

ORIGINAL ARTICLE

## Use of Continuous Glucose Monitoring as an Outcome Measure in Clinical Trials

Roy W. Beck, M.D., Ph.D., Peter Calhoun, M.A., and Craig Kollman, Ph.D.

### Abstract

**Objective:** Although developed to be a management tool for individuals with diabetes, continuous glucose monitoring (CGM) also has potential value for the assessment of outcomes in clinical studies. We evaluated using CGM as such an outcome measure.

**Research Design and Methods:** Data were analyzed from six previously completed inpatient studies in which both CGM (Freestyle Navigator™ [Abbott Diabetes Care, Alameda, CA] or Guardian® [Medtronic, Northridge, CA]) and reference glucose measurements were available. The analyses included 97 days of data from 93 participants with type 1 diabetes (age range, 5–57 years; mean, 18 ± 12 years).

**Results:** Mean glucose levels per day were similar for the CGM and reference measurements (median, 148 mg/dL vs. 143 mg/dL, respectively;  $P=0.92$ ), and the correlation of the two was high ( $r=0.89$ ). Similarly, most glycemia metrics showed no significant differences comparing CGM and reference values, except that the nadir glucose tended to be slightly lower and peak glucose slightly higher with reference measurements than CGM measurements (respective median, 59 mg/dL vs. 66 mg/dL [ $P=0.05$ ] and 262 mg/dL vs. 257 mg/dL [ $P=0.003$ ]) and glucose variability as measured with the coefficient of variation was slightly lower with CGM than reference measurements (respective median, 31% vs. 35%;  $P<0.001$ ).

**Conclusions:** A reasonably high degree of concordance exists when comparing outcomes based on CGM measurements with outcomes based on reference blood glucose measurements. CGM inaccuracy and underestimation of the extremes of hyperglycemia and hypoglycemia can be accounted for in a clinical trial's study design. Thus, in appropriate settings, CGM can be a very meaningful and feasible outcome measure for clinical trials.

### Introduction

ALTHOUGH DEVELOPED TO BE a management tool for individuals with diabetes, continuous glucose monitoring (CGM) also has potential value for the assessment of outcomes in clinical studies. This is particularly true for outpatient studies, such as ones focusing on hypoglycemia reduction, short-term glucose levels, glycemic control where hemoglobin A1c (HbA1c) levels are near normal, or glycemic control with a closed-loop system. Even in longer-duration studies where HbA1c is an appropriate primary outcome measure, CGM can be an important secondary outcome, particularly for assessing time within a target range and time in the hypoglycemic range. The value of CGM as a primary outcome measure has been demonstrated in two randomized trials evaluating CGM as an intervention in individuals with type 1 diabetes with HbA1c in the excellent range.<sup>1,2</sup>

Current-generation CGM glucose measurements are not as accurate as home blood glucose meter measurements.<sup>3</sup>

However, in many outpatient studies, measurement of glucose levels with a blood glucose meter is not feasible to use as an outcome. Requiring patients to do six to eight blood glucose measurements per day at specified time intervals for long periods of time is too burdensome for many, and compliance with frequent middle-of-the-night measurements is likely to be low. In addition, blood glucose meter measurements may tend to oversample high and low values because patients will tend to measure the blood glucose when concerned that the glucose level may be low or high. Intermittent seven-point testing (before and after meals and at bedtime) was used in the Diabetes Control and Complications Trial.<sup>4</sup> However, a more recent study found it difficult to achieve a high compliance rate.<sup>5</sup> Even with seven glucose measurements a day, much can be missed, particularly related to hypoglycemia.

For some studies evaluating hypoglycemia reduction, CGM may be the optimal outcome measure. Although the goal of an intervention may be to prevent severe hypoglycemic events, the event rate may be too low for the reduction in clinical events to be a feasible outcome measure because of the

TABLE 1. PERCENTAGES OF GLUCOSE MEASUREMENTS IN DIFFERENT RANGES COMPARING CONTINUOUS GLUCOSE MONITORING AND REFERENCE GLUCOSE MEASUREMENTS

	Overall		Daytime		Nighttime	
	CGM	Reference	CGM	Reference	CGM	Reference
All values ( <i>n</i> )	52,146	5,195	27,944	2,830	24,202	2,365
71–180 mg/dL (%)	59	60	58	56	61	65
>180 mg/dL (%)	28	31	35	36	21	24
>250 mg/dL (%)	9	10	10	12	8	8
≤70 mg/dL (%)	12	9	8	7	17	11
≤60 mg/dL (%)	6	4	4	3	9	5
Mean glucose (mg/dL)	148	154	159	164	134	142
Area under the curve 70 mg/dL	1.31	0.77	0.80	0.59	1.89	0.98
Low Blood Glucose Index	2.4	1.8	1.6	1.4	3.4	2.2
Area under the curve 180 mg/dL	16.04	19.16	19.18	23.83	12.41	13.58
High Blood Glucose Index	6.6	7.5	7.8	8.9	5.1	5.8
Glucose CV (%)	47	47	43	46	50	47
Limited to paired values ( <i>n</i> )	5,206	5,206	2,950	2,950	2,256	2,256
71–180 mg/dL (%)	60	60	58	56	63	65
>180 mg/dL (%)	31	31	37	36	23	23
>250 mg/dL (%)	9	10	10	12	8	7
≤70 mg/dL (%)	9	9	6	7	14	12
≤60 mg/dL (%)	5	4	3	3	7	5
Mean glucose (mg/dL)	153	154	163	164	139	141
Area under the curve 70 mg/Dl	0.97	0.83	0.60	0.61	1.45	1.11
Low Blood Glucose Index	1.9	1.8	1.3	1.4	2.7	2.4
Area under the curve 180 mg/dL	17.38	19.01	20.75	23.57	12.97	13.06
High Blood Glucose Index	7.1	7.4	8.3	8.8	5.5	5.6
Glucose CV (%)	45	47	42	46	47	47

CGM, continuous glucose monitoring; CV, coefficient of variation (SD/mean).

need for an impractically large sample size.<sup>6</sup> In outpatient studies of overnight hypoglycemia where frequent blood glucose monitoring is not possible, CGM may be the only feasible outcome measurement to use.

In order to design studies using CGM as an outcome measure, it is helpful to have an understanding of how CGM-measured outcomes would compare with outcomes determined from blood glucose measurements. Using data from several inpatient studies, we compare CGM and reference blood glucose measurements for a variety of glycemic indices and discuss the effect of CGM-based outcomes on clinical trial design and interpretation of results.

### Research Design and Methods

Data from six previously completed studies<sup>7–10</sup> were used in the analyses. All of the studies were conducted in an inpatient clinical research center setting and had glucose data from both CGM (Freestyle Navigator™ [Abbott Diabetes Care, Inc., Alameda, CA] or Guardian® [Medtronic MiniMed, Northridge, CA]) and reference measurements (Supplementary Table S1; Supplementary Data are available online at [www.liebertonline.com/dia](http://www.liebertonline.com/dia)). Two protocols included an exercise session (38 days in 38 participants). Data from four protocols included CGM measurements every minute (three with Navigator and one with Guardian, 44 days in 40 participants), data from one protocol included Guardian CGM measurements every 5 min (24 days in 24 participants), and data from one protocol included Navigator CGM measurements every 10 min (29 days in 29 participants). One protocol

(12 days in 12 participants) had calibrations every 6 h done with a YSI analyzer (YSI Life Sciences, Yellow Springs, OH); five protocols (85 days in 85 participants) had calibrations done with a fingerstick sample according to the manufacturer's recommended schedule. Two protocols (53 days in 53 participants) allowed additional calibrations if the study personnel believed the sensor was improperly calibrated. Insulin-induced hypoglycemia testing and closed-loop sessions were not included. Reference blood glucose measurements were obtained every 15 or 30 min (four and two studies, respectively) using a YSI instrument or by sending samples to a central laboratory at the University of Minnesota in Minneapolis.

Glucose data were analyzed over a 24-h period and separately for daytime (6 a.m. to 10 p.m.) and nighttime (10 p.m. to 6 a.m.). The daytime and nighttime periods each required at least 6 h of CGM and reference data to be included; to be included in the 24-h analysis, 6 h of data in both daytime and nighttime periods was required.

The studies included 97 days of glucose data obtained from 93 participants with type 1 diabetes using an insulin pump. The 93 participants had a mean age of 18 ± 12 years (range, 5–57 years); 47% were female, and 97% were white. Mean HbA1c was 7.7 ± 1.1%. There were 61 days (61 participants) of data using a Navigator and 36 days (36 participants) using a Guardian.

The purpose of the analyses was to assess how glycemic summary metrics that might be used in a clinical trial would differ using CGM data compared with summary metrics using reference glucose values. Although the purpose was

TABLE 2. COMPARISON OF CONTINUOUS GLUCOSE MONITORING AND REFERENCE GLUCOSE MEASUREMENTS COMPUTED OVER EACH HOSPITAL DAY

	Median (25 <sup>th</sup> , 75 <sup>th</sup> percentile) for overall (n=79)					
	Reference	CGM	Difference <sup>a</sup>	Absolute value of the difference	r <sup>b</sup>	P value <sup>c</sup>
Mean glucose (mg/dL)	143 (124, 176)	148 (125, 178)	-1 (-11, 15)	13 (6, 20)	0.89	0.92
Percentage of values between 71 and 180 mg/dL (%)	65 (45, 78)	65 (47, 81)	1 (-6, 8)	7 (3, 13)	0.85	0.41
Percentage of values ≤70 mg/dL (%)	4 (0, 10)	2 (0, 10)	0 (-5, 2)	3 (0, 9)	0.59	0.30
Nadir glucose (mg/dL)	59 (54, 72)	66 (50, 90)	3 (-8, 15)	13 (7, 20)	0.52	0.05
Area under the curve 70 mg/dL	0.33 (0.00, 0.77)	0.07 (0.00, 0.82)	0.00 (-0.35, 0.22)	0.29 (0.02, 0.75)	0.55	0.80
Low Blood Glucose Index	1.1 (0.4, 2.0)	0.8 (0.2, 2.0)	0.0 (-0.8, 0.4)	0.6 (0.2, 1.4)	0.71	0.24
Percentage of values >180 mg/dL (%)	25 (12, 46)	27 (11, 49)	-1 (-6, 5)	5 (3, 13)	0.86	0.67
Peak glucose (mg/dL)	262 (205, 320)	257 (202, 313)	-14 (-36, 9)	26 (12, 48)	0.89	0.003
Area under the curve 180 mg/dL	11.63 (1.90, 23.82)	9.06 (2.33, 18.00)	-0.51 (-6.86, 2.07)	4.08 (1.26, 12.67)	0.86	0.12
High Blood Glucose Index	5.0 (2.8, 9.6)	5.2 (2.5, 8.4)	-0.4 (-1.9, 1.1)	1.5 (0.6, 3.4)	0.87	0.19
Glucose CV (%)	35 (26, 43)	31 (25, 40)	-2 (-7, 1)	4 (1, 8)	0.87	<0.001

<sup>a</sup>Continuous glucose monitoring (CGM) – reference.

<sup>b</sup>Spearman correlation.

<sup>c</sup>Wilcoxon signed rank test used to test difference between CGM and reference metrics.

CV, coefficient of variation (SD/mean).

not to assess point-by-point accuracy, to provide a perspective on the outcome metric comparison, results of a point-by-point accuracy analysis are provided in Supplementary Table S2. Reference and CGM differences were compared for each visit. Outcome measures included mean glucose level, percentage 71–180 mg/dL, percentage ≤70 mg/dL, nadir glucose, Low Blood Glucose Index, percentage >180 mg/dL, peak glucose level, High Blood Glucose Index, area under the curve for glucose values ≤70 mg/dL and >180 mg/dL, and glucose coefficient of variation.<sup>11</sup> Wilcoxon signed rank tests and Spearman correlations were used to compare reference and CGM metrics. A separate analysis was conducted to assess the correlation of various CGM-measured hypoglycemic metrics with each other using data from the Juvenile Diabetes Research Foundation CGM randomized trial.<sup>12</sup> Analyses were performed using SAS version 9.2 software (SAS Institute, Cary, NC).

## Results

Tables 1 and 2 show summary statistics for possible glycemic outcome metrics comparing CGM-measured glucose values with reference blood glucose values. In Table 1, glycemia metrics were computed by pooling data across all participant-days, whereas in Table 2 (and Supplementary Table S3, which shows the results separately for daytime and nighttime), the glycemia metrics first were computed for each participant-day, and then summary statistics were computed using the participant-day values. It can be seen in Table 2 that the medians of the participant-day mean glucose levels were similar for the CGM and reference measurements (respective median, 148 mg/dL vs. 143 mg/dL;  $P=0.92$ ), and the correlation of the two was high ( $r=0.89$ ). The high concordance of the participant-day mean glucose values can be seen in Figure 1A. Similarly, most glycemia metrics showed no significant differences when comparing CGM and reference values (Table 2). Results and the probability of detecting a true hy-

poglycemic event computed separately for the Navigator and Guardian CGM devices are shown in Supplementary Table S4 and Supplementary Figure S1, respectively. Results for each of the six studies are shown in Supplementary Table S5.

Although there were no significant differences in several hypoglycemia and hyperglycemia metrics (areas under the curve, Low and High Blood Glucose Indexes, and percentage of values in the hypoglycemic or hyperglycemic ranges), the nadir glucose level tended to be slightly lower and peak glucose level slightly higher with reference measurements compared with CGM measurements (respective median, 59 mg/dL vs. 66 mg/dL [ $P=0.05$ ] and 262 mg/dL vs. 257 mg/dL [ $P=0.003$ ]). Perhaps reflective of the differences in nadir and peak glucose values, glucose variability as measured with the coefficient of variation was lower with CGM than reference measurements (respective median, 31% vs. 35%;  $P<0.001$ ).

With respect to hypoglycemia outcomes, several different metrics can be used such as those included in Tables 1 and 2. In addition, binary outcomes can be defined such as the occurrence of a single or consecutive glucose values below a threshold. As can be seen in Table 3, all of these metrics are highly correlated.

## Discussion

The main role for using CGM to assess glycemia outcomes likely will be in outpatient studies where frequent and unbiased blood glucose monitoring generally is not feasible. Although the current analyses were conducted using inpatient data, generalization to the outpatient setting seems appropriate. These analyses have demonstrated that glycemia metrics computed from CGM approximate glycemia metrics computed from reference blood glucose measurements with the exception that CGM tends to slightly underestimate the extremes of hypoglycemia and hyperglycemia as well as glucose variability.

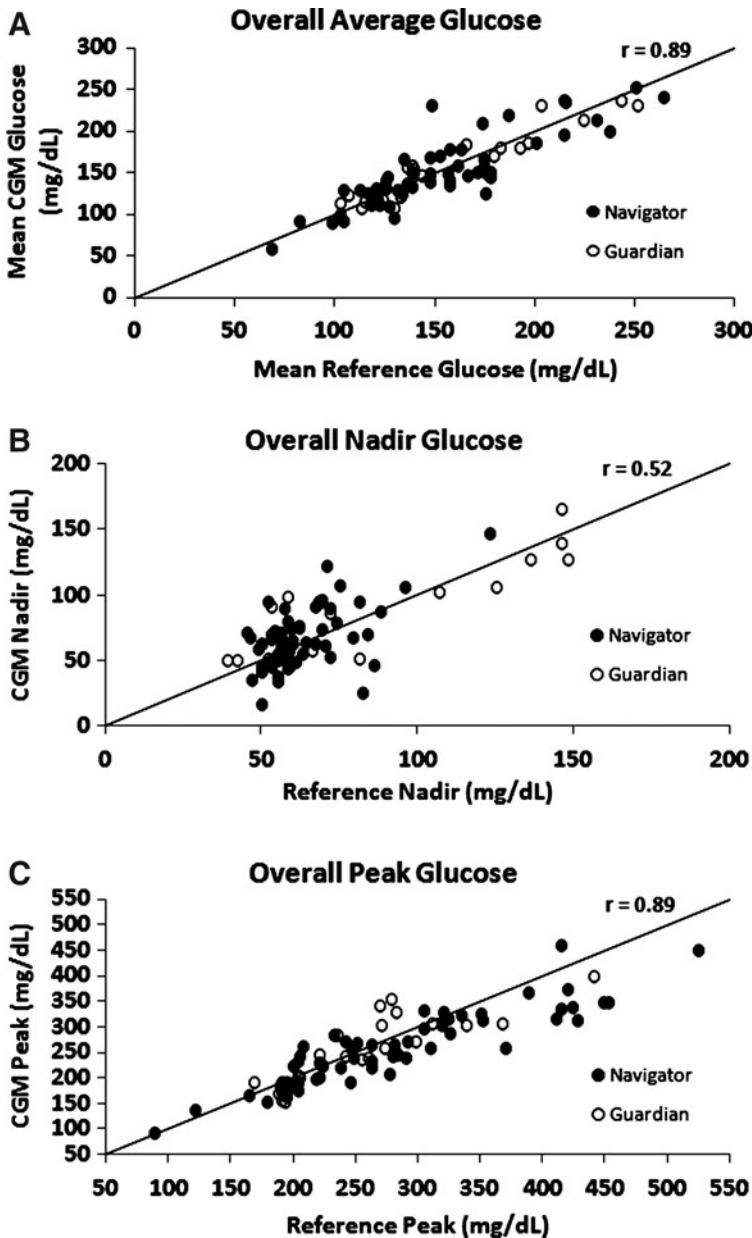


FIG. 1. Comparison of (A) mean, (B) nadir, and (C) peak continuous glucose monitoring (CGM) and reference glucose measurements for each participant-day.

Because CGM glucose measurements are not as accurate as blood glucose measurements, knowledge of the degree of concordance of CGM measurements with blood glucose measurements is important in designing clinical trials. For clinical trials using continuous outcome glycemia measures, knowledge of the expected variance is needed to compute the required sample size. Inaccuracy of CGM, which increases the variance, can be accounted for by increasing the sample size. For binary outcomes created from CGM glucose data, such as the occurrence of hypoglycemia, the relevant issue is the frequency of misclassification of the outcome due to CGM inaccuracy. Misclassification, which includes false positives (CGM indicating hypoglycemia when it did not actually occur) and false negatives (CGM misses true occurrence of hypoglycemia), dilutes the treatment effect in a clinical trial assuming that a true treatment group difference exists. However, with knowledge of the expected false-positive and false-negative rates, sample size and/or the duration of the

study can be adjusted as we have described in a prior publication.<sup>6</sup> For instance, if the false-positive and false-negative rates were each 20%, the dilution of the treatment effect would be 60% of the true treatment effect ( $100\% - 20\% - 20\%$ ), and the number of outcome measurements would need to be increased by a factor of approximately 3 ( $60\%^{-2}$ ) by increasing the sample size and/or extending the study duration.

A similar dilution of the treatment effect can occur using CGM for a continuous outcome measure. For example, suppose a clinical trial is performed using mean glucose level as the outcome measure. Further suppose that a regression equation  $\text{CGM} = (0.75 \times \text{blood glucose}) + 25 + \text{measurement error}$  exists between the CGM value and the actual blood glucose level. If the average mean glucose level is 180 mg/dL in the control group and 160 mg/dL in the intervention group, then the treatment effect,  $\Delta$ , would be 20 mg/dL. However, if CGM were used instead to assess the outcome, then the expected mean glucose in the control group would be



TABLE 3. SPEARMAN CORRELATION AMONG DIFFERENT CONTINUOUS GLUCOSE MONITORING-MEASURED DEFINITIONS OF NOCTURNAL HYPOGLYCEMIA

	Definition <sup>a</sup>							
	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>	<i>h</i>
A	1	0.95	0.99	0.94	0.93	0.92	0.93	0.91
B	0.95	1	0.93	0.99	0.87	0.92	0.88	0.91
C	0.99	0.93	1	0.94	0.95	0.93	0.95	0.91
D	0.94	0.99	0.94	1	0.89	0.94	0.90	0.93
E	0.93	0.87	0.95	0.89	1	0.96	0.99	0.95
F	0.92	0.92	0.93	0.94	0.96	1	0.97	0.98
G	0.93	0.88	0.95	0.90	0.99	0.97	1	0.96
H	0.91	0.91	0.91	0.93	0.95	0.98	0.96	1

Data of 223 continuous glucose monitoring (CGM) users in the Juvenile Diabetes Research Foundation CGM randomized controlled trial<sup>12</sup> who had at least 25 nights, each with at least 6 h of data ( $n=25,370$  nights).

<sup>a</sup>Defined as follows: A/a, single CGM value  $\leq 60$  mg/dL; B/b, single CGM value  $\leq 70$  mg/dL; C/c, two consecutive values  $\leq 60$  mg/dL; D/d, two consecutive CGM values  $\leq 70$  mg/dL; E/e, time CGM below 60 mg/dL; F/f, time CGM below 70 mg/dL; G/g, area under the curve ( $\leq 70$  mg/dL); and H/h, Low Blood Glucose Index.

$(0.75 \times 180) + 25 = 160$  mg/dL in the control group versus  $(0.75 \times 160) + 25 = 145$  mg/dL in the intervention group. Thus, in this example CGM dilutes the treatment effect,  $\Delta$ , from 20 mg/dL to 15 mg/dL. It is possible that the regression equation between CGM and blood glucose could vary between the control and intervention groups,<sup>13,14</sup> further affecting the CGM estimate of  $\Delta$  in a clinical trial. Under the null hypothesis where the intervention does not affect glucose values, the regression equation would be the same in both groups. Thus, the CGM would still give the correct type I error rate. Measurement error from CGM will generally increase the type II error rate (i.e., decrease power), which can be accounted for by increasing the sample size as noted above.

Retrospective recalibration of the CGM glucose data using blood glucose meter measurements and using regression methods to estimate the probability that the true glucose is above or below a certain threshold given the CGM reading offers the promise of reducing the variance and for binary outcomes, reducing the misclassification. Further evaluation of the role of these approaches is warranted.

For use of CGM as an outcome measure, the analyses reported herein are more relevant than point-by-point accuracy analyses, with the latter being more relevant when CGM is used by a patient for diabetes management. Accuracy metrics will generally be better for a composite measure involving multiple glucose readings than for individual data points. The data presented in this study should therefore not be confused with point accuracy.

We have demonstrated a reasonably high degree of concordance of composite outcomes based on CGM measurements with the corresponding outcomes based on reference blood glucose measurements. CGM inaccuracy and underestimation of the extremes of hyperglycemia and hypoglycemia can be accounted for in a clinical trial's study design, although the magnitude of any treatment effect will generally be rescaled by CGM. Supplementary Table S6 gives the regression equation for each CGM outcome in our study, but these could vary in different patient cohorts. As noted above, under the null hypothesis where the intervention does not affect glucose values ( $\Delta=0$ ), CGM should still give the correct type I error rate. Thus, in appropriate settings, CGM can be a very meaningful and feasible outcome measure for clinical trials.

## Acknowledgments

Study funding was provided by the Juvenile Diabetes Research Foundation, Inc. (grant 22-2011-643). We thank Roman Hovorka, Ph.D., the Juvenile Diabetes Research Foundation Continuous Glucose Monitoring Study Group, and the DirecNet Study Group for providing the data included in this article.

## Author Disclosure Statement

No competing financial interests exist.

## References

- Battelino T, Phillip M, Bratina N, Nimri R, Oskarsson P, Bolinder J: Effect of continuous glucose monitoring on hypoglycemia in type 1 diabetes. *Diabetes Care* 2011;34:795–800.
- Juvenile Diabetes Research Foundation Continuous Glucose Monitoring Study Group: The effect of continuous glucose monitoring in well-controlled type 1 diabetes. *Diabetes Care* 2009;32:1378–1383.
- Kovatchev B, Anderson S, Heinemann L, Clarke W: Comparison of the numerical and clinical accuracy of four continuous glucose monitors. *Diabetes Care* 2008;31:1160–1164.
- The Diabetes Control and Complications Trial (DCCT). Design and methodologic considerations for the feasibility phase. The DCCT Research Group. *Diabetes* 1986;35:530–545.
- Diabetes Research in Children Network (DirecNet) Study Group: Eight-point glucose testing versus the continuous glucose monitoring system in evaluation of glycemic control in type 1 diabetes. *J Clin Endocrinol Metab* 2005;90:3387–3391.
- Beck RW, Kollman C, Xing D, Buckingham BA, Chase HP: Outcome measures for outpatient hypoglycemia prevention studies. *J Diabetes Sci Technol* 2011;5:999–1004.
- Diabetes Research in Children Network (DirecNet) Study Group: The accuracy of the FreeStyle Navigator continuous glucose monitoring system in children with type 1 diabetes. *Diabetes Care* 2007;30:59–64.
- Hovorka R, Kumareswaran K, Harris J, Allen J, Elleri D, Xing D, Kollman C, Nodale M, Murphy H, Dunger D, Amiel S, Heller S, Wilinska M, Evans M: Overnight closed loop insulin delivery (artificial pancreas) in adults with type 1

- diabetes: crossover randomised controlled studies. *BMJ* 2011;342:d1855.
9. Hovorka R, Allen J, Elleri D, Chassin L, Harris J, Xing D, Kollman C, Hovorka T, Larsen AM, Nodale A, De Palma A, Wilinska M, Acerini C, Dunger D: Manual closed-loop insulin delivery in children and adolescents with type 1 diabetes: a phase 2 randomised crossover trial. *Lancet* 2010; 375:743–751.
  10. 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;10:266–272.
  11. Kovatchev BP, Straume M, Cox D, Farhy LS: Risk analysis of blood glucose data: a quantitative approach to optimizing the control of insulin dependent diabetes. *Theor Med* 2000; 3:1–10.
  12. Juvenile Diabetes Research Foundation Continuous Glucose Monitoring Study Group: Continuous glucose monitoring and intensive treatment of type 1 diabetes. *N Engl J Med* 2008;359:1464–1476.
  13. Rebrin K, Sheppard NF Jr, Steil G: Use of subcutaneous interstitial fluid glucose to estimate blood glucose: revisiting delay and sensor offset. *J Diabetes Sci Technol* 2010;4:1087–1098.
  14. Steil G, Langer M, Jaeger K, Alexander J, Gaies M, Agus M: Value of continuous glucose monitoring for minimizing severe hypoglycemia during tight glycemic control. *Pediatr Crit Care Med* 2011;12:643–648.

Address correspondence to:

Roy W. Beck, M.D., Ph.D.  
c/o Jaeb Center for Health Research  
15310 Amberly Drive, Suite 350  
Tampa, FL 33647

E-mail: [jdrfapp@jaeb.org](mailto:jdrfapp@jaeb.org)