A Critical Appraisal of the Continuous Glucose–Error Grid Analysis

IRIS M. WENTHOLT, MD[1]
JOOST B. HOEKSTRA, MD, PHD[1]
J. HANS DEVRIES, MD, PHD[1]

OBJECTIVE — There is no consensus on how to optimally assess the accuracy of continuous glucose sensors. We examined the continuous glucose–error grid analysis (CG-EGA) and compared it with classical accuracy assessment methods, using data from a previously reported study comparing two different continuous glucose sensors in type 1 diabetic patients.

RESEARCH DESIGN AND METHODS — Drift, delay, mean absolute difference (MAD), sensitivity, and specificity for detecting hypo- and hyperglycemia were calculated, and a Clarke error grid and a CG-EGA were constructed for both sensors, also including an examination of the influence of choosing different time intervals for paired sensor and reference glucose values.

RESULTS — For sensor II, there was a delay between blood glucose and sensed glucose (7.1 min, $P < 0.001$). Sensor II was more accurate than sensor I during hypo- and hyperglycemia (e.g., smaller MAD, $P = 0.011$ and $P = 0.024$, respectively; better sensitivity for detecting hypoglycemia, $P = 0.018$). Correction for the 7-min delay improved sensor II MAD with 2.2% in every range. In contrast, CG-EGA did not reveal a difference in accuracy between the sensors. Paradoxically, CG-EGA results for sensor II deteriorated when corrected for the delay. CG-EGA calculated with shorter time intervals resulted in worsening accuracy for both sensors.

CONCLUSIONS — CG-EGA did not detect differences in accuracy whereas conventional methods did. CG-EGA is time demanding; results are hard to interpret and seem to vary with chosen time intervals. At present, CG-EGA does not contribute to a combination of various established assessment methods.

Diabetes Care **29:1805–1811, 2006**

Continuous glucose sensors are promising devices for the treatment of diabetes. They provide much more information on daytime and nighttime glucose patterns than glucose spot measurements. In recent studies, an association between continuous glucose monitoring and improved glycemic control has been reported (1,2). During the last 5 years, various continuous glucose sensors have been developed (3). To receive Conformité Européenne or U.S. Food and Drug Administration approval, but, more importantly, to allow clinical action upon (online) sensor readings, continuous glucose sensors must exhibit a certain reliability. However, both defining and assessing sensor accuracy appear to be difficult (4). Although many methods have been proposed, consensus on the ideal accuracy assessment method or combination of methods has not been achieved yet.

In 1987, the Clarke error grid analysis (EGA), designed by and named for Dr. William Clarke, was first reported (5). This method was innovative because it took take into account not only the difference between the system-generated and reference blood glucose values but also the clinical significance of this difference (5). Like other established accuracy assessment methods, such as correlation, linear regression, and mean absolute difference (MAD), the Clarke EGA evaluates the correspondence and discrepancy between blood glucose values and sensor readings at isolated static points in time, while neglecting the relationship of data points in time. In continuous glucose monitoring systems, data at a given point are related to those nearby. In other words, the rate of change in glucose per se is not assessed by the above-mentioned conventional measures.

In 2004, a novel accuracy assessment method called continuous glucose–error grid analysis (CG-EGA) was proposed by Clarke's group, which had been specifically designed for evaluation of continuous glucose sensors. With this new approach, the study group wanted to "initiate a debate on the current and important problem of accuracy assessment" rather than "to set a method in stone" (6). CG-EGA takes into account the interdependency of successive data by combining point accuracy with rate accuracy (6,7). In a recent study, the clinical accuracy of two continuous glucose sensors were analyzed using the CG-EGA. Other established accuracy assessment methods, such as EGA, MAD, correlation, and sensitivity and specificity for hypo-, normo-, and hyperglycemia, were not reported, which makes it hard to assess the exact contribution of the novel method (8).

The aim of the current study was to explore the principles underlying the CG-EGA and to compare the CG-EGA with classical accuracy assessment methods using data obtained during an earlier reported direct comparison of two subcutaneous continuous glucose sensors in type 1 diabetic patients (9).

RESEARCH DESIGN AND METHODS

— A full description of the study protocol has been reported before; a summary is given below (9). The study was approved by the local ethics committee, and participants gave written informed consent. Thirteen type 1 diabetic patients (9 men) were enrolled. Mean ± SD HbA_{1c} (A1C) was 8.2 ± 0.8%, BMI

From the Academic Medical Center, Department of Internal Medicine, Amsterdam, the Netherlands.
Address correspondence and reprint requests to I.M. Wentholt, MD, Department of Internal Medicine, Academic Medical Center, P.O. Box 22660, 1100 AD Amsterdam, Netherlands. E-mail: i.m.wentholt@amc.uva.nl.
Received for publication 12 January 2006 and accepted in revised form 19 April 2006.
Abbreviations: CG-EGA, continuous glucose–error grid analysis; EGA, error grid analysis; MAD, mean absolute difference; P-EGA, point–error grid analysis; R-EGA, rate–error grid analysis.
A table elsewhere in this issue shows conventional and Système International (SI) units and conversion factors for many substances.
DOI: 10.2337/dc06-0079

23.8 ± 3.0 kg/m^2, age 34.3 ± 10.7 years, and diabetes duration 17.2 ± 9.5 years. Other than one patient who had a history of panretinal coagulation, participants were free from diabetes complications.

Two commercially available glucose sensors were inserted subcutaneously in the abdominal area: a needle-type sensor (sensor I) and a microdialysis-based sensor (sensor II). Both sensors were attached strictly according to the manufacturer's instructions. Sensor I reports glucose values every 5 min and sensor II every 3 min (10,11). Approximately 12 h after insertion of either sensor, patients were admitted to the hospital for 1 night. Blood sampling twice hourly for glucose determination using the hexokinase/glucose-6-phosphate dehydrogenase method (Roche Hitachi) started at 10:00 P.M. and continued until 8:00 A.M. the next morning.

The next morning at 8:00 A.M., an increase in glucose was induced by postponing the usual rapid-acting insulin injection until 30 min after breakfast. A decrease in glucose was induced by an additional amount of rapid-acting insulin (on average 5 units added to the usual morning dose) injected subcutaneously.

Starting 45 min after breakfast, venous blood was frequently sampled (once every minute) for 30 min to record the glucose peak. At 105 min after breakfast (75 min after insulin injection), frequent sampling was again done for 30 min, this time to record the glucose nadir (Fig. 1 in ref. 9).

Data analysis

The following accuracy assessment methods were used. Data were pooled, and a MAD was calculated for each sensor, as well as sensitivity and specificity for detecting hypo- and hyperglycemia. In addition, a Clarke error grid plot was constructed. CG-EGA was performed in accordance with the original report, as described below.

CG-EGA

For each sensor, readings were paired with blood glucose values over constant time intervals. If concomitant blood glucose values were not available, interpolated blood glucose values were used. We chose identical time intervals of 15 min for both sensors as used in the original CG-EGA report (6). The in-between blood glucose readings were discarded for analysis. For the paired values, both point and rate accuracies were calculated and plotted in the respective point–error grid analysis (P-EGA) and rate–error grid analysis (R-EGA) plots. Point accuracy reflects the difference between two paired samples at one point in time ($SG_{t1} - BG_{t1}$, where SG is sensor glucose measurement

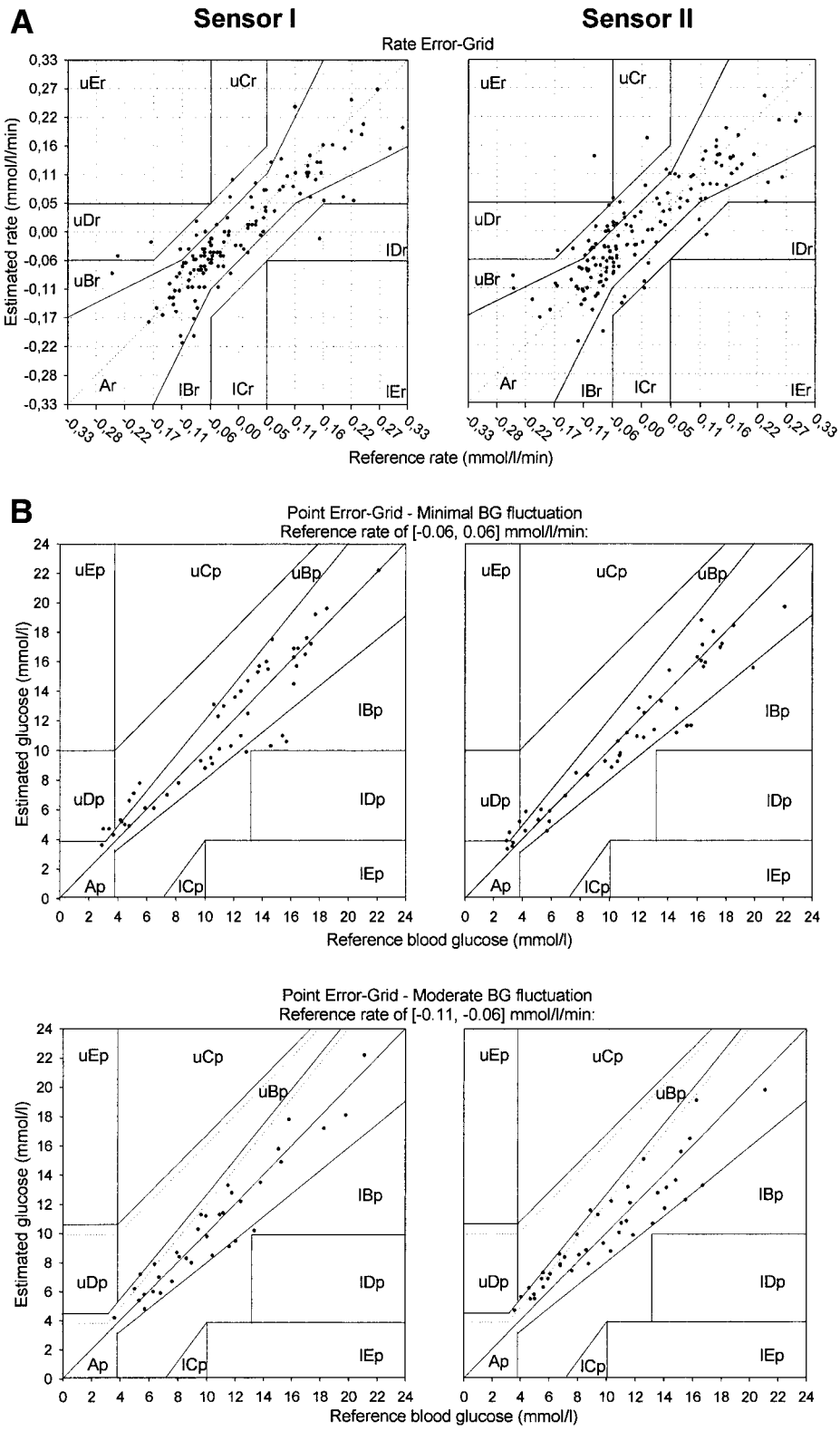

Figure 1—*Glucose values plotted in the R-EGA grids* (A) *and scatter plots of the glucose point values superimposed over the P-EGA grids* (B). *The CG-EGA plots points on grids with dynamically adjusted boundaries. l, lower; p, point; r, rate; u, upper.*

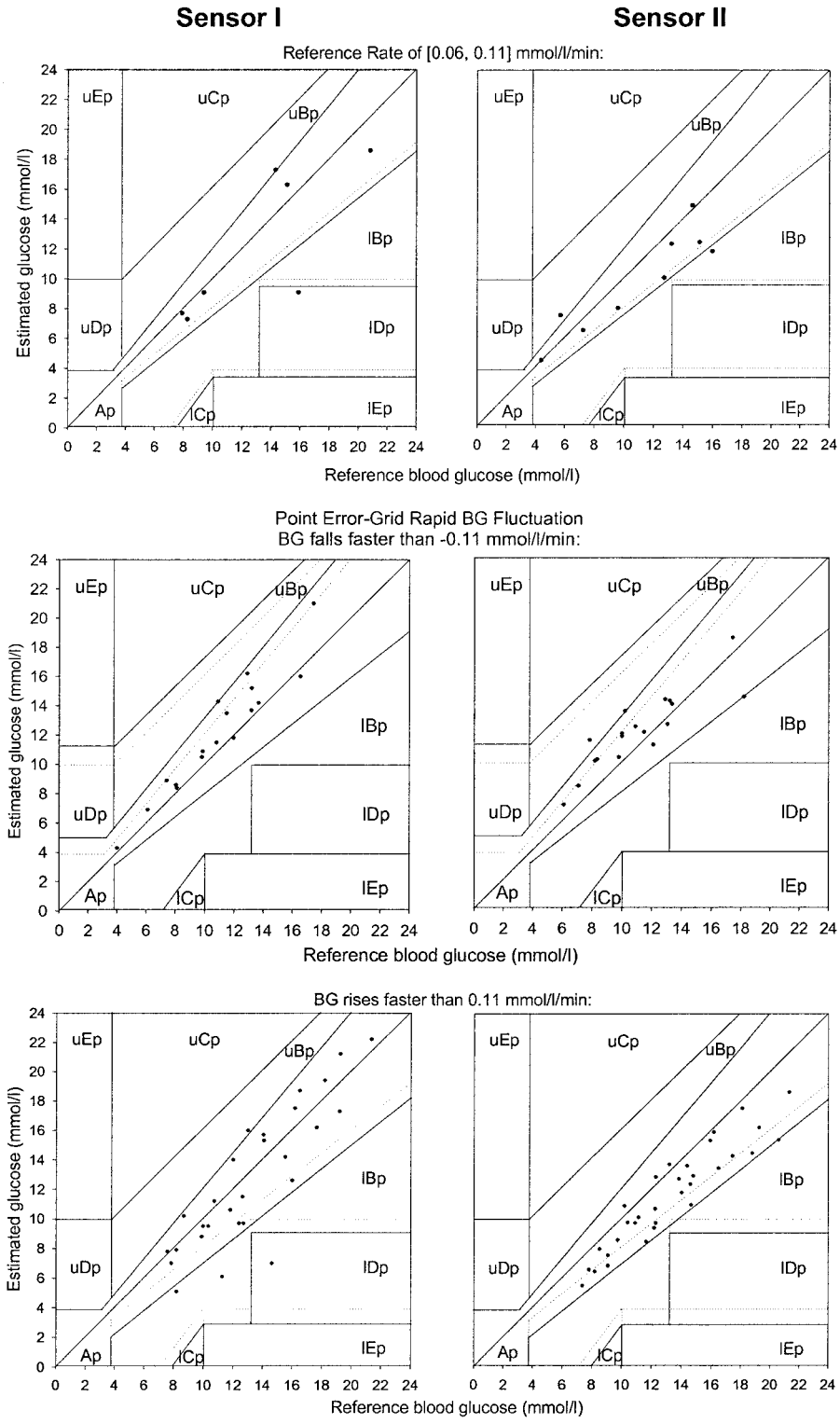

Fig. 1—*Continued.*

and BG is blood glucose measurement). Rate is defined as the difference in glucose level between two consecutive measurements, divided by the time interval between those measurements [(glucose$_{t2}$ − glucose$_{t1}$)/Δt]. Rate accuracy reflects the difference in rate between sensor and corresponding blood glucose measurements [(SG$_{t2}$ − SG$_{t1}$) − (BG$_{t2}$ − BG$_{t1}$)/Δt] and indicates how well the sensor is capable of following direction and tempo of the blood glucose changes.

The rates of sensor readings are plotted against blood glucose rates in the R-EGA, consisting of zones A–E, which are clinically comparable to the Clarke error grid zones: points in zone A represent rates of change in sensor readings similar to blood glucose rates, points in zones B–D represent rates of change with increasing inaccuracy, and points in zone E represent readings with a rate opposite to the reference rate.

The P-EGA is also similar to the Clarke error grid, except that it allows for a possible shift of the upper and lower boundaries of zones A, B, and D, in proportion to the rate (6). To what extent zones A, B, and D are expanded is determined by multiplying the mean rate by 7 min. This period of time originates from the assumption of a general 7-min delay between interstitial (sensor) and blood glucose. Rate is divided into five categories: minimal fluctuation of blood glucose (-0.06 to 0.06 mmol$\cdot$l$^{-1}\cdot$min^{-1}), moderate rise and fall (± 0.11 to ± 0.06 mmol/l$\cdot$l$^{-1}\cdot$min^{-1}), and rise and fall faster than ± 0.11 mmol/l$\cdot$l$^{-1}\cdot$min^{-1}.

Finally, the results of the separately calculated rate and point accuracy are combined in one matrix and presented per blood glucose range: hypo-, normo-, and hyperglycemia. Sensor readings are considered clinically accurate when they fall into the A and B zones in both the point and rate error grids. Clinically benign errors are those with acceptable point accuracy (A or B zones in the point error grid) but significant errors in rate accuracy (C, D, or E zones in the rate error grid), which are unlikely to have clinically negative therapeutic consequences (6).

To evaluate the influence of the length of time interval chosen on the CG-EGA results, we repeated the analysis with 5- and 3-min intervals for sensors I and II, respectively, based on the reporting frequency of each sensor.

We also used the results of a novel analysis method, reported in our earlier study (9). In this so-called combined curve fitting, the sensor-reported glucose values were fitted into curves using least-squares linear regression, while they were assumed to have the same shape as the blood glucose curve. This assumption allowed for possible vertical and horizontal shifts indicating drift and delay, respectively (9). Results of this analysis showed no significant drift (vertical shift) for either sensor. The horizontal shift for sensor II revealed a significant mean ± SD delay between blood glucose and sensed glucose of 7.1 ± 5.5 min ($P < 0.001$, one-sample t test); for sensor I there was

Table 1—*CG-EGA rate accuracy*

	15 min		5 min	3 min
Zone	Sensor I	Sensor II	Sensor I	Sensor II
A	82.4	74	71	52.7
B	14.7	21.3	18.5	27
C	0.7	2	5	10.2
D	2.2	2	2.8	5
E	0	0.7	2.6	5.1

Data are percent. Rate accuracy was calculated over a time interval of 15 vs. 5 min for sensor I and 15 vs. 3 min for sensor II.

no delay (9). Subsequently for sensor II, MAD was recalculated after correction for the delay. Also, CG-EGA sensitivity to the delay was examined. The readings of sensor II were paired with blood glucose measurements given 6 min earlier. This time was necessarily a multiplication of the sensor's reporting frequency (once every 3 min); therefore, 6 min serves as a proxy of the reported 7-min delay.

RESULTS — MADs were 15.0 ± 12.2% (735 paired samples) for sensor I and 13.6 ± 10.2% (1,156 paired samples) for sensor II, the latter being significantly more accurate ($P = 0.013$, Student's *t* test). After correction for the 7-min delay (9), MAD for sensor II further improved to 11.7 ± 9.8% ($P < 0.0001$ vs. both sensor II MAD without correction and sensor I MAD).

MADs calculated separately over the hypoglycemic range (≤3.9 mmol/l) were 24.9 and 17.5% for sensors I and II, respectively; over the normoglycemic range (3.9 < blood glucose ≤ 10.0 mmol/l) MAD was 15.5% for both sensors; and over the hyperglycemic range (>10.0 mmol/l), MADs were 12.5 and 11.1% for sensors I and II, respectively, with sensor II being significantly more accurate than sensor I during hypo- and hyperglycemia ($P = 0.011$ and $P = 0.024$, respectively, Student's *t* test). Correction for the 7-min delay of sensor II resulted in a MAD improvement of 2.2% on average in every range with *P* values varying from 0.001 during hypo- and normoglycemia to <0.0001 during hyperglycemia.

Sensitivity for detecting hypoglycemia was significantly better for sensor II (75.0%) than for sensor I (55.9%, $P = 0.018$, Pearson χ^2). Specificity was similar with values of 96.3 and 97.4% for sensors I and II, respectively ($P = 0.153$). Neither sensitivity (84.7 and 90.1% for sensors I and II, respectively, $P = 0.207$) nor specificity for detecting hyperglycemia (92.7 and 90.4%, $P = 0.274$) differed significantly between the sensors.

Clarke EGA indicated for sensor I that 95.9% of the readings fell in the clinically acceptable zones A or B and 4.1% fell in zone D; no readings ended up in zones C or E. For sensor II, 98.3% of the readings fell in zones A or B, which was significantly more compared with sensor I ($P < 0.0001$, Pearson χ^2); readings fell 0.1% in zone C and 1.5% in zone D, the latter percentage being significantly lower than for sensor I ($P < 0.0001$). There were no readings in zone E.

The sensor rates are plotted against the blood glucose rates in the rate-error grid (Fig. 1*A*), and rate accuracy is summarized in Table 1. For sensors I and II, 97.1 and 95.3% of the rates, respectively, ended up in the clinically acceptable zones (A or B; $P = 0.449$, Pearson χ^2). The sensor and blood glucose readings are plotted in the point error grid (Fig. 1*B*), and point accuracy is listed in Table 2. Point accuracies of sensors I and II were also similar with 97.2 and 99.0% of the sensor readings, respectively, in zones A or B ($P = 0.607$). Point and rate accuracy are combined and reported per glucose range in the matrix in Fig. 2*A* and *B*. A simplified version of the matrix (Fig. 2*C*) reveals that both sensors exhibit similar accuracy in every glucose range, with only 60 and 57.2% of the readings of sensors I and II, respectively, being clinically acceptable, i.e., accurate readings or benign errors, during hypoglycemia ($P = 0.921$), 100% during normoglycemia, and 97.6 and 97.8% during hyperglycemia ($P = 0.926$). Thus, the CG-EGA did not reveal a difference in accuracy between the sensors.

As mentioned above, the MAD for sensor II improved after correction for its delay. Paradoxically, according to the CG-EGA, accuracy of sensor II after correction for the delay with 6 min seemed to deteriorate rather than to improve during hypo- and normoglycemia with percentages of clinically acceptable readings of 51.5% (vs. 57.2% without correction for delay) and 96.8% (vs. 100%), respectively ($P = 0.782$ and $P = 0.194$, Pearson χ^2). During hyperglycemia, accuracy was significantly worse: 89.3% (vs. 97.8%) of the corrected sensor II readings were clinically acceptable ($P = 0.011$).

Reducing the time intervals from 15 to 5 and 3 min resulted in a significantly worse rate error grid with 89.5% of the rates falling in zones A or B (compared with 97.1% with 15-min intervals, $P = 0.006$, Pearson χ^2) for sensor I and 79.7% for sensor II (compared with 95.3%, $P < 0.0001$) (Table 1), and with increases of 2.6% ($P = 0.057$) and 4.4% ($P = 0.015$) of values falling in the clinically danger-

Table 2—*CG-EGA point accuracy*

	15 min		5 min	3 min	15 min	5 min	
Zone	Sensor I	Sensor II	Sensor I	Sensor II	Ref. 6	Ref. 8	Ref. 8
A	82.2	87.2	80.5	84.9	75	71	69.7
B	15	11.8	16.3	13.8	23.7	26.9	24
C	0	0.1	0	0.1	0.1	0	0
D	2.8	1	3.2	1.2	1.2	2.1	6.3
E	0	0	0	0	0	0	0

Data are percent. CG-EGA point accuracy was calculated over 15- and 5-min time intervals for sensor I and 15- and 3-min time intervals for sensor II in the first four columns; for comparison, the last three columns represent CG-EGA point accuracy of other sensors in two previous studies. In all studies, hardly any reading ended up in zone C of the point error grid.

A

Rate Error-Grid Zones sensor I	Point Error-Grid Zones **sensor I**										
	Hypoglycemia BG ≤ 3.89 mmol/l			Normoglycemia 3.89 < BG ≤ 10 mmol/l			Hyperglycemia BG > 10.0 mmol/l				
	A	D	E	A	B	C	A	B	C	D	E
A	60%	40%	0%	72.9%	10.4%	0%	68.7%	10.8%	0%	1.2%	0%
B	0%	0%	0%	10.4%	2.1%	0%	13.3%	2.4%	0%	1.2%	0%
uC	0%	0%	0%	0%	0%	0%	0%	1.2%	0%	0%	0%
lC	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
uD	0%	0%	0%	2.1%	0%	0%	1.2%	0%	0%	0%	0%
lD	0%	0%	0%	0%	2.1%	0%	0%	0%	0%	0.4%	0%
uE	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
lE	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

B

Rate Error-Grid Zones sensor II	Point Error-Grid Zones **sensor II**										
	Hypoglycemia BG ≤ 3.89 mmol/l			Normoglycemia 3.89 < BG ≤ 10 mmol/l			Hyperglycemia BG > 10.0 mmol/l				
	A	D	E	A	B	C	A	B	C	D	E
A	42.9%	42.9%	0%	71.2%	11.5%	0%	58.2%	9.9%	0%	0%	0%
B	14.3%	0%	0%	13.5%	1.9%	0%	24.2%	1.1%	0%	0%	0%
uC	0%	0%	0%	0%	0%	0%	2.2%	0%	0%	0%	0%
lC	0%	0%	0%	0%	0%	0%	1.1%	0%	0%	0%	0%
uD	0%	0%	0%	1.9%	0%	0%	1.1%	0%	0%	0%	0%
lD	0%	0%	0%	0%	0%	0%	1.1%	0%	0%	0%	0%
uE	0%	0%	0%	0%	0%	0%	1.1%	0%	0%	0%	0%
lE	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

C

Zone	*Hypoglycemia*		*Normoglycemia*		*Hyperglycemia*	
	sensor I ΔT = 15 min	**sensor II** ΔT = 15 min	**sensor I** ΔT = 15 min	**sensor II** ΔT = 15 min	**sensor I** ΔT = 15 min	**sensor II** ΔT = 15 min
Accurate readings	60%	57.1%	95.8%	98.1%	95.2%	93.4%
Benign errors	0%	0%	4.2%	1.9%	2.4%	4.4%
Erroneous readings	40%	42.9%	0%	0%	2.4%	2.2%

D

Zone	*Hypoglycemia*		*Normoglycemia*		*Hyperglycemia*	
	sensor I ΔT = 5 min	**sensor II** ΔT = 3 min	**sensor I** ΔT = 5 min	**sensor II** ΔT = 3 min	**sensor I** ΔT = 5 min	**sensor II** ΔT = 3 min
Accurate readings	52.9%	59.4%	88.9%	84.0%	87.5%	75.5%
Benign errors	0%	8.1%	9.6%	13.5%	5.1%	14.5%
Erroneous readings	47.1%	32.4%	1.4%	2.5%	7.4%	10.0%

Figure 2—*Error matrix for sensor I with 15-min intervals, combining rate and point accuracy stratified by glucose range. The white, light-gray, and dark-gray squares represent accurate readings, benign errors, and erroneous readings, respectively, for sensors I* (A) *and II* (B). *C: By merging the white, light-gray, and dark-gray squares into three categories, the matrix can be collapsed into a more accessible 3 × 3 table, combined for both sensors.* D: *Collapsed error matrix calculated with 5- and 3-min time intervals for sensor I and sensor II, respectively. BG, blood glucose; l, lower; u, upper.*

ous zone E for sensor I and sensor II rates, respectively. Point accuracy, with correction for rate, revealed comparable accuracy when calculated with shorter intervals, with 96.8% (vs. 97.1%, $P = 0.869$) and 98.7% (vs. 98.0%, $P = 0.502$) of the readings of sensors I and II, respectively, in zones A or B (Table 2).

The final (collapsed) error matrix (Fig. 2*D*), calculated with shorter time intervals, showed for sensor I 7.1% more erroneous readings during hypoglycemia ($P = 0.781$, Pearson χ^2) and 5.0% more during hyperglycemia ($P = 0.102$) compared with the error matrix calculated with longer intervals (Fig. 2*C*). The error matrix for sensor II contained 7.8% significantly more erroneous readings during hyperglycemia ($P = 0.016$), when calculated with shorter intervals. However, the matrix outcome seemed to improve during hypoglycemia, with 10.5% fewer erroneous readings in that range ($P = 0.594$).

CONCLUSIONS — We applied two conventional methods, combined curve fitting and the recently proposed CG-EGA, to evaluate the contribution of the CG-EGA in comparing two subcutaneous continuous glucose sensors. MAD revealed a clear distinction in accuracy, with the MAD for sensor II being significantly lower than that for sensor I. MADs calculated per glucose range showed that, during hypoglycemia in particular, sensor II exceeded sensor I in accuracy, whereas performance was similar at normoglycemia. Sensitivity for detecting hypoglycemia was 19.1% worse for sensor I than for sensor II. According to the novel method of curve fitting, there was no significant drift for either sensor, as indicated by the vertical shift. Vertical shift of a fitted curve probably flattens out over- and underestimated glucose values but offers insight in the sensor's systematic over- or underestimation. Another advantage of combined curve fitting may be the inclusion of all measurements into the analysis in contrast to the use of paired samples only, such as with MAD, EGA, or CG-EGA. Its main advantage is its ability to assess a delay in vivo. The latter advantage was clinically relevant and statistically significant (7 min) for sensor II but not apparent for sensor I. Over all ranges, the sensor II MAD after correction for the 7-min delay improved from 13.6 to 11.7% and with ~2% in every range. Thus, according to various classical accuracy assessments and the method of combined curve fitting (horizontal shift), sensor performance differed significantly. In contrast, the results obtained through CG-EGA did not reveal a difference between the two sensors in any glucose range. Regarding the basic underlying philosophy of the CG-EGA, that the rate of glucose change is seemingly not assessed by conventional measures of sensor accuracy, it may well be argued that a significant rate error will immediately translate into a point error, as long as the sampling frequency is not too low.

When applying the CG-EGA method, we discovered some of its other limitations. We list them in the following to advance the discussion on accuracy assessment of continuous monitoring devices.

First, CG-EGA analysis is very time-consuming because a proper rate accuracy assessment requires frequent blood sampling. This raises questions concerning the applicability of CG-EGA in, for example, large-scale investigations with novel glucose sensors. Another practical drawback is that the final 9×11 matrices representing the CG-EGA results are not easy to interpret for those not thoroughly familiar with the method.

Second, despite the rate accuracy of sensor II with 3-min intervals being considerably worse than with 15-min intervals, the final CG-EGA matrix combining rate and point error grid unexpectedly indicates a better accuracy during hypoglycemia. Apparently, poor rate accuracy is barely noticeable in the final CG-EGA outcome, which disputes the value of the rate error grid part.

Third, the results based on CG-EGA vary with the time intervals. A longer duration of time intervals apparently diminishes the effect of noise and flattens out the gradual errors seen at shorter intervals. Thus, if the time intervals can be arbitrarily chosen by the investigator, CG-EGA is sensitive to interobserver variability.

Fourth, the formula to shift the grid lines in the point accuracy plot in proportion to the rate differences is based on the assumption that interstitial (sensor) glucose is delayed 7 min from the blood glucose concentration. To verify this assumption, we calculated the delay of sensor II caused by the instrument itself, making use of the volume of the transported perfusion fluid, the flow rate (11,12), and the time needed to receive the sensor signal (12). We found the instrument delay to be 6.2 min plus the unknown glucose exchange time. Thus, according to our data the 7-min delay of sensor II can be explained mainly by the instrument delay. In addition, sensor I did not show a delay, although this sensor also measures interstitial glucose concentrations. So, our data argue against the assumption regarding the 7-min delay of interstitial glucose from blood glucose, which is in accordance with another sensor accuracy study (10). This finding raises questions about the appropriateness of one of the formulas underlying the CG-EGA.

Remarkably, as in all other studies that have used CG-EGA, hardly any sensor readings in our study ended up in zone C of the point accuracy error grid (Table 2). This might be due to the shifted grid lines resulting in a decrease of the total C zone surface. The C zone is intended to represent erroneous sensor readings resulting in overcorrecting treatment, i.e., hyperglycemic sensor readings, when reference values are within the target range (5). To gain insight into the degree of deviation associated with each grid zone, we calculated the MAD over all readings per point error grid zone. Zone A included readings with mean ± SD MADs of 9.4 ± 6.1 and 11.3 ± 7.4% for sensors I and II, respectively; however, MADs of the readings in the (so-called) clinically acceptable zone B were 29.6 ± 7.9 and 28.6 ± 7.4% with a maximum MAD of 46% in that zone. MADs in the C zone could not be calculated because of the absence of readings in that zone; MADs in zone D were 47.4 ± 8.4 and 38.8 ± 5.6%. Perhaps this explains why so many readings end up in the desired zones A and B and so few in zones C and D. It seems that the borders applied in the error grid, especially after they are shifted depending on the rate, are too forgiving for accuracy assessment.

Although the effort to design a new assessment tool for continuous glucose monitors is commendable, there are major shortcomings with the CG-EGA. We emphasize the importance of using multiple established assessment methods to evaluate sensor accuracy and recommend the combination of MAD calculated per glucose range and combined curve fitting with assessment of horizontal and vertical shift and sensitivity for detecting hypoglycemia. Not only is application of these methods less time consuming for investigators, but also the final results are easier to interpret for clinicians. MAD calculated per glucose range specifies the accuracy

compared with the reference value in every glucose range; combined curve fitting allows for detection of a systematic under- or overestimation and delay. Sensitivity for detecting hypoglycemia evaluates one of the main tasks of the sensors in clinical practice, i.e., reducing the risk of hypoglycemia in intensively treated diabetic patients. This combination of measures provides the most comprehensive approach to assess continuous glucose sensor performance at this time.

References

1. Bolinder J, Deiss D, Riveline J, Battelino T, Bosi E, Tubiana-Rufi N, Kerr D, Phillip M: Guardian RT continuous glucose monitoring system with real time glucose values and alarms functions: a new tool for improving glucose control in patients with type 1 diabetes mellitus? *Diabetologia* 48:A48, 2005
2. Kaufman FR, Gibson LC, Halvorson M, Carpenter S, Fisher LK, Pitukcheewanont P: A pilot study of the continuous glucose monitoring system: clinical decisions and glycemic control after its use in pediatric type 1 diabetic subjects. *Diabetes Care* 24: 2030–2034, 2001
3. Klonoff DC: A review of continuous glucose monitoring technology. *Diabetes Technol Ther* 7:770–775, 2005
4. Kollman C, Wilson DM, Wysocki T, Tamborlane WV, Beck RW: Limitations of statistical measures of error in assessing the accuracy of continuous glucose sensors. *Diabetes Technol Ther* 7:665–672, 2005
5. Clarke WL, Cox D, Gonder-Frederick LA, Carter W, Pohl SL: Evaluating clinical accuracy of systems for self-monitoring of blood glucose. *Diabetes Care* 10:622–628, 1987
6. 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* 27:1922–1928, 2004
7. Clarke WL: The original Clarke error grid analysis (EGA). *Diabetes Technol Ther* 7:776–779, 2005
8. 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* 28:2412–2417, 2005
9. Wentholt IM, Vollebregt MA, Hart AA, Hoekstra JB, Devries JH: Comparison of a needle-type and a microdialysis continuous glucose monitor in type 1 diabetic patients. *Diabetes Care* 28:2871–2876, 2005
10. Boyne MS, Silver DM, Kaplan J, Saudek CD: Timing of changes in interstitial and venous blood glucose measured with a continuous subcutaneous glucose sensor. *Diabetes* 52:2790–2794, 2003
11. Maran A, Crepaldi C, Tiengo A, Grassi G, Vitali E, Pagano G, Bistoni S, Calabrese G, Santeusanio F, Leonetti F, Ribaudo M, Di MU, Annuzzi G, Genovese S, Riccardi G, Previti M, Cucinotta D, Giorgino F, Bellomo A, Giorgino R, Poscia A, Varalli M: Continuous subcutaneous glucose monitoring: a multicenter analysis. *Diabetes Care* 25:347–352, 2002
12. Poscia A, Mascini M, Moscone D, Luzzana M, Caramenti G, Cremonesi P, Valgimigli F, Bongiovanni C, Varalli M: A microdialysis technique for continuous subcutaneous glucose monitoring in diabetic patients (part 1). *Biosens Bioelectron* 18: 891–898, 2003