The International Diabetes Closed Loop (iDCL) trial: A Randomized Crossover Comparison of Adaptive Model Predictive Control (MPC) Artificial Pancreas Versus Sensor Augmented Pump (SAP)/Predictive Low Glucose Suspend (PLGS) in the Outpatient Setting in Type 1 Diabetes (DCLP4)

Statistical Analysis Plan

Version 1.0

September 1, 2020

Based on Protocol Version 4.1

Version History

Version	Author	Approvers	Effective Date	Study Stage	Protocol Version
1.0	Dan Raghinaru	Craig Kollman	9/1/2020	Pilot Phase	4.1

Approvals

Role	Digital Signature or Handwritten Signature/Date
Author and Statistician: Dan Raghinaru, JCHR	Dan Raghinaru I am the author of this document 2020-09-01 10:12-04:00
Senior Statistician: Craig Kollman, JCHR	Craig Kollman I am digitally signing this document 2020-09-11 15:49-04:00
Coordinating Center Director: John Lum, JCHR	John Lum I am electronically signing this document 2020-09-11 16:01-04:00
Sponsor: Jordan Pinsker, Sansum Diabetes Research Institute	Jordan Pinsker Digitally signed by Jordan Pinsker Date: 2020.09.11 13:34:49 -07'00'

1. Study Overview

The following table gives an overview of the DCLP4 study.

Table 1. Study Overview

Title	The International Diabetes Closed Loop (iDCL) trial: A Randomized Crossover Comparison of Adaptive Model Predictive Control (MPC) Artificial Pancreas Versus Sensor Augmented Pump (SAP)/Predictive Low Glucose Suspend (PLGS) in the Outpatient Setting in Type 1 Diabetes (DCLP4)
Précis	A randomized crossover trial will compare the efficacy and safety of an automated insulin delivery (AID) study system using an adaptive Model Predictive Control (MPC) algorithm versus SAP (which may or may not include PLGS; to be referred to as SAP) therapy in people with type 1 diabetes. A Pilot Phase involving at least 5 participants using the study system for 10-14 days will be conducted prior to the crossover trial.
Investigational Device	The AID study system is composed of a Tandem t:AP pump, a Dexcom G6 continuous glucose monitoring sensor, and a smart phone that contains the adaptive algorithm and communicates with the other devices.
Objectives	To compare the efficacy and safety of an AID system using an adaptive MPC algorithm versus SAP or PLGS therapy in people with type 1 diabetes.
Study Design	Randomized Crossover Trial with two 13-week periods, preceded by a CGM run-in phase.
Number of Sites	5-7 U.S. sites
Major Endpoints	Superiority for time in range 70-180 mg/dL and non-inferiority for time <54 mg/dL measured with CGM will be considered primary endpoints, analyzed using a hierarchical gatekeeping testing procedure Secondary outcomes will include HbA1c and the following CGM metrics • Mean glucose • Time >180 mg/dL • Time >250 mg/dL • Time <70 mg/dL • Time <54 mg/dL (superiority) • Coefficient of variation
Population	**Inclusion** • Clinical diagnosis, based on investigator assessment, of type 1 diabetes for at least one year and using insulin for at least 1 year • Using an insulin pump for at least 3 months (which may include use of automated features) • Familiarity and use of a carbohydrate ratio for meal boluses • Age ≥18.0 years old • For females, not currently known to be pregnant *If female and sexually active, must agree to use a form of contraception to prevent pregnancy while a participant in the study. A negative serum or urine pregnancy test will be required for all females of child-bearing potential. Participants who become pregnant will be discontinued from the study. Also, participants who during the study develop and express the intention to become pregnant within the timespan of the study will be discontinued.* • If using a personal CGM, willingness to use a Dexcom G6 CGM and discontinue personal CGM use during the study • Willing not to begin use of, or not to continue use of if currently using, a personal AID (closed loop control) system during the study; *note if the system offers an open-loop mode or*

	can be switched to a PLGS mode that is compatible with the Dexcom G6, the system may be used during the study in these modes only • Willingness to switch to lispro (Humalog) or aspart (Novolog) if not using already, and to use no other insulin besides lispro (Humalog) or aspart (Novolog) during the study • Willingness not to start any new non-insulin glucose-lowering agent during the course of the trial, and not to use Afrezza during the trial • Investigator believes that the participant can successfully and safely operate all study devices and is capable of adhering to the protocol **Exclusion Criteria** 1) Use of Afrezza or any non-insulin glucose-lowering agent other than metformin (including GLP-1 agonists, DPP-4 inhibitors, SGLT-2 inhibitors, sulfonylureas) unless participant is willing to discontinue during the trial. 2) Two or more episodes of DKA requiring an emergency room visit or hospitalization in the past 6 months 3) Two or more episodes of severe hypoglycemia with seizure or loss of consciousness in the last 6 months 4) Hemophilia or any other bleeding disorder 5) A medical or other condition that in the opinion of the investigator could create a safety concern for the participant or put the study at risk a. History of frequent severe hypoglycemia or history of frequent severe hyperglycemia and/or ketosis, without emergency room visit or hospitalization, due to poor diabetes self-management may be disqualifying per investigator judgment 6) Participation in another pharmaceutical or device trial at the time of enrollment or during the study
Sample Size	Up to 35 individuals randomized in the crossover trial with the goal of at least 32 completing the trial
Treatment Groups	Random assignment (1:1) to begin period 1 either with the AID study system or with SAP using home pump and then cross over to the other respective treatment in period 2
Participant Duration	6-7 months
Protocol Overview/ Synopsis	**Screening and Enrollment** • Informed consent will be signed, and eligibility will be assessed • Medical history and physical examination • HbA1c measurement using point-of-care device (if not available from prior 3 months) • Urine pregnancy test (if applicable) • Baseline questionnaires completed **Pilot Phase** The Pilot Phase will include 1-2 participants at each site, using a personal Dexcom G5/6 with use on at least 11 of the prior 14 days and the investigator believes that a CGM run-in is unnecessary. The Pilot Phase is intended to (1) test the functionality of all aspects of the study System and (2) train the clinical staff on the execution of the clinical protocol, including hands-on training with the device prior to initiating the RCT Period. Screening and study system start up procedures will be the same as for the Crossover Trial. Participants will use the study system for 10-14 days. After at least 5 total participants from at least 3 different sites have completed the Pilot Phase, the data will be reviewed for preset safety criteria before proceeding to the RCT. Assuming there are no significant safety or device issues that occur, the Crossover Trial will begin. **CGM Run-In (2-4 weeks)** Eligible participants using a personal Dexcom G5/G6 sensor, with use on at least 11 of the prior 14 days may skip the CGM run-in unless the investigator believes that there is a need to optimize insulin pump settings prior to randomization. Participants skipping the run-in can go directly to the randomization visit.

	Eligible participants not currently using a personal Dexcom G5/G6 sensor or using a Dexcom G5/G6 sensor with readings captured on less than 11 out of the prior 14 days will initiate a CGM run-in phase with a study Dexcom G6 sensor. • *Users of a sensor other than Dexcom G5/G6 (eg, Medtronic, Abbott) must be willing to discontinue use of that sensor for the duration of the study and users of an AID (closed loop control) system must be willing to discontinue its use during the study (PLGS mode can be used during the study if feasible).* Duration of the run-in will be 2-4 weeks at investigator discretion. During this time, Study Staff will review uploaded CGM data and may optimize insulin pump settings as needed by phone contact with subjects. Prior to each phone contact, the participant will be asked to upload the CGM and if possible pump data for study staff review. Successful completion of the run-in will require use of the sensor for at least 11 days during a final assessment period of 14 days. The run-in may be extended at investigator discretion if the investigator believes that additional training is needed or the participant should have another opportunity to achieve the 11 day use criterion. **Randomized Crossover Trial** Eligible participants will be randomly assigned to one of two treatment groups for 13 weeks each: 1. Group A: study system for period 1 for 13 weeks, then SAP for period 2 for 13 weeks 2. Group B: SAP for period 1 for 13 weeks, then study system for Period 2 for 13 weeks Study System Period: The participant will be trained regarding AID study system use including meal announcement, meal bolusing, correction doses, and exercise. Participants will be expected to use the study system at all times at home with the exception of times of illness. SAP Period: During the SAP period, participants will use their own insulin pump and a study-provided Dexcom G6 sensor. • *PLGS mode on the pump can be used if feasible but an AID system cannot be used.* **Study Flow** Period 1 will commence on the day of randomization with initiation of the study system or continued use of personal pump SAP. Both groups will have the same contact and visit schedule in both period 1 and period 2 • Visits at 2 weeks, 6 weeks, and 13 weeks • Phone contacts at 3 days, 4 weeks, and 9 weeks At randomization, and at each end of period visit, central lab HbA1c measurement will be made, a pregnancy test will be performed in applicable subjects, and quality of life and treatment satisfaction questionnaires will be completed. At the end of period 1 visit, the intervention will be switched to the study system or SAP for period 2. There will not be a washout period. Throughout the study, the occurrence of adverse events and device issues will be solicited and recorded.

The following table provides an overview of the schedule of study visits, phone contacts, and key procedures.

Table 2: Schedule of Visits and Procedures

	Screening Visit	CGM Run-in Visit	Randomization Visit	Days (d) or Weeks (w) from the Start of Each of Two 13-Week Study Periods[1]						
				0d[2]	3d	2w	4w	6w	9w	13w
Visit (V) or Contact (C)			V	V	C	V	C	V	C	V
Informed Consent	X									
Eligibility Assessment	X									
Medical history/ physical exam	X									
Height, weight, blood pressure and pulse	X									X
HbAlc (POC or local lab, if needed)	X									
HbA1c (Central lab)			X							X
C-peptide and blood glucose (Central Lab)			X							
Pregnancy test (females of child-bearing potential)	X		X	X[3]						
Questionnaires	X									X
Assessment of CGM use	X	X								
Study system training			X	X[4]						
AE Assessment		X	X		X	X	X	X	X	X
Upload device data from home					X		X		X	
Download device data at clinic visit		X				X		X		X

1. Two 13-Week treatment periods with no intervening washout
2. Will coincide with Randomization Visit for study period 1
3. For study period 2 only
4. For Group B participants (those assigned to use SAP during period 1) only

2. Comparison with Protocol

The author (Dan Raghinaru) attests the consistency between this document and Protocol v4.1.

The following analyses were added here that are not in the protocol:

- Binary CGM outcomes:
  - % in range 70-180 mg/dL >70%
  - % <54 mg/dL <1%
  - % in range 70-180 mg/dL >70% and % <54 mg/dL <1%
- Sensitivity Analyses – First Week of CGM data
- Additional Tabulations and Analyses – weekly CGM metrics during MPC period
- Hypoglycemia events (defined as at least 15 consecutive minutes <54 mg/dL) was moved from safety to secondary outcomes and replaced the previous hypoglycemia events (defined as at least 15 consecutive minutes <70 mg/dL)
- CGM-measured hyperglycemic events (≥15 minutes with glucose concentration >300 mg/dL) moved from safety to secondary outcomes.

3. Statistical Hypotheses

The primary outcome for this study is CGM-measured % in range 70-180 mg/dL over 12 weeks in each one of the two crossover periods. The intervention will be considered effective if % time in range during the MPC period is superior to the SAP period using a statistical significance of $\alpha=0.05$ and the model specified below in Section 7 (i.e. $p < 0.05$)

The null/alternative hypotheses are:

a. *Null Hypothesis*: There is no difference in mean CGM-measured % in range 70-180 mg/dL over 12 weeks between SAP and MPC periods

b. *Alternative Hypothesis*: The mean CGM-measured % in range 70-180 mg/dL over 12 weeks is different for SAP and MPC periods.

4. Sample Size

Sample size has been computed for the primary outcome (CGM-measured % in range 70-180 mg/dL). Data from the SAP group in DCLP3 (The International Diabetes Closed Loop Protocol 3 Pivotal Trial of t:slim X2 with Control-IQ Technology) were used to calculate sample size specific to ≥18 years age group. For the SAP group, the standard deviation for time in range 70-180 mg/dL over the course of the first 3 months was 12% (95% CI 10% to 16%) and the correlation between the first 3 months and last 3 months was 0.92.

A total sample size was computed to be N=31 for the following assumptions: (1) two 3-month periods [MPC:SAP] crossover randomization, (2) 90% power, (3) a 8% absolute increase in % time in range 70-180 mg/dL, (4) a SD of 17%, (5) a correlation between the two periods of 0.70, and (6) 2-sided type 1 error of 0.05.

The total sample size has been increased to N=35 to account for potential dropouts.

5. Outcome Measures

5.1. Primary Efficacy Endpoint:

- CGM-measured % in range 70-180 mg/dL over 12 weeks in each one of the two crossover periods

5.2. Secondary Efficacy Endpoints

5.2.1. Hierarchical Endpoint

The following secondary endpoint will be tested in a hierarchical fashion only if the primary efficacy endpoint reached statistical significance as described in Section 8.1 below:

- CGM-measured % below 54 mg/dL over 12 weeks in each one of the two crossover periods (non-inferiority outcome with a non-inferiority limit of 1.0%). This is the only non-inferiority and one-sided outcome in the study.

5.2.2. Other Secondary Endpoints

The following endpoints are considered secondary.

- CGM metrics related to overall control over 12 weeks
  - % in range 70-140 mg/dL
  - mean glucose
  - glucose variability measured with the coefficient of variation
  - glucose variability measured with the standard deviation
  - % in range 70-180 mg/dL >70%
  - % in range 70-180 mg/dL >70% and % <54 mg/dL <1%
- CGM metrics related to hypoglycemia over 12 weeks
  - % <70 mg/dL
  - % <60 mg/dL
  - % <54 mg/dL (superiority – see details in Section 5.2.1 above)
  - low blood glucose index
  - hypoglycemia events (defined as at least 15 consecutive minutes <54 mg/dL)
  - % <54 mg/dL <1%
- CGM metrics related to hyperglycemia over 12 weeks
  - % >180 mg/dL
  - % >250 mg/dL
  - % >300 mg/dL
  - high blood glucose index
  - CGM-measured hyperglycemic events (≥15 minutes with glucose concentration >300 mg/dL)
- CGM metrics by time of day
  - Calculate all CGM metrics listed above (including the primary outcome) for:
    - All 24 hours of the day (the primary outcome is over 24hr)
    - Daytime only (06:00AM to 00:00AM)

        - Nighttime only (00:00AM to 06:00AM)
- HbA1c at 13 weeks
    - HbA1c
    - HbA1c <7.0%
    - HbA1c <7.5%
- Insulin at 13 weeks
    - Total daily insulin (units/kg)
    - Basal: bolus insulin ratio
- Weight and body mass index at 13 weeks
- Diabetes Distress Scale at 13 weeks – total score and 4 subscales:
    - Emotional burden
    - Physician-related distress
    - Regimen-related distress
    - Interpersonal distress
- Glucose Monitoring Satisfaction Survey – measures treatment satisfaction and burden
- Hypoglycemia Confidence Scale
- Diabetes Technology Attitudes survey
- INSPIRE survey scores - MPC period only
- System Usability Survey - MPC period and at 13 weeks only

5.3 Calculation of CGM Metrics (primary and secondary):

- Baseline: The last 2 weeks of personal CGM data or last 2 weeks of CGM run-in data will serve as baseline. If a participant has <72hr of data in the last two weeks of CGM run-in, then would go back to the 3rd week and, if still <72hr, then would repeat and go back to the 4th week.
- Each one of the Two Follow-up Periods: Since participants will receive up to 7 days of training before taking the MPC system home, the first 7 days in the SAP and MPC periods will not be included in the CGM analyses. CGM data starting from day 8 following randomization visit through the 13-week visit for first period and from day 8 following treatment reassignment through the second 13-week visit for the second crossover period will be included in the calculation of CGM metrics.
- At least 72hr of CGM data at baseline and in each one of the two periods are needed for the metrics to be calculated.
- All CGM metrics at baseline and follow-up will be calculated giving equal weight to each sensor reading for each participant.

5.4. Questionnaires

All questionnaires will be administered online, and participants are allowed to skip specific questionnaires or items within a questionnaire. The data might therefore include some missing items. All questionnaires will be scored according to the instructions given in the manual. In case no manual exists

for a given questionnaire or the manual does not provide guidance on how to handle missing items, then the following criteria will be applied.

At least 75% of the questions must be completed to be included in the analysis. This 75% rule will be applied separately for the total score and each subscale so it is possible the sample size will be different for some subscales. The score used for analysis will be based on the average among the questions that were answered and then scaled accordingly.

5.5 Analysis Windows

Analysis windows apply to the following outcomes measured at 13-week visit in each one of the two crossover periods:

- HbA1c
- Insulin metrics
- Height/Weight
- Questionnaires

In order to be included in the corresponding analyses, these data must be collected between days 76 and 106 for both the first and second 13-week visit date within each period. The target date is 91 days from the beginning of each one of the two crossover periods.

This does not apply to the CGM metrics which are calculated as described above.

6. Description of Statistical Methods

6.1. General Approach:

- All analyses comparing the MPC with the SAP period will follow the intention-to-treat (ITT) principle with each participant analyzed according to the treatment period assigned by randomization.
- In the unlikely event that a randomized participant is later found to be ineligible, the participant will be excluded from all analyses.
- Otherwise, all randomized participants with at least 72h of CGM data in at least 1 follow-up period will be included in the primary and secondary analyses.
- All p-values – apart from the secondary hierarchical CGM-measured % below 54 mg/dL- will be two-sided.
- Standard residual diagnostics, based on the differences between the two periods, will be performed for all analyses. If values are highly skewed, then a winsorizing at $10/90^{th}$ percentile, nonparametric, or MM estimation methods will be used instead for the primary and secondary outcomes. Previous experience suggests that no transformation, nonparametric, or MM estimation analyses will be necessary for % time in range 70-180 mg/dL, % above 180 mg/dL, mean glucose, or HbA1c. Other outcomes like % below 54 mg/dL are skewed; however, the differences between the two periods may follow a normal distribution and there may be no need for transformation, nonparametric, or MM estimation.

6.2 Analysis Cohorts

Primary and Secondary Analyses:

- All randomized participants and periods will be analyzed according to the ITT principle as described above.
- All participants with at least 72hr of CGM data in at least one randomized period will be included in the CGM-based analyses.
- All randomized periods with a HbA1c measurement at 13 weeks will be included in HbA1c analyses. Similar approaches will be followed for the other secondary outcomes like insulin, weight/BMI, and questionnaires analyses.

Per Protocol (PP) Analyses:

Three different per protocol analyses will be considered:

- If more than 10% of participants have fewer than 168 hours CGM data in either treatment period, the primary and secondary hierarchical analyses will be replicated excluding such participants.
- The primary and secondary hierarchical analyses will be replicated restricting to participants who:
  - used the system in CL mode for >80% (i.e. >1,613 hours) during the MPC period, and
  - used the sensor for >80% (i.e. >1,613 hours) during the SAP period.

Safety Analyses:

- Safety outcomes will be reported for all randomized participants by period. Separately, any reported adverse events pre-randomization will be tabulated.

Sensitivity Analysis:

- Missing Data: As noted below in Section 6.2, all participants with ≥72hr of CGM data in at least one of the two crossover periods will be included in primary analyses. It is also worth emphasizing that any statistical method for handling missing data makes untestable assumptions. The goal will be to minimize the amount of missing data in this study so that results and conclusions will not be sensitive to which statistical method is used. To that end, sensitivity analyses will be performed to explore whether results are similar for primary and secondary hierarchical analysis when using different methods. The following methods will be applied:
  - Available cases – participants with ≥72hr of CGM data in at least one of the two crossover periods (this is the default method)
  - Complete cases only – participants with ≥72hr of CGM data in both crossover periods
  - Rubin's multiple imputation

- Carry-over Effect: As noted in Section 7 below, the model will include pre-randomization baseline, the two crossover periods, and will adjust for period and site (as random effect). A model will be run for the primary outcome to test for any carry-over effects, by adding a treatment by period interaction term.
- First Week of Data: The primary and secondary hierarchical CGM analyses will be repeated without excluding the first week of CGM data.

7. Primary Analysis

The primary outcome is CGM measured % time in range 70-180 mg/dL over the two randomized crossover periods.
Summary statistics (mean ± SD or median (quartiles)) will be reported for the CGM-measured % in range 70-180 mg/dL at baseline, over the two randomized crossover periods, and the difference between the two periods.

A repeated measures linear regression model will be fit for CGM-measured % in range 70-180 mg/dL. The model will include pre-randomization baseline, the two crossover periods, and will adjust for period and site (as random effect). Primary analysis will report the point estimate, 95% confidence interval and p-value for the randomized periods difference at follow-up. Inclusion of the pre-randomization baseline value as a third observation for each participant in the model gives a variance reduction analogous to adjusting for it as a covariate. Baseline is not modeled as a covariate in this analysis because there is no corresponding baseline for the second period, but only pre-randomization. The model will account for correlated data from the same participant. Residual values will be examined for an approximate normal distribution. If residuals are highly skewed, then a winsorizing at $10/90^{th}$ percentile or robust statistical method (e.g., non-parametric or MM estimation) will be used instead. It is expected that the residual values for CGM-measured % in range 70-180 mg/dL will follow an approximate normal distribution.

If a participant has <72hr of data for one of the two cross-over periods or at baseline, then the participant and the periods with ≥72hr of data will be included in the primary analysis.

8. Analysis of the Secondary Endpoints

Point estimates and confidence intervals for the treatment period differences will be presented for all secondary metrics. Models similar with the one described above will be implemented.

8.1. Hierarchical Analysis

The only hierarchical outcome in this study will be non-inferiority in CGM-measured % below 54 mg/dL.

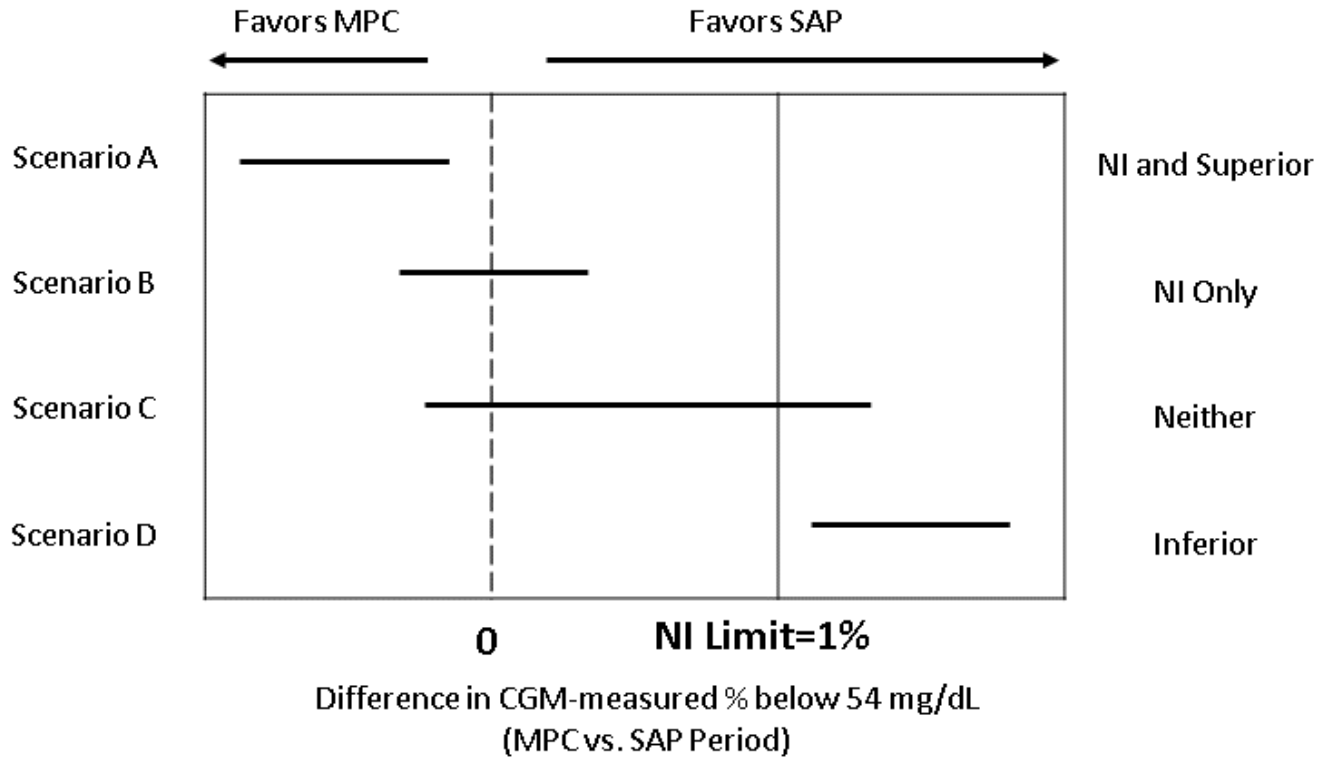
Regression Model

A repeated measures linear regression model will be fit for CGM-measured % below 54 mg/dL. The model will include pre-randomization baseline, the two crossover periods, and will adjust for period and

site (as random effect). Will report the point estimate, two-sided 95% confidence interval and one-sided p-value for the randomized periods difference at follow-up. Inclusion of the pre-randomization baseline value as a third observation for each participant in the model gives a variance reduction analogous to adjusting for it as a covariate. The model will account for correlated data from the same participant. Residual values will be examined for an approximate normal distribution. If residuals are highly skewed, then a winsorizing at 10/90th percentile or robust statistical method (e.g., non-parametric or MM estimation) will be used instead. Missing data will be handled as described above for the primary analyses.

Non-inferiority

CGM-measured % below 54 mg/dL will be a non-inferiority test with limit 1%. Since non-inferiority is typically framed in terms of a one-sided test, it is worth noting that the left half of a two-sided test at alpha = 0.05 gives the same rejection region as a one-sided test at alpha = 0.025. Therefore, reporting a two-sided 95% confidence interval will provide flexibility to also test for inferiority if non-inferiority cannot be declared, or superiority if non-inferiority is declared. Note that since all tests for non-inferiority/inferiority/superiority are based on the same confidence interval, the overall type 1 error rate is maintained at 5%. The following figure shows examples of the inference to be drawn for a non-inferiority analyses based on the two-sided 95% confidence interval in various scenarios:



Hierarchical Hypothesis Testing

To preserve the overall type 1 error for non-inferiority in CGM-measured % below 54 mg/dL, a hierarchical testing procedure will be used. A formal statistical assessment of non-inferiority (limit 1%)

for time below 54 mg/dL will only be performed if the primary analysis for CGM-measured % in range 70-180 mg/dL described above results in a statistically significant result ($p < 0.05$).

If a non-significant result is encountered for CGM-measured % in range 70-180 mg/dL, then no formal statistical hypothesis testing for non-inferiority CGM-measured % below 54 mg/dL will be performed, and analysis of this variable become exploratory. As mentioned above in section 5.2.2, superiority in CGM-measured % below 54 mg/dL is additionally listed as an exploratory outcome.

Regardless of the results of the hierarchical testing, summary statistics appropriate to the distribution of CGM-measured % below 54 mg/dL will be tabulated by treatment period. A two-sided 95% confidence interval for the treatment period and their difference will also be calculated. However, a confidence interval that excludes zero will not be considered a statistically significant result if for CGM-measured % in range 70-180 mg/dL failed reach statistical significance.

8.2 Secondary CGM Metrics

The analyses for the continuous CGM-measured outcomes, including missing data, will parallel those mentioned above for the primary outcome.

For the secondary and binary CGM outcomes listed above, risk-adjusted percentages by treatment period and the difference between treatment arms will be computed from a logistic regression model [1]. The logistic regression will adjust for the same factors mentioned above for the primary outcome (i.e., period, site, and baseline value).

8.3 HbA1c

Summary statistics [mean ± SD or median (IQR)] will be reported for the central lab HbA1c at randomization and 13 weeks in each treatment period.

Lab HbA1c values at 13 weeks will be compared between the two treatment periods using a repeated linear model while adjusting for period and site (random effect). Inclusion of the randomization HbA1c value as a third observation for each participant in the model gives a variance reduction analogous to adjusting for it as a covariate. The model will account for correlated data from the same participant. All participants with at least one HbA1c lab measurement at 13 weeks in one of the two periods will be included in the analysis. Regression diagnostics will be employed analogous to as described in Section 7 for the primary outcome.

For the binary HbA1c outcomes listed above, similar analyses will be done as described above for the binary CGM outcomes.

8.4. Insulin Analyses

Summary statistics appropriate to the distribution for total daily insulin and the bolus:basal ratio will be given by treatment group at baseline and at 13 weeks in each one of the two crossover periods. All participants with at least one CFR-reported insulin metric at 13 weeks in one of the two periods will be included in the analysis. Regression diagnostics will be employed analogous to as described in Section 7 for the primary outcome.

8.5. Weight and Body Mass Index Analyses

Summary statistics appropriate to the distribution for weight and BMI will be given by treatment group at baseline and 13 weeks in each one of the two crossover periods. All participants with at least one reported weight metric at 13 weeks in one of the two periods will be included in the analysis. Regression diagnostics will be employed analogous to as described in Section 7 for the primary outcome.

8.6. Questionnaires

For each questionnaire, mean ± SD values or percentiles appropriate to the distribution will be given by randomized period for the total score and each subscale at baseline and at 13 weeks in each period.

For questionnaires administered in both randomization periods, comparisons will be made using similar longitudinal models as described above for the primary outcomes. Separate models will be run for the total score and each of the subscales listed above, and the models will adjust for baseline questionnaire score. A point estimate, confidence interval and p-value will be given for the treatment period difference at 13 weeks.

9. Safety Analyses

All enrolled participants will be included in these analyses and all their safety events up to the second and final 13-week visit will be reported.

The circumstances of all reportable cases of the following will be summarized and tabulated by treatment period:

- Severe hypoglycemia
- Diabetic ketoacidosis
- Ketone events defined as day with ketone level ≥1.5 mmol/L
- Worsening of HbA1c from start of the period to 13 weeks by >0.5%
- Other serious adverse events (SAE) and serious adverse device events (SADE)
- Adverse device effects (ADE)
- Unanticipated adverse device effects (UADE)

Statistical analyses to compare rates of severe hypoglycemia and rates of diabetic ketoacidosis between treatment arms will only be performed if there are at least 10 events after randomization. If yes, the numbers will be compared between the two treatment periods using a robust Poisson regression. The amount of follow up will be included as an offset covariate to compare the rates.

Comparison of safety outcomes between the two treatment periods only include those events occurring on or after randomization until the final 13-week visit.

The analyses for the two continuous CGM-measured outcomes will parallel those mentioned above for the primary outcome.

The CGM metrics listed in the above efficacy section 5.2.2. also serve as safety measures if there is worsening in the intervention group.

Any pre-randomization adverse events will be tabulated separately and will include participants who were never randomized.

10. Device Issues

Reported device malfunctions or other issues for each type of study device (e.g., closed loop system, CGM, blood glucose meter)

11. Protocol Adherence and Retention

The following tabulations and analyses will be performed by treatment period to assess protocol adherence for the study:

- Number of protocol and procedural deviations per participant along with the number and percentage of participants with each number of deviations
- Number of protocol and procedural deviations by severity with brief descriptions listed
- Flow chart accounting for all participants at all scheduled visits and phone contacts post treatment initiation to assess visit and phone completion rates – will include missing, in and out of window visit and phone contacts
- Number of and reasons for unscheduled visits and phone calls
- Number of participants who stopped treatment and reasons

12. Baseline Descriptive Statistics

Baseline demographic and clinical characteristics of the cohort of all randomized participants will be summarized in a table using summary statistics appropriate to the distribution of each variable. Will include:

- Age
- Gender
- Race/ethnicity
- Income, education, and/or insurance status
- Diabetes duration
- Daily SMBG at enrollment
- HbA1c

- BMI
- C-peptide
- Participant-reported number of SH and DKA 12 months prior to the start of the study

13. Other Tabulations

Individual listings for each participant will include the following:

- Treatment period order, age, gender, race/ethnicity, duration, height, weight, and BMI
- Study related information (like enrollment and randomization dates, enrollment and randomization HbA1c, randomization C-peptide, status, run-in requirement)
- Previous SMBG use, non-insulin medications, devices months used and manufacturer
- Past SH and DKA events
- Physical exam results
- Income, education, and insurance
- Pre-existing medical conditions other than diabetes
- Medication at enrollment
- Baseline glucose metrics

The following tabulations and analyses will be performed by treatment period:

- Percent sensor use – overall and by month
- The daily frequency of downloaded SBGM use - overall and by month

The following tabulations and analyses will be performed for MPC period only:

- Percent time in different operational modes - overall and by month
- Percent change of clinical and controller adaptation parameters throughout closed-loop use
- Rate of different failure events and alarms per 24 hours recorded by the system – overall and by month

14. Planned Interim Analyses

No formal interim efficacy analyses are planned for this study.

The DSMB will review all cases of severe hypoglycemia and diabetic ketoacidosis irrespective of device relationship, all device related SAEs, and all UADEs at the time that they occur during the study and will review compiled safety data at periodic intervals. The DSMB can request modifications to the study protocol or suspension or outright stoppage of the study if deemed necessary.

The data to be reviewed periodically by DSMB will include information regarding the following:

- Status of randomized participants
- Recruitment rates by month and by site
- Baseline demographic and clinical characteristics
- Dropped participants and reasons for discontinuing
- Protocol deviations
- Device issues
- Scheduled and unscheduled visits and contacts
- Frequency of CGM and system use over time and by site
- Reportable adverse events as described in Section 10 of the protocol
- CGM-based hypo- and hyper-glycemic events during baseline and all available post randomization data

15. Subgroup Analyses

In exploratory analyses, the primary outcome (i.e. CGM-measured % in range 70-180 mg/dL) will be assessed separately in various subgroups and for continuous variables according to the baseline value as defined below. Modification of the treatment effect by baseline variables will be assess by including an interaction term in the model described above for the primary analysis.

Interpretation of subgroup analyses will be viewed with caution, particularly in the absence of an overall significant difference. For continuous variables, results will be displayed in subgroups based on cutpoints although the analysis will utilize the variable as continuous. If there is insufficient sample size in each subgroup, the cutpoints for continuous measures may be adjusted per the observed distribution of values. Cutpoint selection for display purposes will be made masked to the outcome data and generally based on means or medians.

- Baseline HbA1c
- Baseline CGM time 70-180 mg/dL
- Age
- Sex
- Race

16. Multiple Comparison/Multiplicity

Primary Analysis

Since there will be a single comparison for the primary outcome (CGM-measured % 70-180 mg/dL), no adjustment is needed.

Secondary Hierarchical Analysis

The hierarchical testing procedure described above in Section 8.1 will be used to control the overall type 1 error for the primary outcome and the other key secondary outcome (CGM time spent <54 mg/dL) identified above.

All Other Secondary Analyses

For comparison of other efficacy endpoints considered secondary, the false discovery rate (FDR) will be calculated using the Benjamini-Hochberg method adapted using the two-stage test. FDR adjusted p-values will be calculated separately for the following categories:

- CGM metrics over 24hr
- CGM metrics during daytime periods
- CGM metrics during nighttime periods
- HbA1c analyses
- Insulin, weight, BMI
- Questionnaires
- Subgroup analyses

P-values from safety analyses, sensitivity analyses and per-protocol analyses will not be adjusted for multiple comparisons.

17. Exploratory analyses

No p-values will be calculated for these analyses.

The following metrics will be reported with the appropriate descriptive statistics by the associated MPC operational mode:

- % below 70 mg/dL
- % above 180 mg/dL
- % time in range 70-180 mg/dL

- mean glucose
- coefficient of variation

18. Additional Tabulations and Analyses

- 24 hours profiles with mean (or medians) and quartiles lines and 4-week interval boxplots by treatment arms for:
    - % below 70 mg/dL
    - % above 180 mg/dL
    - % time in range 70-180 mg/dL
    - mean glucose
    - coefficient of variation

- During the MPC period, the closed-loop system will include an automated algorithm adaptation approximately once every week. Tabulations, 24 hours profiles, and boxplots over each one of the approximately 13 adaptation periods will be generated for:
    - % below 70 mg/dL
    - % above 180 mg/dL
    - % time in range 70-180 mg/dL
    - mean glucose
    - coefficient of variation

19. Pilot Phase

After at least 5 total participants from at least 3 different sites have completed the pilot phase, the data will be reviewed by DSMB. If there are no significant safety, usability, or system functionality issues, as specified in Section 3.4 of the protocol, the randomized crossover trial may begin for sites who have completed at least 1 pilot phase participant. The following descriptive analyses will be generated:

- Baseline demographics and clinical characteristics (Section 12)
- Device issues (Section 10)
- Safety (Section 9):
    - Severe hypoglycemia
    - Diabetic ketoacidosis

- Ketone events defined as day with ketone level ≥1.5 mmol/L
- Other serious adverse events (SAE) and serious adverse device events (SADE)
- Adverse device effects (ADE)
- Unanticipated adverse device effects (UADE)

- CGM-metrics during the entire pilot phase and baseline (as defined above in Section 4.3):
  - % <70 mg/dL
  - % <60 mg/dL
  - % <54 mg/dL
  - low blood glucose index
  - hypoglycemia events (defined as at least 15 consecutive minutes <54 mg/dL)
  - % in range 70-180 mg/dL
  - % in range 70-140 mg/dL
  - mean glucose
  - glucose variability measured with the coefficient of variation
  - glucose variability measured with the standard deviation
  - % >180 mg/dL
  - % >250 mg/dL
  - % >300 mg/dL
  - high blood glucose index
  - CGM-measured hyperglycemic events (≥15 minutes with glucose concentration >300 mg/dL)
- All these factors will be given at the participant level and with summary statistics appropriate to the distribution
- Other analyses as requested by DSMB

None of the pilot phase data will be used in the RCT analyses.

Reference:

1. Kleinman LC, Norton EC. "What's the Risk? A simple approach for estimating adjusted risk measures from nonlinear models including logistic regression". Health Serv Res 2009;44:288-302.