DIABETICMedicine

DOI: 10.1111/dme.12823

Research: Care Delivery

The artificial pancreas: evaluating risk of hypoglycaemia following errors that can be expected with prolonged at-home use

H. Wolpert[1], M. Kavanagh[1], A. Atakov-Castillo[1] and G. M. Steil[2]

[1]Joslin Diabetes Center, Boston, USA and [2]Division of Medicine Critical Care, Boston Children's Hospital, Boston, USA

Accepted 29 May 2015

Abstract

Aims Artificial pancreas systems show benefit in closely monitored at-home studies, but may not have sufficient power to assess safety during infrequent, but expected, system or user errors. The aim of this study was to assess the safety of an artificial pancreas system emulating the β–cell when the glucose value used for control is improperly calibrated and participants forget to administer pre-meal insulin boluses.

Methods Artificial pancreas control was performed in a clinic research centre on three separate occasions each lasting from 10 p.m. to 2 p.m. Sensor glucose values normally used for artificial pancreas control were replaced with scaled blood glucose values calculated to be 20% lower than, equal to or 33% higher than the true blood glucose. Safe control was defined as blood glucose between 3.9 and 8.3 mmol/l.

Results Artificial pancreas control resulted in fasting scaled blood glucose values not different from target (6.67 mmol/l) at any scaling factor. Meal control with scaled blood glucose 33% higher than blood glucose resulted in supplemental carbohydrate to prevent hypoglycaemia in four of six participants during breakfast, and one participant during the night. In all instances, scaled blood glucose reported blood glucose as safe.

Conclusions Outpatient trials evaluating artificial pancreas performance based on sensor glucose may not detect hypoglycaemia when sensor glucose reads higher than blood glucose. Because these errors are expected to occur, in-hospital artificial pancreas studies using supplemental carbohydrate in anticipation of hypoglycaemia, which allow safety to be assessed in a controlled non-significant risk environment should be considered as an alternative. Inpatient studies provide a definitive alternative to model-based computer simulations and can be conducted in parallel with closely monitored outpatient artificial pancreas studies used to assess benefit.

Diabet. Med. 33, 235–242 (2016)

Introduction

There have been significant advances over the past decade in insulin pump and continuous glucose monitors. Improvements in continuous glucose monitors have allowed artificial pancreas (AP) studies to be conducted outside the closely supervised clinic environment [1–4], setting the stage for larger outpatient trials. New insulin pumps have also become available that can suspend basal insulin delivery based on a continuous glucose monitor input [5], but where the sensor glucose (SG) value is not allowed to influence when, or at what rate, the basal delivery is restarted. This reflects ongoing concerns that the sensor value may not be sufficiently accurate to ensure that there is no over-delivery of insulin [6].

However, unlike drug development, where it is difficult to predict the risks associated with a new drug, risks related to system or user errors in an AP system can easily be foreseen, and the ability of an AP system to cope with the errors can be evaluated under controlled clinical conditions without putting the patient at significant risk of harm. Of all the putative errors that might lead to harm, sensor calibration error is perhaps most concerning because these errors lead to prolonged periods when sensor glucose reads higher or lower than blood glucose (BG) [7]. Missed or delayed meals also increase the risk of hypoglycaemia, and AP algorithms relying on the user to provide insulin in advance of meals have the added risk that the individual will forget or give an

Correspondence to: Garry M. Steil. E-mail: garry.steil@childrens.harvard.edu

What's new?

- A novel method to assess the impact of sensor and other errors on the safety of an artificial pancreas is introduced in which sensor glucose values normally used for control are replaced with near-perfect minute-to-minute glucose values calculated from reference blood glucose values, and errors that reflect infrequent, but expected, events are added to the signal.
- We show that an artificial pancreas system emulating the β–cell is safe when the correct glucose values, or values 20% lower than correct are used, but that values 33% higher than correct result in control that is too aggressive to be considered safe.

inappropriate bolus. Finally, although many AP algorithms are able to adapt to changes in an individual's insulin sensitivity, the adaptation requires time, and it is during this time that the algorithms can be expected to be most unsafe. This is particularly true during instances when algorithms are adapting to an increase in the user's insulin sensitivity (S_I). In this study, we assess whether an AP system emulating the β–cell [8] can safely cope with these challenges. We defined safe control as night-time blood glucose values between 3.9 and 8.3 mmol/l, and daytime values between 3.9 and 12.3 mmol/l, with the added criteria that there should be no use of supplemental carbohydrates to prevent or treat hypoglycaemia.

Methods

Experimental procedures

Participants with Type 1 diabetes were admitted to the Beth Israel Deaconess Medical Center/Harvard Medical School Clinical Research Center (CRC) on three occasions. Participants were asked to arrive at the CRC before 4 p.m. for a scheduled dinner at 6:00 p.m., with glucose levels managed according to usual care pump therapy until 9:00 p.m. At 9:00 p.m., AP control was initiated using an algorithm developed at Boston's Children's Hospital and Joslin Diabetes Center [9–11], but with modifications to allow sensor glucose to be replaced with a scaled blood glucose (SBG) value calculated from blood glucose determinations made every 10–15 min with a YSI glucose analyser. Scaling factors (calibration factors) of 0.8, 1.0 and 1.33 were introduced to create conditions analogous to AP control with sensor glucose reading 20% lower than, equal to and 33% higher than blood glucose. For the AP algorithm studied here, this has the same effect as multiplying the algorithm's gain – a determinant of how aggressive the algorithm is – by the calibration factor, while at the same time dividing the target by the same calibration factor. Asymmetric calibration errors (–20% and +33%) were chosen to yield symmetric errors (± 1.67 mmol/l) in target, as described in the Supporting Information. A second blood glucose determination, to be used in instances in which the YSI was unavailable for any reason, was obtained with the point-of-care meter normally used by the hospital to manage blood glucose (Abbott Precision Xceed Pro).

A 3-h initialization period was allowed for the controller to stabilize (9 p.m. to midnight), after which the control was continued through the night (midnight to 8:00 a.m.) and breakfast (8:00 a.m. to 2:00 p.m.). Lunch was served at 2 p.m. to recreate the condition where a meal might be delayed. During the initialization and night intervals, changes in insulin delivery were effected every 15 min as each new blood glucose value became available; during the breakfast interval, changes were effected every minute using extrapolated blood glucose values obtained every 10 min (8–10 a.m.) or 15 min (10–2 p.m.).

Meal content, announcement and timing

Dinner consisted of brown rice, chicken, steamed broccoli and carrots, salad and fruit (83 g carbohydrate). Breakfast consisted of oatmeal, an omelette with cheese and toast (60 g carbohydrate). To create the condition where a meal insulin bolus might be missed, preprandial insulin boluses previously shown to improve control were not given [12]. Supplemental carbohydrate (15 g in juice or tablet) was provided in anticipation of blood glucose falling below 3.33 mmol/l, with the anticipated blood glucose value calculated based on rate-of-change since the previous sample. Use of supplemental carbohydrate was treated as identical to having a hypoglycaemic event (blood glucose < 3.33 mmol/l).

AP algorithm configuration

Participants were controlled without adjusting for differences in body weight or total daily dose of insulin, creating the condition where a participant's S_I may have changed rapidly, but with the algorithm not having had sufficient time to identify and adapt to that change. Target glucose was set at 6.67 mmol/l with the expectation that asymmetric calibration errors would lead to symmetric changes in target (± 1.67 mmol/l) or fasting glucose (see Supporting Information for details).

Outcomes

Primary outcome was *a priori* defined as the blood glucose area-under-the-curve (AUC) in the interval 8 a.m. to 12 p.m. (registered ClinicalTrials.gov NCT02065895). Secondary outcomes – peak postprandial glucose concentration, nadir postprandial glucose concentration, night-time time-in-target range 5.0–8.33 mmol/l and use of supplemental carbohydrate to correct or prevent blood glucose < 3.3 mmol/l within 15 min – were used to further characterize control. Repeated instances of supplemental carbohydrate without an interven-

ing blood glucose value > 3.9 mmol/l were considered a single event. Mean absolute relative difference between blood glucose and the scaled blood glucose used to effect AP control was characterized using standard formulae (100·|SG – BG|/BG) and Clark Error Grid analysis [13].

Statistical procedures

The study was powered to detect 20% differences in $AUC_{8–12}$ using an estimate of variability obtained from previous data [9]. Outcomes are reported for both scaled blood glucose and blood glucose, recognizing that the two measures are not independent (analysis was performed to assess the impact of using sensor glucose, rather than blood glucose, to report AP outcome metrics). For outcomes assumed to be normally distributed, differences were assessed by repeated measures ANOVA using Sidak's correction for multiple comparisons; non-normally distributed data were assessed using Freidman's test. Instances in which a participant was unable to complete all three scheduled visits per protocol were excluded from analysis comparing differences by calibration error, but included in intent-to-treat analysis characterizing control without regard for calibration error, as would normally be done in clinical studies where the calibration error would be unknown. Orders of admission were randomized using routines available in Excel. Results are reported as mean [± 95% confidence interval (CI)], or median [interquartile range (IQR) or range; minimum maximum] as appropriate. Power calculations were performed using Nquery v. 2.0; statistical analysis was performed using GraphPad Prism v. 6.

Risk analysis

The study was deemed non-significant risk by the investigators based on the frequency of blood glucose determinations and the use of explicit instructions for providing supplemental carbohydrate. Risk analysis was concurred by the local institutional review board and by the Food and Drug Administration. The protocol was approved by the institutional review board and all participants gave written informed consent.

Results

Participants

Consent was obtained from eight participants [median age 57 (range 40–71) years; median duration of diabetes 47 (range 11–62) years; daily insulin use 35.4 (range 27.0–45.2) units; HbA_{1c} 60 (range 51–68) mmol/mol, 7.7% (range 6.8%–8.4%), all male]. Technical difficulties during participant 1's (S1) first admission (SBG = 1.0·BG) required the admission to be repeated, with the subsequent two admissions (repeat SBG = 1.0·BG and SBG = 1.33·BG) showing blood glucose to increase > 2.8 mmol/l in multiple 10-min sample intervals. Based on this observation, the toast was removed from subsequent breakfast meals and the size of remaining meal portions adjusted as needed. Because of scheduling conflicts, S5 was only able to complete the first admission (SBG = 1.33·BG). S6 informed the study team on the morning of admission 2 (SBG = 1.33·BG) that he would be unable to complete breakfast due to an urgent unanticipated issue at work. He was admitted for the night-time control period and completed admission 3 per protocol. Individual blood glucose and insulin delivery tracing for all available data (intent-to-treat) are provided in the Supporting Information. At no point during the study did the YSI glucose analyser fail to provide blood glucose values and at no time was the point-of-care meter value used to effect control. A comparison of the YSI and point-of-care meter glucose is, however, provided in the Supporting Information.

Per protocol night-time performance

In the six participants completing all nights per protocol, blood glucose was significantly elevated [9.23 CI (7.8, 10.6); range 4.89–14.2 mmol/l vs. 6.7 mmol/l target] on starting AP control (9 p.m.). Following the 3-h controller initialization period blood glucose remained at or near target (Fig. 1a); 6 a.m. fasting scaled blood glucose was not significantly different from target as the calibration factor was altered ($P = 0.81$; Table 1). Blood glucose was not different from levels expected levels at each calibration error (not different from 5.00, 6.67, 8.33 mmol/l for sensors

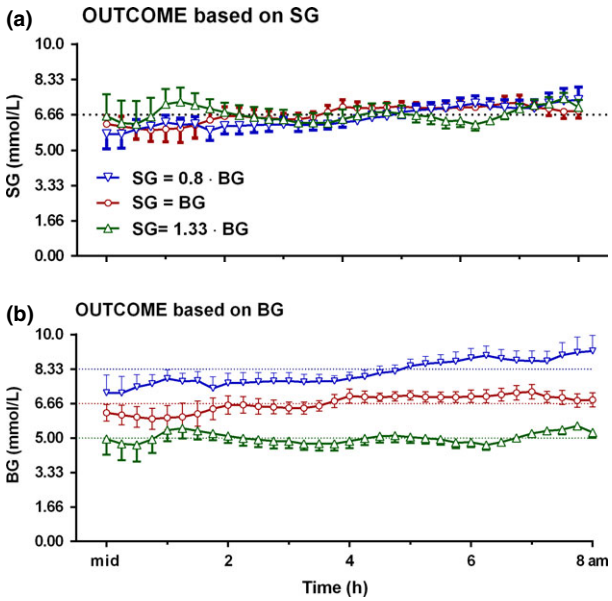

FIGURE 1 Per protocol analysis (a) Average SG obtained with a PID-AP control algorithm in which the glucose value used for contol (SG) was 1/3 higher than BG (green line), equal BG (red line), or 20% lower than BG (blue line); dashed line indicates target. (b) Same results reported as BG; dashed lines indicate expected BG values. All results reported mean±sem.

Table 1 Night-time outcomes. Per protocol analysis on six participants completed on three nights. Sensor value used for control was calculated from the YSI glucose measure, but with the calculation introducing errors either 20% lower than the true value, or 33% higher than the true value

Sensor error	Time in target$_{70-150}$(%) Median [IQR]	*P*	Midnight to 6 a.m. (mmol/l) Median [IQR]	Supplemental carbohydrate used to correct or prevent hypoglycaemia (E/S)	P	Fasting (mmol/l) Mean [95% CI]	*P*	Insulin delivered Units	*P*
Night-time outcomes based on scaled blood glucose (SBG)									
–20%	100 [100 100]	0.3333	6.2 [5.8 6.9]	none	NA	7.3 [6.0 8.6]	0.81	5.6 [4.2 6. 9]	0.07
None	100 [100 100]		6.5 [5.8 7.3]	none		7.1 [6.2 7.8]		6.1 [4.5 7.6]	
+33%	100 [90 100]		6.5 [5.7 7.2]	none		7.2 [6.3 8.0]		7.0 [5.4 8.6]	
Night-time outcomes based on blood glucose (BG)									
–20%	80 [54 86]	0.0165	7.8 [7.2 8.6]	0/0	0.1005	9.1 [7.5 10.7]	0.001	5.6 [4.2 6.9]	0.07
None	100 [100 100]		6.5 [5.8 7.3]	0/0		7.1 [6.2 8.0]		6.1 [4.5 7.6]	
+33%	88 [87 100]		4.9 [4.3 5.5]	3/2		5.4 [4.8 6.0]		7.0 [5.4 8.6]	

E, number of events; S, number of subjects.

reading 33% higher, equal to and 20% lower than blood glucose; dashed lines Fig. 1b; fasting values not different from expected value; Table 1, 95% CI). Fasting blood glucose levels were significantly different from each other as calibration error was introduced. Total insulin delivered during the night tended to increase as the calibration factor was changed from 0.8 to 1.33, but this did not achieve statistical significance ($P = 0.07$; Table 1).

All blood glucose values were within the target range when scaled blood glucose was equal to blood glucose. When scaled blood glucose was set lower or higher than blood glucose, 100% of scaled blood glucose values remained in target, but only 50% of the participants continued to be reported as having 100% of blood glucose values in target range (Table 1). With scaled blood glucose reading 20% lower than blood glucose, median time-in-target was 80%. With scaled blood glucose 33% higher than blood glucose, supplemental carbohydrate to prevent or treat hypoglycaemia was required on two occasions (two of six participants). All interventions were based on blood glucose values with scaled blood glucose, representing sensor glucose with calibration error, never fell below 4.5 mmol/l and was never predicted to fall below 3.33 mmol/l.

Intent-to-treat night-time performance

Combining all data from all participants at all calibration factors yielded 21 nights of data (see Supporting Information Figs S1–S3 for individual subject responses profiles at each of the 3 calibration errors). Median IQR was within the target zone throughout the night, irrespective of whether blood glucose or scaled blood glucose was used to calculate outcomes, but scaled blood glucose reported better performance (tighter control; Fig, 2; blue versus green shading). A histogram of all individual blood glucose and scaled blood glucose values (Fig. 2b) likewise showed better AP performance when scaled blood glucose was used to calculated time-in-target (95% vs. 85% for blood glucose). Supplemental carbohydrates were not needed for any night during which scaled blood glucose was lower than blood glucose, were required on one night with scaled blood glucose equal to blood glucose (3–4 h after control was initialized) and three nights when SBG = 1.33 BG (two of the three nights occurring 4–5 h after control was initialized). In all but one instance, the AP system was unaware that the blood glucose was anticipated to fall below 3.33 mmol/l.

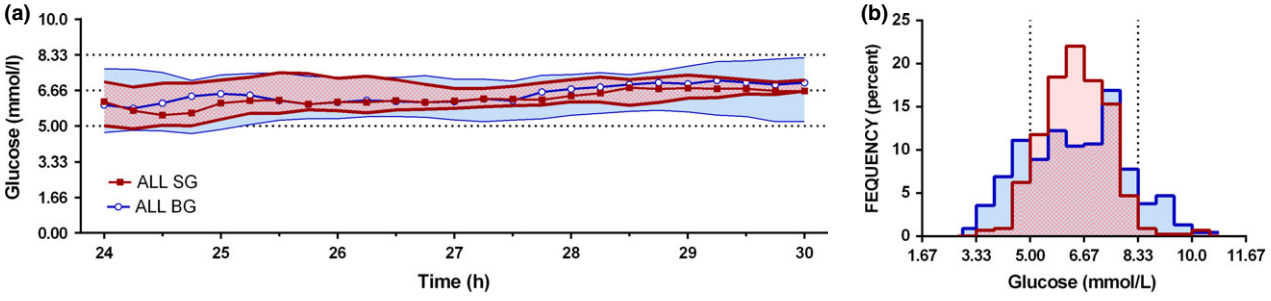

FIGURE 2 Intent to Treat analysis (all subjects). (a) Median IQR of all nighttime blood glucose (BG) and "SG" values. SG=0.8, 1.0, and 1.33 times BG. Dashed line at 6.67 mmol/l indicates target; dashed lines at 5.05 and 8.33 mmol/l indicate target range. (b) same data without regard for time-of-night.

Table 2 Meal response outcomes. Per protocol analysis on five particiapnts completing closed-loop control with all three calibrations

Calibration error	AUC_{8-2} (mmol/l/min) Mean [95% CI]	*P*	Peak (mmol/l) Mean [95% CI]	*P*	Nadir (mmol/l) Mean [95% CI]	*P*	Supplemental carbohydrate used to correct or prevent hypoglycaemia (E/S)		Insulin delivered Units	*P*
Meal outcomes based on scaled blood glucose (SBG)										
–20%	53.4 [42.6 64.2]	0.4343	11.3 [10.2 12.4]	0.0205	6.6 [5.6 7.5]	0.0174	none	NA	8.8 [6.9 10.6]	0.29
None	48.8 [41.1 56.4]		11.8 [9.8 14.0]		4.8 [3.7 6.5]		2/1		11.3 [8.2 14.4]	
+33%	49.8 [44.4 55.3]		13.3 [11.3 15.4]		4.5 [3.8 5.2]		none		11.3 [6.7 15.7]	
Meal outcomes based on blood glucose (BG)										
–20%	66.8 [53.3 80.3]	0.0011	14.1 [12.7 15.5]	0.0059	8.2 [7.1 9.3]	0.0001	none	0.0052	8.8 [6.9 10.6]	0.29
None	48.8 [41.1 56.4]		11.8 [9.8 14.0]		4.8 [3.2 6.5]		2/1		11.3 [8.2 14.4]	
+33%	37.4 [33.4 41.5]		10.1 [8.5 11.6]		3.4 [2.9 4.0]		8/5		11.3 [6.7 15.7]	

Per protocol meal performance

AUC_{8a-2p} (primary outcome) for the five participants completing all breakfast meals per protocol differed significantly when a calibration error was introduced (Table 2; $P = 0.0011$), with the highest AUC_{8a-2p} observed when the scaled blood glucose value used for control was 20% lower than blood glucose [66.776 (95% CI 53.2–80.3) mmol/l/h], and the lowest AUC_{8a-2p} was observed when the scaled blood glucose value was 33% higher than blood glucose [37.4 (33.4 41.5) mmol/l/h]. Calibration errors leading to scaled blood glucose readings higher or lower resulted in AUC_{8a-2p} values significantly different from that obtained with the correct calibration [AUC = 48.8 (41 56.4) mmol/l/h]. Fasting, postprandial peak and postprandial nadir blood glucose levels also differed significantly (Table 2; $P < 0.05$ all), with the peak postprandial value significantly higher than the maximum deemed acceptable when the scaled blood glucose was reading 20% lower than blood glucose (Table 2; maximum acceptable: 12.2 mmol/l). Nadir glucose was not different from the pre-specified minimum of 3.9 mmol/l at any calibration error. Insulin delivered in response to the meal did not differ significantly as calibration error was introduced (Table 2; $P = 0.29$).

Intent-to-treat differences between the glucose value used to effect control and true blood glucose

Introducing fixed calibration errors of –20% and +33% resulted in scaled blood glucose having a mean absolute relative difference of 18.8% (see Supporting Information Figure S4, left panel). This error included both the error in calibration and the error due to delays inherent in extrapolating blood glucose samples obtained every 10 to 15 min. Mean absolute relative difference between blood glucose obtained from the bedside hospital meter and bedside YSI analyser was 6.5%, with differences sufficient to generate sensor calibration errors of +33% occurring in 3 of 1503 paired samples, and differences sufficient large to generate errors of –20% occurring in 2 of 1503 paired samples (see Supporting Information Figure S4, right panel).

Discussion

AP studies often have power to show improvements in control during frequent events such as meals or sleep, but lack power to characterize control during infrequent events such as sensor calibration errors that might result in serious hypo- or hyperglycaemia. To characterize control during these events, at-home studies need to be conducted over long periods and in large numbers of participants. For example, if the point-of-care meter values obtained here had been used to calibrate an otherwise perfect sensor once per day, the sensor would be expected to read 33% higher than the blood glucose measured with the YSI on 3 of 1503 days. This suggests that ~ 500 patient-days would be needed before the effect of the error on the breakfast meal could be expected to be characterized in even one participant. To obtain the response in six participants, as was done here, would require six participants to be studied for well over a year, with the added risk that the hypoglycaemia might go undetected. This highlights the advantages of inpatient AP studies, in which virtually any putative high-risk factor can be assessed using small numbers of participants studied over brief periods. The number of participants, and the time needed being determined by the incidence rates and magnitude and variability of risk factor introduced, which can be set at the discretion of the study investigators.

We used this approach to characterize how our AP algorithm would behave during instances when the glucose value used by the algorithm to effect insulin delivery was incorrectly calibrated, the user fails to provide a pre-meal insulin bolus, the subsequent meal (lunch) is delayed, and the

algorithm is unaware of any recent change in insulin requirement [9,10]. Our night-time results showing a median blood glucose concentration of 6.39 mmol/l, with 85% of values between 5.00 and 8.33 mmol/l, and only one instance in which blood glucose was even anticipated to fall below 3.33 mmol/l despite the fact that introducing errors in calibration should alleviate many of concerns related to the use of an AP system at night. However, care should be taken in interpreting the results, as the differences in S_I within the population were not as large as we might have liked, with the total daily dose of insulin only varying from 27.0 to 45 U/day. The low range was likely due to the small number of participants studied. This, combined with the limited duration, also limits our ability to establish a narrow confidence interval around the estimated per night incidence rate of hypoglycaemia, which if we allow one additional hour for initialization was 1/21 nights (4.8%, 95% CI 0–24%).

However, our daytime data clearly showed that the algorithm was not robust to calibration errors during the postprandial period. More than 50% of participants (five of eight) required supplemental carbohydrates in the 5–6–h interval following breakfast when scaled blood glucose was set 33% higher than blood glucose. Of particular concern was that in every instance, the AP system was unaware of the immediate risk of hypoglycaemia, because the scaled blood glucose level never fell below 4.5 mmol/l and no value was ever predicted to fall below 3.33 mmol/l (see individual tracings, Supporting Information). We believe this result to be relevant to AP systems that rely on glucagon to prevent hypoglycaemia [2], because they count on the sensor being able to detect impending or acute hypoglycaemia. Our results also highlight a limitation to outpatient studies using sensor glucose to report outcomes, as these instances might never be reported, thereby biasing the results in favour of the AP. This bias was evident in every endpoint reported in this study, with sensor glucose consistently reporting better performance than was actually achieved based on blood glucose. Our prior studies of insulin delivery in the ICU [11,14,15] and in ambulatory participants with Type 1 diabetes [9,10,12,16,17] have all reported sensor glucose to overestimate blood glucose in the low range and underestimate values in the high range, as identified by regression analysis [18]. This characteristic can be expected to inflate AP performance both in the high and low glucose range. It is possible that these characteristics are changing or that sensor glucose values might be adjusted using advanced stochastic methodology [19,20].

It is not clear whether the results obtained with the AP algorithm studied here will be applicable to systems relying on alternate approaches and algorithms [8]. For example, model predictive control algorithms may be less sensitive to sensor calibration errors. The relative merits of the model predictive control approach vs. the approach used here have recently been debated [21,22]; however, it should be expected that any successful approach to closing-the-loop (see recent review [23]) will lead to sensor glucose values being at target under fasting conditions. If so, instances in which the sensor glucose reads higher than blood glucose will result in lower than desired blood glucose values. Approaches using insulin boluses to cover meal carbohydrate may also be less reliant on having an accurate sensor; however, instances where the sensor does read higher than blood glucose and the user gives the correct bolus may still increase the risk of hypoglycaemia. In these instances, the starting blood glucose will be lower. Confirming in each system that calibration errors expected to occur do not increase the risk of hypoglycaemia to an unacceptable level seems prudent. Likewise, instances in which the sensor reads lower than blood glucose and the subject forgets to give the pre-meal bolus should be checked to ensure they do not lead to unacceptable hyperglycaemia. These analyses can potentially be performed using computer simulation models; however, before relying on simulations, investigators should confirm that the simulations produce results consistent with the expected behaviour of the controller or the behaviour observed in prior clinical studies, as differences have been noted [24]. Investigators should also be aware that multiple simulation models have been developed with possibility that different models will yield different results [25–28].

In summary, this study failed to demonstrate that the meal response obtained with our AP algorithm is sufficiently robust to sensor calibration error resulting in sensor values 33% higher than the true blood glucose and users forget to administer pre-meal insulin boluses [9,10]. However, our night-time results show a high percentage of night-time glucose values in the target range, despite introducing errors in the glucose signal used to effect control. This observation supports the use of such a system at home during the night. Our results indicate that at-home studies using sensor glucose to quantify outcomes may underestimate hypoglycaemia when the sensor reports values higher than blood glucose are concerning because this introduces bias. The bias may even be greater than that observed here in that point-of-care meters used at home are less accurate than those used in a hospital [19]. The problem may be alleviated in part in continuous glucose monitor systems that include algorithms that protect against bad calibration points. Use of continuous glucose monitors to compare different algorithms should also be valid, as the same bias should appear in each algorithm. Overall, our findings suggest that inpatient studies validating mechanisms to ensure patients are safe under high-risk conditions expected to occur infrequently could proceed in parallel with the closely monitored outpatient studies showing the benefits of these systems.

Funding sources

Funding was received from the Juvenile Diabetes Research Foundation.

Competing interests

None declared.

Author contributions

GMS and HW contributed equally to study design. GMS analysed data and wrote first draft of manuscript. MK assisted with studies and reviewed/edited manuscript. AAC contributed to study design and manuscript.

Acknowledgements

This study was supported by a grant from the JDRF to GMS. The investigators wish express thanks to the study participants, and to the nurses of Beth Israel Deaconess Clinical Research Center.

References

1 Phillip M, Battelino T, Atlas E, Kordonouri O, Bratina N, Miller S *et al.* Nocturnal glucose control with an artificial pancreas at a diabetes camp. *N Engl J Med* 2013; **368**: 824–833.
2 Russell SJ, El-Khatib FH, Sinha M, Magyar KL, McKeon K, Goergen LG *et al.* Outpatient glycemic control with a bionic pancreas in Type 1 diabetes. *N Engl J Med* 2014; **371**: 313–325.
3 Kovatchev BP, Renard E, Cobelli C, Zisser HC, Keith-Hynes P, Anderson SM *et al.* Safety of outpatient closed-loop control: first randomized crossover trials of a wearable artificial pancreas. *Diabetes Care* 2014; **37**: 1789–1796.
4 Thabit H, Lubina-Solomon A, Stadler M, Leelarathna L, Walkinshaw E, Pernet A *et al.* Home use of closed-loop insulin delivery for overnight glucose control in adults with Type 1 diabetes: a 4–week, multicentre, randomised crossover study. *Lancet Diabetes Endocrin* 2014; **2**: 701–709.
5 Garg S, Brazg RL, Bailey TS, Buckingham BA, Slover RH, Klonoff DC *et al.* Reduction in duration of hypoglycemia by automatic suspension of insulin delivery: the in-clinic ASPIRE study. *Diabetes Technol Ther* 2012; **14**: 205–209.
6 Sherr JL, Palau Collazo M, Cengiz E, Michaud C, Carria L, Steffen AT *et al.* Safety of nighttime 2-hour suspension of basal insulin in pump-treated Type 1 diabetes even in the absence of low glucose. *Diabetes Care* 2014; **37**: 773–779.
7 *The Content of Investigational Device Exemption (IDE) and Premarket Approval (PMA) Applications for Artificial Pancreas Device Systems 11/9/2012*. Available at http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM259305.pdf Last accessed 20 February 2013.
8 Steil GM, Grodsky GM. The artificial pancreas: is it important to understand how the beta cell controls blood glucose? *J Diabetes Sci Technol* 2013; **7**: 1359–1369.
9 Wolpert HA, Atakov-Castillo A, Smith SA, Steil GM. Dietary fat acutely increases glucose concentrations and insulin requirements in patients with Type 1 diabetes: implications for carbohydrate-based bolus dose calculation and intensive diabetes management. *Diabetes Care* 2013; **36**: 810–816.
10 Dauber A, Corcia L, Safer J, Agus MS, Einis S, Steil GM. Closed-loop insulin therapy improves glycemic control in children aged < 7 years: a randomized controlled trial. *Diabetes Care* 2013; **36**: 222–227.
11 Agus MS, Steil GM, Wypij D, Costello JM, Laussen PC, Langer M *et al.* Tight glycemic control versus standard care after pediatric cardiac surgery. *N Engl J Med* 2012; **367**: 1208–1219.
12 Weinzimer SA, Steil GM, Swan KL, Dziura J, Kurtz N, Tamborlane WV. Fully automated closed-loop insulin delivery versus semiautomated hybrid control in pediatric patients with Type 1 diabetes using an artificial pancreas. *Diabetes Care* 2008; **31**: 934–939.
13 Parkes JL, Slatin SL, Pardo S, Ginsberg BH. A new consensus error grid to evaluate the clinical significance of inaccuracies in the measurement of blood glucose. *Diabetes Care* 2000; **23**: 1143–1148.
14 Steil GM, Langer M, Jaeger K, Alexander J, Gaies M, Agus MS. Value of continuous glucose monitoring for minimizing severe hypoglycemia during tight glycemic control. *Pediatr Crit Care Med* 2011; **12**: 643–648.
15 Steil GM, Alexander J, Papas A, Monica L, Modi BP, Piper H *et al.* Use of a continuous glucose sensor in an extracorporeal life support circuit. *J Diabetes Sci Technol* 2011; **5**: 93–98.
16 Steil GM, Rebrin K, Darwin C, Hariri F, Saad MF. Feasibility of automating insulin delivery for the treatment of Type 1 diabetes. *Diabetes* 2006; **55**: 3344–3350.
17 Steil GM, Palerm CC, Kurtz N, Voskanyan G, Roy A, Paz S *et al.* The effect of insulin feedback on closed loop glucose control. *J Clin Endocrinol Metab* 2011; **96**: 1402–1408.
18 Rebrin K, Sheppard NF Jr, Steil GM. Use of subcutaneous interstitial fluid glucose to estimate blood glucose: revisiting delay and sensor offset. *J Diabetes Sci Technol* 2010; **4**: 1087–1098.
19 Freckmann G, Pleus S, Link M, Zschornack E, Klotzer HM, Haug C. Performance evaluation of three continuous glucose monitoring systems: comparison of six sensors per subject in parallel. *J Diabetes Sci Technol* 2013; **7**: 842–853.
20 Hovorka R, Nodale M, Haidar A, Wilinska ME. Assessing performance of closed-loop insulin delivery systems by continuous glucose monitoring: drawbacks and way forward. *Diabetes Technol Ther* 2013; **15**: 4–12.
21 Bequette BW. Algorithms for a closed-loop artificial pancreas: the case for model predictive control. *J Diabetes Sci Technol* 2013; **7**: 1632–1643.
22 Steil GM. Algorithms for a closed-loop artificial pancreas: the case for proportional-integral-derivative control. *J Diabetes Sci Technol* 2013; **7**: 1621–1631.
23 Peyser T, Dassau E, Breton M, Skyler JS. The artificial pancreas: current status and future prospects in the management of diabetes. *Ann NY Acad Sci* 2014; **1311**: 102–123.
24 Laxminarayan S, Reifman J, Steil GM. Use of a food and drug administration-approved Type 1 diabetes mellitus simulator to evaluate and optimize a proportional-integral-derivative controller. *J Diabetes Sci Technol* 2012; **6**: 1401–1412.
25 Kovatchev BP, Breton M, Man CD, Cobelli C. In silico preclinical trials: a proof of concept in closed-loop control of Type 1 diabetes. *J Diabetes Sci Technol* 2009; **3**: 44–55.
26 Wilinska ME, Hovorka R. Glucose control in the intensive care unit by use of continuous glucose monitoring: what level of measurement error is acceptable? *Clin Chem* 2014; **60**: 1500–1509.
27 Kanderian SS, Weinzimer SA, Steil GM. The identifiable virtual patient model: comparison of simulation and clinical closed-loop study results. *J Diabetes Sci Technol* 2012; **6**: 371–379.
28 Steil GM, Hipszer B, Reifman J. Update on mathematical modeling research to support the development of automated insulin delivery systems. *J Diabetes Sci Technol* 2010; **4**: 759–769.
29 Ogata K. *Modern control engineering*, 4th ed. Upper Saddle River, NJ: Prentice Hall, 2002.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Individual tracings (Intent-to-treat). Top panel. Blood Glucose (BG; left axis) control achieved after replacing sensor glucose values normally used for control with Scaled Blood Glucose (SBG; right axis) values equal to 0.8 × BG. Bottom panel. Closed-loop Insulin delivery. Dark grey shading indicates open-loop control, light grey indicates closed-loop initialization; dashed lines indicate target (6.7 mmol/L) and hypoglycaemic threshold (3.33 mmol/L); triangle symbols indicate supplemental carbohydrates given in anticipation of BG<hypoglycaemic threshold. Bottom axis is time of day.

Figure S2. Individual tracings (Intent-to-treat). Top panel. Blood Glucose (BG; left axis) control achieved after replacing sensor glucose values normally used for control with Scaled Blood Glucose (SBG; right axis) values equal to 1.0 × BG. Bottom panel. Closed-loop insulin delivery. Dark grey shading indicates open-loop control, light grey indicates closed-loop initialization; dashed lines indicate target (6.7 mmol/L) and hypoglycaemic threshold (3.33 mmol/L); triangle symbols indicate supplemental carbohydrates given in anticipation of BG<hypoglycaemic threshold. Bottom axis is time of day

Figure S3. Individual tracings (Intent×to×treat). Top panel. Blood Glucose (BG; left axis) control achieved after replacing sensor glucose values normally used for control with Scaled Blood Glucose (SBG; right axis) values equal to 1.33 × BG. Bottom panel. Closed-loop insulin delivery. Dark grey shading indicates open-loop control, light grey indicates closed-loop initialization; dashed lines indicate target (6.7 mmol/L) and hypoglycaemic threshold (3.33 mmol/L); triangle symbols indicate supplemental carbohydrate given in anticipation of BG<hypoglycaemic threshold. Bottom axis is time of day.

Figure S4. Clarke Error Grid analysis. Left Panel. Error grid for reference YSI Blood Glucose (BG) values versus scaled blood glucose (SBG) values used for control. Mean Absolute Relative Difference (MARD) between SBG and YSI was 18.8%. Right Panel shows error grid for reference YSI Blood Glucose (BG) versus hospital meter. Green symbols indicate meter values that would have resulted in sensor glucose (SG) values reading less than 0.8 times BG; blue symbols indicate meter values that would have resulted in SG values reading greater than 1.33 times BG. MARD YSI versus meter = 6.8%.