARD improved only for two of the three systems used.

In one published study, Luijf and coauthors<sup>17</sup> concluded that CGM accuracy in a clinical research center and CGM accuracy at home differ due to higher measurement frequency at clinical research centers. The authors claim that the effect of changing their reference methods was minimized because their SMBG meter was factory calibrated with YSI values.<sup>17</sup> However, the Navigator-BG device used in our study was also factory calibrated against YSI 2300, yet a considerable bias was found.

One of this study's goals was to implement the recommendations of POCT05-A concerning testing of point and trend accuracy. POCT05-A suggests many analyses and evaluations, making the study design and evaluation very complex. This study focused on point and trend accuracy to provide results of parameters that are often reported in the context of performance of CGM systems. Glucose excursions with high-frequency data collection were induced on two days to achieve this goal. In this study, one excursion was at the beginning of sensor lifetime and calibration interval, whereas the other excursion was performed at a later time without taking sensor lifetime and calibration cycle into account, thus deviating from the testing protocol POCT05-A, which requires more segments with high-frequency data collection. Most recommendations for this testing protocol, such as achieving rates of glucose change exceeding 3 mg/dl/min, obtaining pairs of sensor reference readings every 15 min, and having adequate numbers of measurements in the glucose ranges <70, 70 to 180, and >180 mg/dl, were implemented, as well as numerical and clinical assessment of point and trend accuracy.

In conclusion, the CGM systems tested in this study showed different performance, with the Navigator achieving lower (i.e., better) MARD and lower PARD results than Seven Plus and the Guardian. This difference was observed

when comparing the systems under identical conditions (core phase), when comparing each system's complete run time, and when only comparing phases of rapidly changing glucose. For the Navigator and the Seven Plus, sensor-to-sensor precision improved during the course of the study, while, for MARD, no improvement was found over time. One major issue of all CGM systems was the poor performance in the hypoglycemic range, which was markedly poorer than in the other glucose concentration ranges.

---

**Funding:**

This study was funded by Roche Diagnostics GmbH, Mannheim, Germany.

---

**Disclosures:**

Guido Freckmann is general manager of the Institute for Diabetes Technology Research and Development GmbH at Ulm University, which performs studies evaluating BG meters and medical devices for diabetes therapy on behalf of various companies; he/the Institute for Diabetes Technology Research and Development GmbH at Ulm University has received speakers' honoraria from Abbott, Bayer, Berlin Chemie, Roche Diagnostics, Ypsomed, and Sanofi. Hans-Martin Klötzer was contracted by Roche Diagnostics, Germany, to evaluate data for this study.

---

**Acknowledgments:**

We thank Prof. Dr. Lutz Heinemann, Dr. Volker Lodwig, and Prof. Dr. Luigi del Re for their valuable input. The authors thank Dr. Elisabeth Ramstetter for her valuable contributions during the study. Scientific writing was supported by Roche Diagnostics GmbH, Mannheim, Germany. Parts of these data were presented at the 72nd Scientific Sessions of the American Diabetes Association (Philadelphia, PA, June 8–12, 2012), the 48th European Association for the Study of Diabetes Annual Meeting (Berlin, Germany, October 1–5, 2012), and the German Diabetes Association Annual Meeting (Leipzig, Germany, May 9–11, 2013).

---

**References:**

1. Clinical and Laboratory Standards Institute. Performance metrics for continuous interstitial glucose monitoring; approved guideline. [http://shopping.netsuite.com/c.1253739/site/Sample\\_pdf/POCT05A\\_sample.pdf](http://shopping.netsuite.com/c.1253739/site/Sample_pdf/POCT05A_sample.pdf). Accessed June 14, 2013.
2. Clarke WL, Anderson S, Farhy L, Breton M, Gonder-Frederick L, Cox D, Kovatchev B. Evaluating the clinical accuracy of two continuous glucose sensors using continuous glucose-error grid analysis. *Diabetes Care*. 2005;28(10):2412–7.
3. Weinstein RL, Schwartz SL, Brazg RL, Bugler JR, Peyser TA, McGarraugh GV. Accuracy of the 5-day FreeStyle Navigator Continuous Glucose Monitoring System: comparison with frequent laboratory reference measurements. *Diabetes Care*. 2007;30(5):1125–30.
4. Wilson DM, Beck RW, Tamborlane WV, Dontchev MJ, Kollman C, Chase P, Fox LA, Ruedy KJ, Tsalikian E, Weinzimer SA; DirecNet Study Group. The accuracy of the FreeStyle Navigator continuous glucose monitoring system in children with type 1 diabetes. *Diabetes Care*. 2007;30(1):59–64.
5. Mastrototaro J, Shin J, Marcus A, Sulur G; STAR 1 Clinical Trial Investigators. The accuracy and efficacy of real-time continuous glucose monitoring sensor in patients with type 1 diabetes. *Diabetes Technol Ther*. 2008;10(5):385–90.
6. Diabetes Research in Children Network (DirecNet) Study Group. The accuracy of the Guardian RT continuous glucose monitor in children with type 1 diabetes. *Diabetes Technol Ther*. 2008;10(4):266–72.
7. Kovatchev B, Anderson S, Heinemann L, Clarke W. Comparison of the numerical and clinical accuracy of four continuous glucose monitors. *Diabetes Care*. 2008;31(6):1160–4.
8. 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(5):1146–54.
9. Bailey T, Zisser H, Chang A. New features and performance of a next-generation SEVEN-day continuous glucose monitoring system with short lag time. *Diabetes Technol Ther*. 2009;11(12):749–55.
10. 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(11):689–95.
11. Nielsen JK, Freckmann G, Kapitza C, Ocvirk G, Koelker KH, Kamecke U, Gillen R, Amann-Zalan I, Jendrike N, Christiansen JS, Koschinsky T, Heinemann L. Glucose monitoring by microdialysis: performance in a multicentre study. *Diabet Med*. 2009;26(7):714–21.
12. Garg SK, Smith J, Beatson C, Lopez-Baca B, Voelmle M, Gottlieb PA. Comparison of accuracy and safety of the SEVEN and the Navigator continuous glucose monitoring systems. *Diabetes Technol Ther*. 2009;11(2):65–72.
13. Valgimigli F, Lucarelli F, Scuffi C, Morandi S, Sposato I. Evaluating the clinical accuracy of GlucoMen(R)Day: a novel microdialysis-based continuous glucose monitor. *J Diabetes Sci Technol*. 2010;4(5):1182–92.
14. Morrow L, Hompesch M, Tideman AM, Matson J, Dunne N, Pardo S, Parkes JL, Schachner HC, Simmons DA. Evaluation of a novel continuous glucose measurement device in patients with diabetes mellitus across the glycaemic range. *J Diabetes Sci Technol*. 2011;5(4):853–9.
15. Adolfsson P, Ornhagen H, Eriksson BM, Cooper K, Jendle J. Continuous glucose monitoring—a study of the Enlite sensor during hypo- and hyperbaric conditions. *Diabetes Technol Ther*. 2012;14(6):527–32.

16. Facchinetti A, Sparacino G, Guerra S, Luijf YM, DeVries JH, Mader JK, Ellmerer M, Benesch C, Heinemann L, Bruttomesso D, Avogaro A, Cobelli C; AP@home Consortium. Real-time improvement of continuous glucose monitoring accuracy: the smart sensor concept. *Diabetes Care*. 2013;36(4):793–800.
17. Luijf YM, Avogaro A, Benesch C, Bruttomesso D, Cobelli C, Ellmerer M, Heinemann L, Mader JK, DeVries JH; AP@home consortium. Continuous glucose monitoring accuracy results vary between assessment at home and assessment at the clinical research center. *J Diabetes Sci Technol*. 2012;6(5):1103–6.
18. Zijlstra E, Heise T, Nosek L, Heinemann L, Heckermann S. Continuous glucose monitoring: quality of hypoglycaemia detection. *Diabetes Obes Metab*. 2013;15(2):130–5.
19. Damiano ER, El-Khatib FH, Zheng H, Nathan DM, Russell SJ. A comparative effectiveness analysis of three continuous glucose monitors. *Diabetes Care*. 2013;36(2):251–9.
20. International Organization for Standardization. ISO 15197:2003. In vitro diagnostic test systems -- requirements for blood-glucose monitoring systems for self-testing in managing diabetes mellitus.
21. Garg S, Zisser H, Schwartz S, Bailey T, Kaplan R, Ellis S, Jovanovic L. Improvement in glycemic excursions with a transcutaneous, real-time continuous glucose sensor: a randomized controlled trial. *Diabetes Care*. 2006;29(1):44–50.
22. Garg S, Jovanovic L. Relationship of fasting and hourly blood glucose levels to HbA1c values: safety, accuracy, and improvements in glucose profiles obtained using a 7-day continuous glucose sensor. *Diabetes Care*. 2006;29(12):2644–9.
23. Obermaier K, Schmelzeisen-Redeker G, Schoemaker M, Kloetzer HM, Kirchsteiger H, Eikmeier H. Performance evaluations of continuous glucose monitoring systems: precision absolute relative difference is part of the assessment. *J Diabetes Sci Technol*. 2013;7(4):824–32.
24. Pleus S, Schmid C, Link M, Zschornack E, Klötzer HM, Haug C, Freckmann G. Performance evaluation of a continuous glucose monitoring system under conditions similar to daily life. *J Diabetes Sci Technol*. 2013;7(4):833–41.
25. Keenan DB, Mastrototaro JJ, Zisser H, Cooper KA, Raghavendhar G, Lee SW, Yusi J, Bailey TS, Brazg RL, Shah RV. Accuracy of the Enlite 6-day glucose sensor with guardian and Veo calibration algorithms. *Diabetes Technol Ther*. 2012;14(3):225–31.