Evaluating the Clinical Accuracy of Two Continuous Glucose Sensors Using Continuous Glucose–Error Grid Analysis

WILLIAM L. CLARKE, MD[1]
STACEY ANDERSON, MD[2]
LEON FARHY, PHD[2]
MARC BRETON, PHD[3]
LINDA GONDER-FREDERICK, PHD[3]
DANIEL COX, PHD[3]
BORIS KOVATCHEV, PHD[3]

OBJECTIVE — To compare the clinical accuracy of two different continuous glucose sensors (CGS) during euglycemia and hypoglycemia using continuous glucose–error grid analysis (CG-EGA).

RESEARCH DESIGN AND METHODS — FreeStyle Navigator (Abbott Laboratories, Alameda, CA) and MiniMed CGMS (Medtronic, Northridge, CA) CGSs were applied to the abdomens of 16 type 1 diabetic subjects (age 42 ± 3 years) 12 h before the initiation of the study. Each system was calibrated according to the manufacturer's recommendations. Each subject underwent a hyperinsulinemic-euglycemic clamp (blood glucose goal 110 mg/dl) for 70–210 min followed by a 1– $mg \cdot dl^{-1} \cdot min^{-1}$ controlled reduction in blood glucose toward a nadir of 40 mg/dl. Arterialized blood glucose was determined every 5 min using a Beckman Glucose Analyzer (Fullerton, CA). CGS glucose recordings were matched to the reference blood glucose with 30-s precision, and rates of glucose change were calculated for 5-min intervals. CG-EGA was used to quantify the clinical accuracy of both systems by estimating combined point and rate accuracy of each system in the euglycemic (70−180 mg/dl) and hypoglycemic (<70 mg/dl) ranges.

RESULTS — A total of 1,104 data pairs were recorded in the euglycemic range and 250 data pairs in the hypoglycemic range. Overall correlation between CGS and reference glucose was similar for both systems (Navigator, $r = 0.84$; CGMS, $r = 0.79$, NS). During euglycemia, both CGS systems had similar clinical accuracy (Navigator zones A + B, 88.8%; CGMS zones A + B, 89.3%, NS). However, during hypoglycemia, the Navigator was significantly more clinically accurate than the CGMS (zones A + B = 82.4 vs. 61.6%, Navigator and CGMS, respectively, $P < 0.0005$).

CONCLUSIONS — CG-EGA is a helpful tool for evaluating and comparing the clinical accuracy of CGS systems in different blood glucose ranges. CG-EGA provides accuracy details beyond other methods of evaluation, including correlational analysis and the original EGA.

Diabetes Care **28:2412–2417, 2005**

From the [1]Division of Pediatric Endocrinology, Department of Pediatrics, University of Virginia Health System, Charlottesville, Virginia; the [2]Division of Endocrinology, Department of Internal Medicine, University of Virginia Health System, Charlottesville, Virginia; and the [3]Division of Behavioral Medicine, Department of Psychiatric Medicine, University of Virginia Health System, Charlottesville, Virginia.

Address correspondence and reprint requests to William L. Clarke, MD, Department of Pediatrics, University of Virginia Health System, Box 800386, Charlottesville, VA 22908. E-mail: wlc@virginia.edu.

Received for publication 3 June 2005 and accepted in revised form 18 July 2005.

W.L.C. has received honoraria and grant/research support from Abbott Laboratories. S.A., L.F., D.C., and B.K. have received grant/research support from Abbott Laboratories.

Abbreviations: CG-EGA, continuous glucose–error grid analysis; CGS, continuous glucose sensor; P-EGA, point–error grid analysis; R-EGA, rate and direction of blood change–error grid analysis.

A table elsewhere in this issue shows conventional and Système International (SI) units and conversion factors for many substances.

Continuous glucose sensors (CGSs) offer the potential to significantly increase the amount of clinically relevant information available to individuals with diabetes and their care providers. Several CGSs have the ability to determine and transmit real time blood glucose values to the individual and/or to warn of impending severe hypoglycemia and/or hyperglycemia with audible alarms (1,2). Such information could be used to alter and possibly improve clinical decision making by the user. Thus, it is crucial that the information being presented to the user be clinically accurate.

Continuous glucose–error grid analysis (CG-EGA) has been developed as a method of evaluating the clinical accuracy of CGS utilizing both blood glucose point accuracy (CGS readings versus reference blood glucose determinations) and rate and direction of change accuracy (3). CG-EGA is a logical extension of the original EGA, which was developed for assessing the clinical accuracy of patient-determined blood glucose values using either estimation or self–blood glucose monitoring systems (4). As such, it is based upon the premise that information being generated by a monitoring system should be reliable enough to result in clinically accurate decision making by the well-educated user. Accuracy and error with EGA and with CG-EGA are categorized into zones of accuracy: zone A, clinically accurate (leading to correct and safe treatment decisions); zone B, benign errors (blood glucose values outside preset precision tolerances (usually within 20% of reference) but probably not resulting in deleterious decision making); zone C, overcorrection errors (results outside of target range when reference is within target range, leading to treatment decisions that could result in blood glucose values outside of the target range); zone D, failure to detect (high or low blood glucose) errors (resulting in failure to treat either low or high blood glucose results appropriately); and zone E, erroneous errors (blood glucose values directly opposite to reference values leading to treatment decisions opposite of that needed). CG-EGA uses these same zones of clinical accuracy/

Table 1—*Point and rate accuracy of FreeStyle Navigator and MiniMed CGMS over the entire blood glucose range*

	P-EGA (point accuracy)		R-EGA (rate accuracy)	
	Navigator	CGMS	Navigator	CGMS
Zone A	71.0	69.7	66.3	63.6
Zone B	26.9	24.0	23.3	26.6
Zone C	0	0	4.2	5.1
Zone D	2.1	6.3	5.3	3.0
Zone E	0	0	0.9	1.8

Data are %.

error to present blood glucose point accuracy (P-EGA) but expands the acceptable target levels to account for physiologic interstitial time lags between blood and interstitial tissue and the rate and direction of glucose change. The expansion of the zones is dynamically adjusted for each data point depending on the blood glucose rate of change at this point. For example, when blood glucose is falling between 1 and 2 mg $\cdot$ dl^{-1} $\cdot$ min^{-1}, the upper limits of the upper A, B, and D zones are raised by 10 mg/dl (1.5 mg/dl [average rate of fall within this range] $\times$ 7 min [average delay between blood and interstitial glucose]). When blood glucose is falling between 2 and 4 mg $\cdot$ dl^{-1} $\cdot$ min^{-1}, the boundaries of zones A, B, and D are expanded by 20 mg/dl (3 mg/dl $\times$ 7 min). Similar expansion of the boundaries of lower zones A, B, and D are made when blood glucose is rising.

A separate error grid has been developed to analyze CGS clinical accuracy in terms of rate and direction of blood glucose change (R-EGA). CGS clinical accuracy is therefore presented with both point and rate determinations that can be viewed separately or combined in accuracy tables stratified by blood glucose level. The accuracy tables, or combined error matrix, display both point and rate accuracy and permit the determination of the clinical accuracy of treatment decisions, considering both concurrently. Since it has also been shown that the performance of CGS as well as self-monitoring of blood glucose devices is less accurate in the hypoglycemic than euglycemic or hyperglycemic ranges, it was decided to evaluate the accuracy of CGS sensors separately for each of these ranges (1,5–8). Thus, CG-EGA results are presented separately for low blood glucose (blood glucose $<$ 70 mg/dl), euglycemia (70 mg/dl $<$ blood glucose $\leq$ 180 mg/dl), and hyperglycemia (blood glucose $>$ 180 mg/dl). This stratification is important because the clinical message sent by an error in the hypoglycemic range is very different from the clinical message sent by the same numerical error in the euglycemic range. The target ranges can be modified, if desired. CG-EGA is compatible with both the original EGA and the consensus error grid (9).

This study evaluated the clinical accuracy of two CGS systems, the FreeStyle Navigator (Abbott Diabetes Care, Alameda, CA) and the MiniMed CGMS Model MMT-7002 (Medtronic, Northridge, CA) used simultaneously in subjects with type 1 diabetes whose blood glucose levels were clamped at euglycemia using a hyperinsulinemic glucose clamp procedure and then subsequently reduced in a controlled manner (1 mg $\cdot$ dl^{-1} $\cdot$ min^{-1}) to hypoglycemia. The data are compared for each system in two of the clinically important ranges (euglycemia and hypoglycemia) using CG-EGA.

RESEARCH DESIGN AND METHODS

— A total of 16 subjects with type 1 diabetes (11 men, 5 women, age 42 $\pm$ 3 [SD] years, duration of diabetes 20 $\pm$ 3 years) participated in the study. Informed consent was obtained from each participant. Subjects were admitted to the General Clinical Research Center in the evening before the study following a physical examination. Both CGS systems (FreeStyle Navigator and MiniMed CGMS) were applied to the abdomen of each subject at the same time, ~12 h before the initiation of the data recording in accordance with the manufacturer's instructions and calibrated as recommended. Specifically, the Navigator was calibrated at 1,3, and 24 h after insertion, and the CGMS was calibrated before operation and then four additional times ~6 h apart over each 24 h of the study. The reference blood glucose for calibration of both systems was determined using a FreeStyle blood glucose monitor (Abbott Laboratories, Alameda, CA) using fingerstick blood sampling.

A euglycemic-hyperinsulinemic clamp (insulin infusion rate 40 mU $\cdot$ kg^{-1} $\cdot$ min^{-1}, variable glucose infusion rate) was used the following morning to achieve and maintain blood glucose levels at ~110 mg/dl. Subsequently, the glucose infusion rate was reduced to permit a controlled decline in blood glucose levels of ~1 mg $\cdot$ dl^{-1} $\cdot$ min^{-1} until the blood glucose level reached 40 mg/dl. The euglycemic clamp portion of the study varied in length from 70 to 210 min, and the duration of the blood glucose reduction procedure ranged from 30 to 60 min.

Arterialized blood was sampled every 5 min, and blood glucose was determined using a Beckman Glucose Analyzer (Beckman Instruments, Fullerton, CA). FreeStyle Navigator glucose readings were recorded each minute, and MiniMed CGMS glucose readings were recorded every 5 min. Because the two systems record data at different time intervals, the CGS records were matched to the reference blood glucose with a 30-s precision, and Navigator data between CGMS readings were discarded. Rate and direction of change of CGS readings were calculated for each system at 5-min intervals.

Data analysis

The P-EGA component of the CG-EGA was used to calculate overall clinical point accuracy for the entire dataset. Next, CG-EGA results combining point and rate accuracy were calculated for each CGS in the euglycemic and hypoglycemic ranges separately (excursions into the hyperglycemic range were not part of this study). In addition to CG-EGA, we used some traditional statistical methods to reaffirm the results: glucose readings from each CGS system were correlated with reference blood glucose readings, and a linear regression model was used to determine the contribution of the readings from each system to the estimation of the reference blood glucose level overall and during hypoglycemia.

RESULTS

— A total of 1,404 matching data points were identified. There were 1,104 readings in the euglycemic (70 mg/dl $<$ blood glucose $\leq$ 180 mg/dl) range and 250 readings in the hypoglycemic range (blood glucose $<$70 mg/dl).

Clinical accuracy of each system over the entire range of glucose readings was assessed first using CG-EGA point accuracy alone as shown in Table 1 and in Fig.

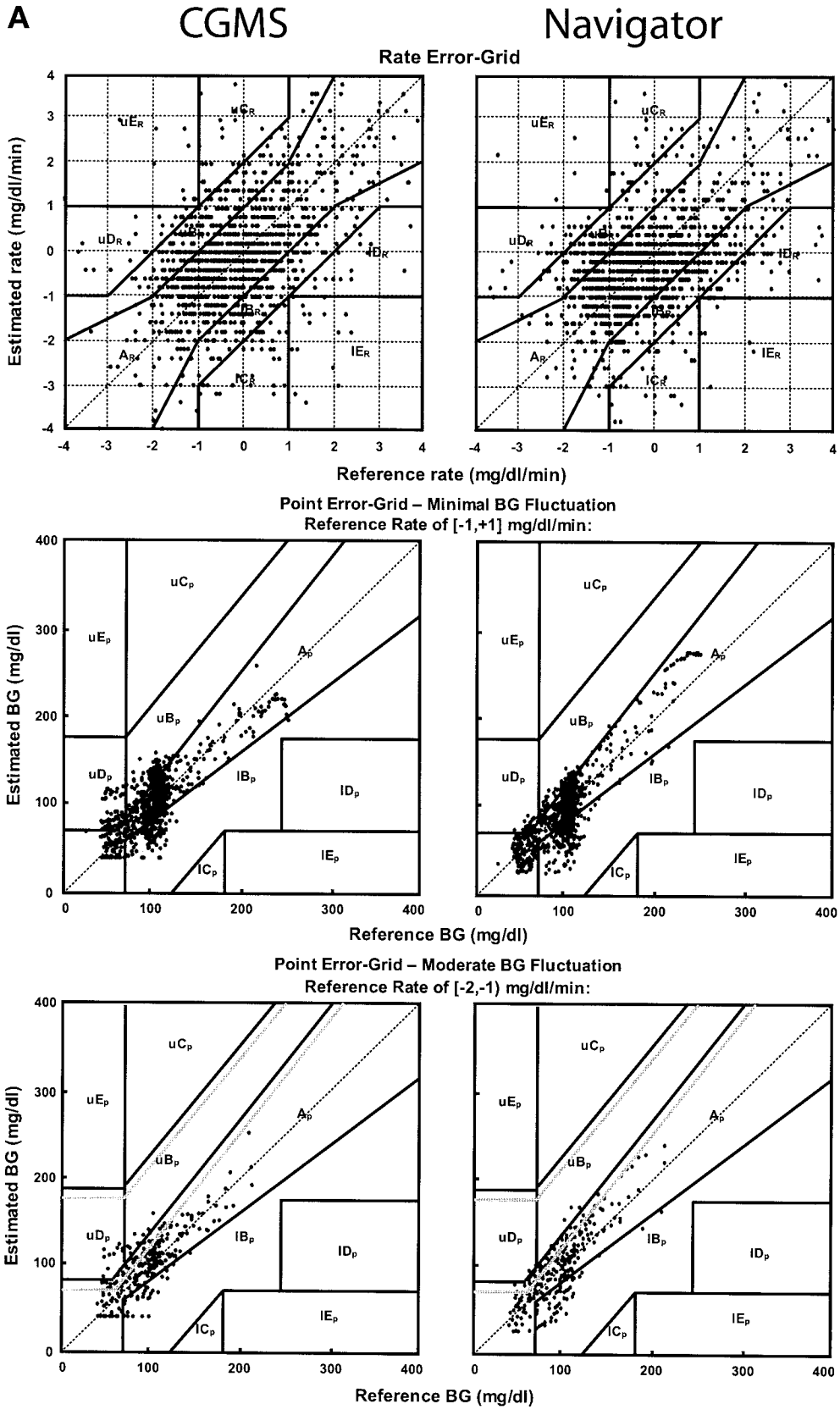

Figure 1—*Visual representations of the results for FreeStyle Navigator and MiniMed CGMS. Shown are scatter plots of the glucose values superimposed over the R-EGA grids* (A) *and scatter plots of the glucose point values superimposed over the P-EGA grids* (A *and* B). *As stated in* RESEARCH DESIGN AND METHODS, *CG-EGA plots points on grids with dynamically adjusted boundaries. BG, blood glucose.*

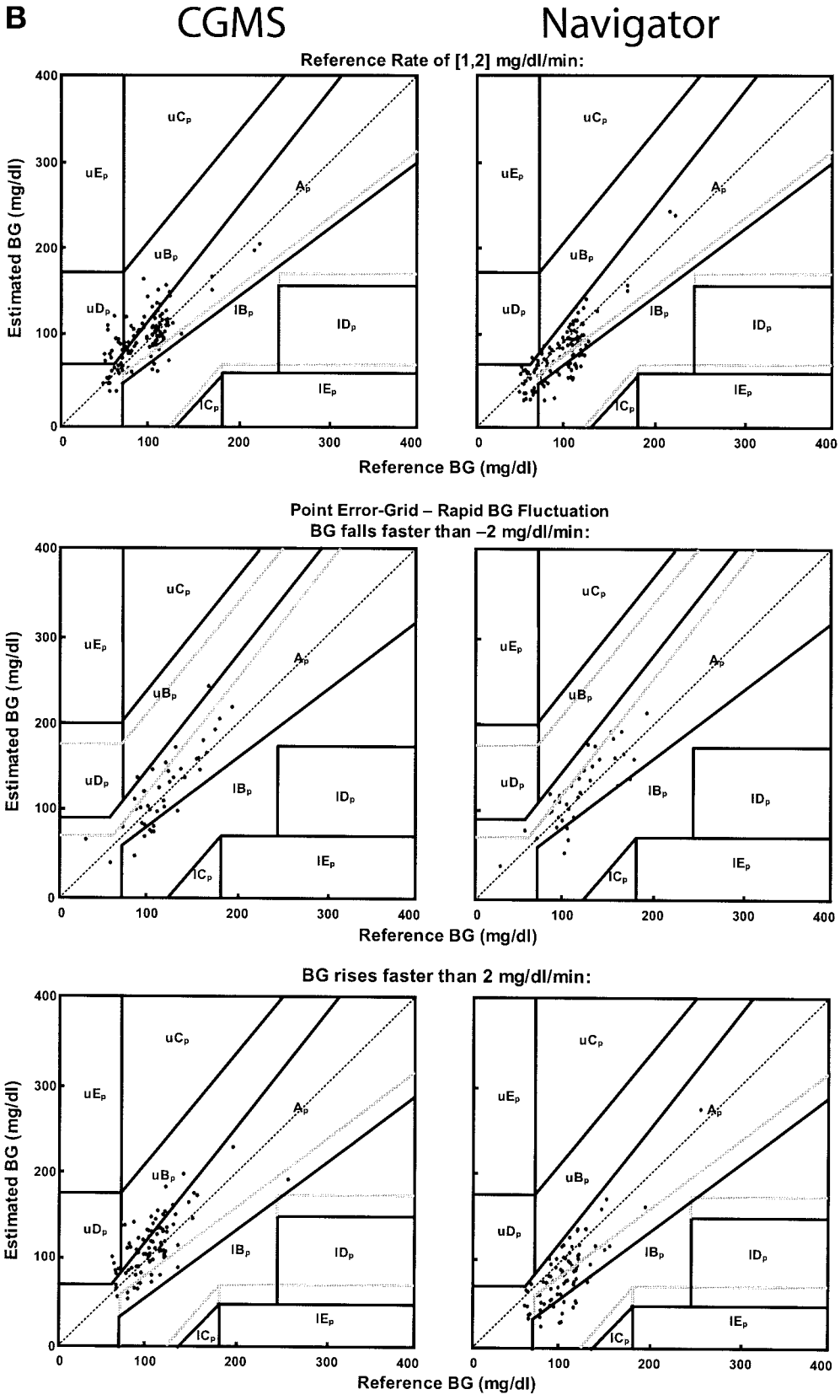

Figure 1—*Continued*

Table 2—*CG-EGA results stratified by blood glucose range*

	Hypoglycemia		Euglycemia	
Zone	Navigator	CGMS	Navigator	CGMS
Accurate readings	82.4	61.6	88.8	89.3
Benign errors	5.6	1.2	10.2	9.0
Inaccurate readings (Σ C + D + E zones)	12.0	37.2	1.0	1.7

Data are %.

1*A* and *B*. FreeStyle Navigator point accuracy was 71% for zone A and 97.9% for zones A + B, whereas MiniMed CGMS point accuracy was similar: zone A, 69.7%; and zones A + B, 93.7%. Rate accuracy was evaluated by R-EGA (Table 1 and Fig. 1*A*): 66.3 vs. 63.6% for zone A and 89.6 vs. 90.2% for zones A + B for Navigator and CGMS, respectively. There were no significant statistical differences between the point or rate accuracy of the two CGSs over the entire range of glucose readings.

The combined point + rate CG-EGA collapses the error-grid zones into three groups, accurate readings, benign errors, and clinically inaccurate readings, and stratifies them by blood glucose zone (3). Combined point and rate accuracy, CG-EGA, as shown in Table 2, revealed similar clinical accuracy for both CGS systems during euglycemia. FreeStyle Navigator data revealed 88.8% accurate readings and 99% accurate + benign error readings, whereas the MiniMed CGMS data revealed 89.3% accurate readings and 98.3% accurate + benign error readings. However, during hypoglycemia, there were significant differences in clinical accuracy between the two systems, with 82.4% accurate readings with the Navigator and 61.6% accurate readings with the CGMS ($P < 0.0005$). This difference persisted when clinical accuracy was assessed by adding accurate and benign error readings together (Navigator zones A + B, 88%; CGMS zones A + B, 62.8%; $P < 0.0005$).

The data were synchronized across subjects by setting zero time at the nadir of reference blood glucose achieved during the hypoglycemic portion of the clamp. Reference, Navigator, and CGMS readings every 5 min for 1 h before and 1/2 h after this nadir were averaged to investigate the performance of the two sensors during induced moderate hypoglycemia (Fig. 2). It is evident that FreeStyle Navigator follows the descent of the reference blood glucose closely and after the nadir overshoots the target during the recovery period, perhaps due to interstitial time lag. MiniMed CGMS fails to follow the descent of the reference blood glucose and tends to skip the nadir, missing hypoglycemia. The average nadir of reference blood glucose across subjects is 50.1 mg/dl, whereas at the same time, the average Navigator reading is 49.1 mg/dl, and the average CGMS reading is 69.3 mg/dl. Repeated measures ANOVA with contrasts shows no statistical difference between reference and Navigator data at blood glucose nadir ($t = 0.3$, NS) and a highly significant difference between reference and CGMS data at the same time point ($t = 4.2$, $P < 0.001$).

In terms of traditional statistics, both CGS systems' readings correlated significantly with the reference blood glucose over the entire range of blood glucose values and were not different from one another (FreeStyle Navigator, $r = 0.84$; MiniMed CGMS, $r = 0.79$; both P levels <0.001).

During euglycemia, regression analysis estimating the reference blood glucose from the Navigator and CGMS readings showed that both Navigator and CGMS contributed significantly to the estimation of blood glucose, with partial correlations of 0.63 for the Navigator and 0.47 for the CGMS, both significant at $P = 0.001$. However, during hypoglycemia, only the Navigator data significantly estimated the reference blood glucose ($P < 0.001$). In the presence of Navigator readings, CGMS data did not contribute to the estimation of reference blood glucose and was dropped by the stepwise regression procedure at P level of 0.1. In this latter case, the CGMS data explained <1% of the variance of hypoglycemic readings, which is consistent with the results of the CG-EGA and with Fig. 2.

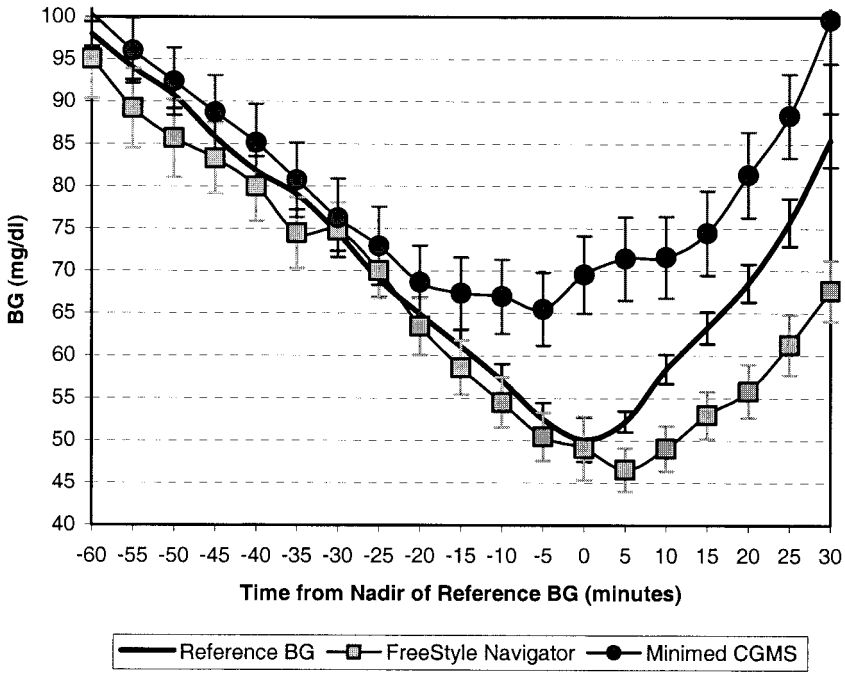

Figure 2—*Performance of the two sensors during induced moderate hypoglycemia.*

CONCLUSIONS— These results demonstrate the usefulness of CG-EGA in evaluating the clinical accuracy of CGS sensors in two clinically important blood glucose ranges. Indeed, correlational analyses, the use of the original EGA or P-EGA alone to express clinical accuracy would have suggested that the two systems tested produced data with similar clinical accuracy. Such analyses would have failed, however, to detect the significant differences in clinical accuracy between the two sensors in the critical hypoglycemic range where not only point accuracy but also the ability to follow the progression of a hypoglycemic episode

are vitally important. In that sense, the dynamic accuracy of a sensor becomes a critical feature, especially during hypoglycemia, which may develop at a much more rapid rate than the modest 1 mg · dl^{-1} · min^{-1} of the present study.

CGS devices have the potential to be important adjuncts in the management of blood glucose levels in both type 1 and type 2 diabetes. However, before their use becomes widespread it will be important for both researchers and clinicians to gain confidence in the systems' ability to produce clinically accurate information. Despite the reports that the currently commercially available systems have low sensitivity and specificity for identifying hypoglycemia (1), several manufacturers are marketing their products as hypoglycemia warning systems to be worn by children while they sleep. Although the current study was not designed to test low blood glucose alarms, the CG-EGA results in the hypoglycemic range suggest that the sensitivity of these systems vary and that further refinements in technology may be required before every system could be used relatively to warn of impending severe hypoglycemia. Indeed, the data presented in Fig. 2 suggest that treatment of hypoglycemia might not occur based on the CGMS readings, whereas overtreatment might occur with the Navigator readings.

This study is the first to report simultaneous accuracy of two different CGS systems applied to the same portion of the body. Particular attention was paid to following the manufacturers' written instructions for the insertion and calibration of the devices. Since each subject was permitted to eat their evening meal following application of the devices, it can be assumed that euglycemic and hyperglycemic excursions were recorded by the devices before the initiation of the clamp study the following morning. In addition, the rate of blood glucose fall was carefully controlled to ~1 mg · dl^{-1} · min^{-1} to mimic a rate of blood glucose reduction that would not be uncommon among persons with type 1 diabetes. Such a rate of fall should not be beyond the limits of detection of any CGS system with commercial potential.

In summary, CG-EGA has been shown to be a useful adjunct for comparing the dynamic clinical accuracy of CGS devices in different clinically important blood glucose ranges. CG-EGA permits the calculation of a numerical measure of accuracy by which different CGS devices may be evaluated and compared and may be useful in setting standards and expectations of such devices. In addition, the CG-EGA is currently the only tool available for the evaluation of sensors' ability to follow blood glucose dynamics, and we suggest its continued use for the evaluation of blood glucose events in motion, particularly for the evaluation of devices that aim to close the loop between blood glucose measurement and insulin delivery. In such a closed-loop scenario, the accurate evaluation of both blood glucose level and its direction and rate of change would be critical.

Acknowledgments— This study was supported in part by National Institutes of Health grants DK 51562, DK 28288, and 5M01 RR00847 and a grant from Abbott Diabetes Care, Abbott Laboratories, Alameda, California.

References

1. The Diabetes Res in Children Network (DirectNet) Study Group: Accuracy of the Glucowatch G2 Biographer and the continuous glucose monitoring system during hypoglycemia: experience of the Diabetes Res in Children Network. *Diabetes Care* 27:722–726, 2004
2. The Diabetes Res in Children Network (DirectNet) Study Group: The accuracy of CGMS in children with type 1 diabetes: results of the Diabetes Res in Children Network (DirectNet) accuracy study. *Diab Tech Ther* 5:781–789, 2003
3. Kovatchev B, Gonder-Frederick L, Cox D, Clarke W: Evaluating the accuracy of continuous glucose monitoring sensors. *Diabetes Care* 27:1922–1928, 2004
4. Clarke W, Cox D, Gonder-Frederick L, Carter W, Pohl S: Evaluating the clinical accuracy of systems for self-monitoring of blood glucose. *Diabetes Care* 10:622–628, 1987
5. Brunner G, Ellmerer M, Sendlhofer G, Wutte A, Trajanoski Z, Schaupp L, Quehenberger F, Wach P, Krejs G, Pieber T: Validation of home blood glucose meters with respect to clinical and analytical approaches. *Diabetes Care* 21:585–590, 1998
6. Weitgasser R, Gappmayer B, Pichler M: Newer portable glucose meters: analytical improvement compared with previous generation devices? *Clin Chem* 45:1821–1825, 1999
7. Chen E, Nichols J, Show-Hong D, Hortin G: Performance evaluation of blood glucose monitoring devices. *Diab Tech Ther* 5:749–768, 2003
8. Klonoff D: The need for separate performance goals for glucose sensors in the hypoglycemic, normoglycemic, and hyperglycemic ranges. *Diabetes Care* 27: 834–836, 2004
9. Parkes J, Slatin S, Pardo S, Ginsberg B: A new consensus error grid to evaluate the clinical significance of inaccuracies in the measurement of blood glucose. *Diabetes Care* 23:1143–1148, 2000