Official Title

Home Use of MD-Logic Automated Insulin Delivery System: Safety and Efficacy

Brief Title: Fuzzy Logic Automated Insulin Regulation (FLAIR)

NCT 03040414

Date: April 10, 2020

Fuzzy Logic Automated Insulin Regulation (FLAIR)

Statistical Analysis Plan

Version 1.1

April 10th, 2020 (Based on Protocol Version 7.0)

Revision History

The following table outlines changes made to the Statistical Analysis Plan.

Version Number	Author	Approver	Effective Date	Study Stage
1.0	Nathan Cohen	Peter Calhoun	05/16/2019	Protocol development
1.1	Ryan Bailey	Peter Calhoun	04/10/2020	Follow-Up

Version Number	Revision Description
1.1	Added a binary indicator for time in range (70-180 mg/dL) >70% and time below 54 mg/dL <1% as an exploratory outcome. Modified section 14.2 to clarify that gaps in the pump data greater than 90 minutes will not be used in the calculation of time in auto mode. Updated subscales for the diabetes distress survey in section 14.3. Added section 14.4 on pump settings.

Author: _____

Senior Statistician: _____

Jaeb Principal Investigator: _____

Investigator: _____

1. Statistical and Analytical Plans

This document outlines the statistical analyses to be performed for the FLAIR study. The approach to sample size and statistical analyses for this study are detailed below.

This is a multicenter, randomized, two period crossover study to compare the efficacy and safety of an automated insulin delivery (AID) system with a proportional-integral-derivative (PID) algorithm (Minimed 670G 3.0; abbreviated as 670G) to an AID system with a PID algorithm enhanced with a Fuzzy Logic algorithm (Minimed 670G 4.0 Advanced Hybrid Closed-Loop; abbreviated as AHCL) in patients with type 1 diabetes. Approximately 112 individuals with type 1 diabetes aged 14-<30 years will enter the RCT, with the expectation that a minimum of 100 will complete both periods of the study. All participants will receive both interventions, and the order of receiving them will be randomized based on a 1:1 ratio. Each period will last for 12 weeks.

Each of the two study periods consists of three visits and four phone contacts. Randomization will be preceded by a run-in period where subjects must demonstrate competency and compliance in using the study pump and CGM device. The run-in period will be divided into a pump and a pump + CGM run-in phase. During the run-in period, a contact will occur at one week, and a visit will occur at two weeks. The length of the pump and pump + CGM run-in phase may vary depending on the participant's comfortability using the devices.

2. Statistical Hypotheses

In this study, there are two co-primary outcomes:

- Superiority in the percentage of time spent with sensor glucose levels above 180 mg/dL during the daytime (6:00 AM to 11:59 PM)
- Non-inferiority in the percentage of time spent with sensor glucose levels below 54 mg/dL calculated over the full 24 hours of the day, with a non-inferiority limit of 2% based on clinical judgment.

Separate analyses will be performed for each co-primary outcome, and both analyses must produce a statistically significant result in order to declare benefit of the intervention.

For each of these analyses, the two corresponding pairs of null and alternative hypotheses are as follows:

<u>Hyperglycemic Outcome:</u>

- *Null Hypothesis:* There is no difference in the percentage of time spent with sensor glucose levels above 180 mg/dL between the two treatments during the daytime for the entire study population.
- *Alternative Hypothesis:* There is a nonzero difference in the percentage of time spent with sensor glucose levels above 180 mg/dL between the two treatments during the daytime for the entire study population.

<u>Hypoglycemic Outcome:</u>

- *Null Hypothesis:* There is a mean difference of at least 2% in the percentage of time spent with sensor glucose levels below 54 mg/dL between the AHCL system and the 670G system over the full 24 hours of the day for the entire study population.
- *Alternative Hypothesis:* There is a mean difference of less than 2% in the percentage of time spent with sensor glucose levels below 54 mg/dL between the AHCL system and the 670G system over the full 24 hours of the day for the entire study population.

3. Sample Size

The sample size calculation for this study is detailed in a separate document (in addition to the protocol) and summarized below. Sample size was calculated based on the following assumptions:

- 90% power
- 5% true absolute reduction in percent time > 180 mg/dL during the daytime with a standard deviation of paired differences of 12% and a two-sided Type 1 error rate of 5%
- Non-inferiority limit of 2% for the treatment group comparison of percent time < 54 mg/dL with a standard deviation of paired differences of 1.3% and a one-sided Type 1 error rate of 2.5%

In the sample size calculation, an adjustment was made to account for the two co-primary analyses. Analyses of both primary outcomes must have a significant result ($p < 0.05$) to declare benefit of the intervention. The two co-primary outcomes therefore do not inflate the Type 1 error, but requiring both to be significant can inflate the Type 2 error. Therefore, a Bonferroni correction was applied to the power calculations by adding the Type 2 error rate (i.e., 100% minus power) for the two analyses. This represents a conservative estimate for composite power as the probability that both tests will reject the null hypothesis.

With the above assumptions, sample size was calculated to be 65. For entry into the trial, this was increased to 112 to provide greater precision for the primary and secondary analyses.

4. Outcome Measures

Co-Primary Efficacy Endpoints*:

1) Percentage of daytime (6:00 AM to 11:59 PM) spent with sensor glucose levels above 180 mg/dL (superiority)
2) Percentage of time spent with sensor glucose levels below 54 mg/dL over the full 24 hours of the day (non-inferiority)

*Both co-primary efficacy endpoints are calculated over each treatment period for the full 24 hours, daytime (6:00 AM – 11:59 PM), and nighttime (12:00 AM – 5:59 AM). They are also calculated for each post-meal period measured for up to three hours after the start of a meal.

Secondary Efficacy Endpoints:

CGM Metrics

Overall Glucose Control

1) Mean glucose
2) Time in range 70 and 180 mg/dL

3) Time in range 70 and 140 mg/dL

Hyperglycemic Outcomes

4) Time above 180 (daytime is a co-primary outcome) and 250 mg/dL
5) Peak glucose and change from baseline to the peak (only calculated for post-meal periods)

Hypoglycemic Outcomes

6) Time below 54 (over 24 hours is a co-primary outcome) and 70 mg/dL

Glycemic Variability

7) Coefficient of variation

Insulin Delivery Metrics

8) Total Daily Insulin Units
9) Basal Daily Insulin Units
10) Bolus Daily Insulin Units
11) Total Daily Insulin Units per Kilogram of Weight
12) Total Basal Insulin Units per Kilogram of Weight
13) Total Bolus Insulin Units per Kilogram of Weight

HbA1c Metrics

14) HbA1c

Clinical Metrics

15) BMI

Additional Exploratory Endpoints:

CGM Metrics:

Continuous Outcomes:

1) Time below 50 and 60 mg/dL
2) Time above 300 mg/dL
3) Area under the curve of sensor glucose levels above 180 mg/dL
4) Area over the curve of sensor glucose levels below 70 mg/dL
5) Standard deviation of sensor glucose levels

Binary Outcome:

1) Time in range (70-180 mg/dL) >70% and time below 54 mg/dL <1%.

HbA1c Metrics:

Binary Outcomes:

1) Improvement in HbA1c $\geq$ 0.5%
2) Improvement in HbA1c $\geq$ 1.0%
3) Relative reduction in HbA1c $\geq$ 10%
4) HbA1c < 7.0%

All secondary efficacy and exploratory CGM metrics will be calculated over each treatment period for the full 24 hours, daytime (6:00 AM – 11:59 PM), and nighttime (12:00 AM – 5:59 AM). They also will be calculated for each post-meal period measured for up to three hours after the start of a meal. The same will be done for all secondary efficacy insulin endpoints. It is

important to note that all hypoglycemic outcomes are considered both efficacy and safety outcomes.

4.1 Calculation of the Co-Primary Outcomes

As noted above, the co-primary outcomes are:

- Daytime Hyperglycemia (percent CGM readings above 180 mg/dL)
- Hypoglycemia (percent CGM readings below 54 mg/dL; daytime/nighttime combined)

For each co-primary outcome, a single value will be calculated for each subject for each period by pooling all CGM readings in the first 84 days of each period or up to and including the 12-week visit date (whichever comes first).

In these calculations, each period will start at the time that the subject turns on auto mode. If the exact time is not known, then the period will start at midnight the day after the subject initiates it. We expect auto mode to be initiated 6-10 days after the randomization visit in period 1 and 6-10 days after the first day of period 2. Participants who are 670G auto mode users at screening may initiate auto mode at the beginning of a study period in which the 670 3.0 HCL pump is being used. The end of the period will be defined as either 11:59 PM on the 12-week visit date or exactly 2,016 hours (i.e. 24 hours/day*84 days) after auto mode is turned on, whichever comes first. If the exact time that auto mode is turned on is unknown, then it will be defined as 11:59 PM on the 12-week visit date or 11:59 PM on the 84th day of the period, whichever comes first. Data will not be truncated due to protocol deviations.

CGM readings will be used according to the time of day as follows:

Hyperglycemic Outcome

CGM readings will be limited to daytime values (between 6:00 AM and 11:59 PM, inclusively).

Hypoglycemic Outcome

All CGM readings will be used with no restriction according to time of day (i.e., all 24 hours).

The baseline value of each co-primary outcome will be calculated by pooling all CGM readings from the last 14 days of the CGM run-in phase. As mentioned above, only readings occurring between 6:00 AM and 11:59 PM (inclusively) during this period will be included in the calculation of the hyperglycemic outcome. If less than 336 hours of CGM data are available during this period, then an additional preceding day will be pooled one at a time until either 336 hours are available or 30 total days are reached (whichever comes first).

4.1.1 Replication of Co-Primary Analyses Omitting the First Four Weeks of Each Period

The co-primary analyses will be replicated after omitting the first four weeks of each crossover period. In these analyses, the first and last day of each crossover period will be the same as defined above. Therefore, each period in these analyses will start exactly 672 hours (i.e. 24 hours/day*28 days) after the start of the period as defined above. For each co-primary outcome metric, a single value will be calculated by pooling all CGM data from this starting point to the final day of the period, inclusively. Only daytime data (6:00 AM to 11:59 PM) will be included in the calculation of the hyperglycemic outcome.

Other than the differences described above, replication of co-primary analyses omitting the first four weeks of each period will be the same as the main analyses.

4.2 Calculation of Secondary Outcomes

All secondary CGM metrics calculated over the full 24 hours of the day will be computed analogously as described above for the primary analysis. Baseline CGM metrics will also be calculated analogously. The same subjects will be included as in the primary analysis. A similar calculation will also be done for the daytime and nighttime metrics at baseline and during each period, except readings will only be pooled if they fall within the appropriate time windows (6:00 AM – 11:59 PM for daytime and 12:00 AM – 5:59 AM for nighttime).

Secondary insulin metrics calculated per kilogram of weight will be computed by dividing total units by weight measured at randomization.

Missing data will not be imputed for any of the secondary analyses in this study, and data will not be truncated due to protocol deviations. Secondary analyses will only include available cases.

4.2.1 Calculation of Post-Meal Secondary Outcomes

The start time of each meal will be defined as the time that the carbohydrate is entered into the pump (t=0). If the next carbohydrate is entered into the pump less than three hours after t=0, then the data will be pooled only from t=0 to the start time of the next meal.

We will assume that each subject will consume up to three meals per calendar day. If more than three carbohydrates are entered into the pump in one day, then only the largest three will be included in post-meal analyses. If less than three are consumed in one day, then only the meals available for that day will be included.

The secondary CGM outcomes calculated during the post-meal periods will be computed by pooling all CGM readings occurring up to three hours after the start of the meal. For the insulin outcomes, all insulin data occurring during this time will be pooled. In the analyses, data will be pooled across all meals within a given period to calculate post-prandial outcomes for each period for each subject. Therefore, analyses of post-prandial endpoints will not be performed on a meal level.

For peak glucose and change from baseline to the peak, the mean will be computed across all meals to calculate the outcome. This will be done to ensure that the peak is not taken to be a single reading from the entire period, because this value would be expected to be extremely large for every subject.

At least 72 hours of post-prandial CGM data will be required to be included in secondary post-meal CGM or insulin analyses.

5. Analysis Datasets and Sensitivity Analyses

5.1 Analysis Cohorts

- All analyses will be performed on an intention-to-treat basis with each period included in the treatment arm assigned by randomization regardless of the actual treatment received.

The primary and secondary efficacy analyses for this study will include all randomized subjects with at least 72 hours of CGM data in either period.

- All subjects with at least 72 hours of insulin data in either period will be included in insulin analyses.
- A per-protocol analysis for the co-primary outcomes will be conducted by limiting to subjects who used the CGM at least 80% of the time in both periods and who used auto mode for at least 80% of the time in each treatment period, unless the sample size for the analysis is at least 95% of the sample size in the primary analysis.
- All participants will be included in safety analyses, regardless of whether the study was completed.

6. Analysis of the Primary Efficacy Endpoints

6.1 Missing Data

Missing data will not be imputed for the primary analysis in this study.

6.2 Statistical Methods

For each co-primary outcome, mean ± SD or summary statistics appropriate to the distribution will be reported by treatment period. A repeated measures least squares regression model with an unstructured covariance structure will be fit for the two co-primary outcomes. Each of these models will adjust for period, pre-study 670G system use, and HbA1c at randomization as fixed effects. The models will include 3 time points: (1) baseline, (2) period 1 outcome, and (3) period 2 outcome. Prior 670G system use in auto-mode at the time of enrollment and HbA1c at randomization will be included as covariates in the model because they will be used to stratify randomization.

A separate model also will be built for each co-primary outcome with the inclusion of a period by treatment interaction term to assess for the presence of a carryover effect.

All covariates obtained on a continuous scale will be entered into the models as continuous variables, unless it is determined that a variable does not have a linear relationship with the outcome. In such a case, categorization and/or transformation will be explored. All p-values will be two-sided. For the co-primary analyses, the intervention will be deemed effective only if the null hypotheses for both co-primary outcomes are rejected.

Standard residual diagnostics will be performed for all analyses. If values are highly skewed, then a transformation or nonparametric methods will be used instead.

6.2.1 Statistical Methods for Percent Time > 180 mg/dL during the Daytime (6:00 AM – 11:59 PM)

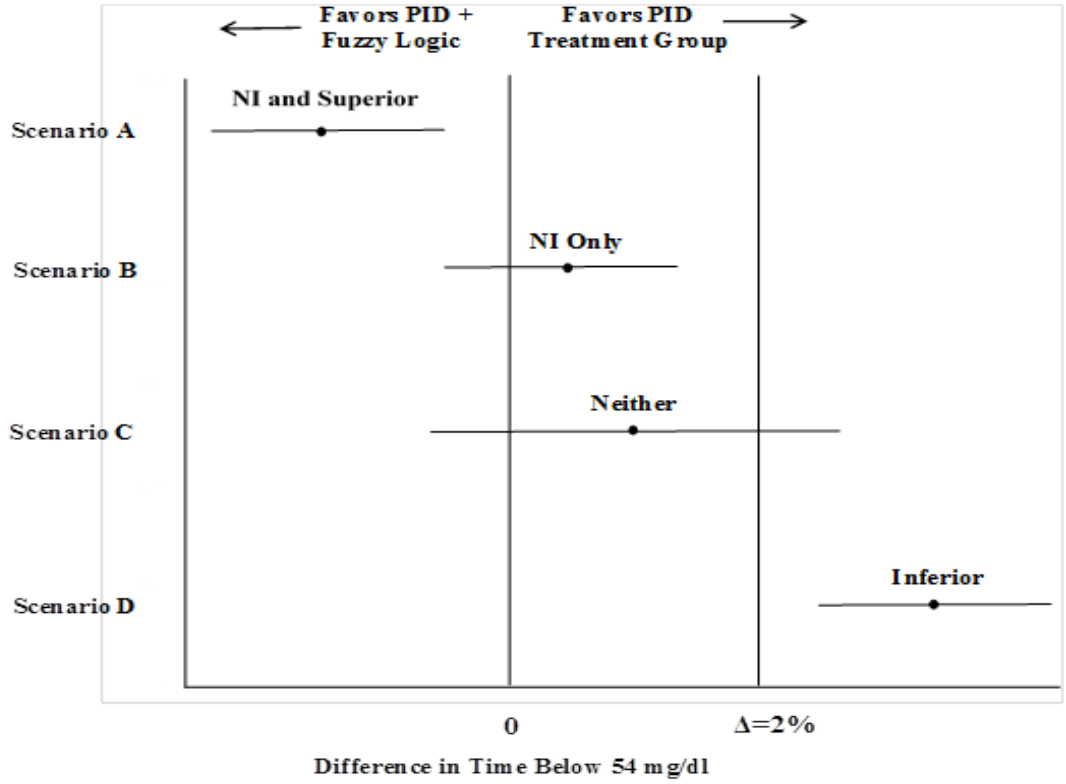
For percent time > 180 mg/dL during the daytime, a 95% confidence interval will be reported for the difference between the two treatment arms based on the least squares model. Residual values will be examined for an approximate normal distribution. If values are highly skewed, then a transformation or nonparametric methods will be used instead. Based on previous data, we do not expect that a transformation will be necessary. A two-sided p-value will be reported, and a 5% significance level will be used to declare statistical significance.

6.2.2 Statistical Methods for Percent Time < 54 mg/dL

For percent time < 54 mg/dL, a two-sided 95% confidence interval will be reported for the difference between the AHCL arm and the 670G arm based on the least squares model. Non-inferiority will be assessed by comparing the upper limit of this confidence interval to a non-inferiority limit of 2% (approximately 30 minutes per day). Residual values will be examined for an approximate normal distribution. If values are highly skewed, then a transformation or nonparametric methods will be used instead. A two-sided p-value will be reported, and a 5% significance level will be used to declare statistical significance.

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 the inference to be drawn for various scenarios of the two-sided 95% confidence interval:

Interpretation of Various Scenarios of Non-inferiority (NI) and Superiority for Treatment Group Comparison of Time Below 54 mg/dl



* NI refers to non-inferiority of PID + Fuzzy Logic compared with PID with a non-inferiority limit of Δ=2% time below 54 mg/dl.

6.2.3 Boxplots of Monthly Trends

Monthly trends for each co-primary outcome will be evaluated using boxplots. For each co-primary outcome, we expect to have 12 weeks' worth of CGM data for each subject for each treatment arm. This amounts to approximately three months' worth of data.

As mentioned above, each period will be defined from the time the subject initiates auto mode until either 84 days have been included or the 12-week visit occurs (whichever comes first). Therefore, the number of days included in each period may vary across subjects. In the monthly boxplots, the first and second months will respectively include all CGM readings from the first 672 hours (i.e. 24 hours/day*28 days) and the next 672 hours. All remaining CGM readings will be included in the third month.

After partitioning the data into three separate months in each period, a single value for each month will be calculated for each co-primary outcome for each treatment arm by pooling all CGM readings falling within the three disjoint windows of days that correspond to each of the three months. For the hyperglycemic outcome, only readings falling between 6:00 AM and 11:59 PM will be included in the calculation. These calculated values will be used to generate boxplots displayed by treatment arm and period for each of the three months over time. Only subjects with at least 72 hours of CGM data in all three months will be included in the monthly boxplots.

6.2.4 Summary Statistics by Hour of the Day

Summary statistics appropriate to the distribution for each co-primary outcome will also be displayed by treatment arm for each hour of the day. For each co-primary outcome for each period, a single value for each hour of the day will be computed by pooling all CGM readings falling within the specified hour. The hourly intervals will be left-closed and right-open (defined as 00:00-<01:00, 01:00-<02:00, etc.). The start and end of each period will be defined in the same way as described above. For the hyperglycemic outcome, summary statistics will not be reported for the nighttime hours (12:00 AM to 5:59 AM). To be included in these analyses, subjects must have at least one hour of CGM data in each hour interval for at least one period.

7. Analysis of the Secondary Endpoints

7.1 Secondary CGM Outcomes

For each secondary CGM outcome, mean ± SD or summary statistics appropriate to the distribution will be reported by treatment arm. All secondary CGM metrics calculated over the full 24 hours of the day, daytime, and nighttime will be compared between treatment arms using the same repeated measures least squares regression model described above for the co-primary outcomes.

Replication of the primary outcome omitting the first four weeks of each period (as described in section 4.1.1) will be considered a secondary analysis. In this analysis, the baseline value included in the model will be the same baseline value as in the primary analysis.

The CGM outcomes listed in Section 4 under "Additional Exploratory Endpoints" will be tabulated separately by treatment arm with no formal statistical comparison.

7.2 Secondary Insulin Outcomes

For each secondary insulin outcome, mean ± SD or summary statistics appropriate to the distribution will be reported by treatment arm. All secondary insulin metrics calculated over the full 24 hours of the day, daytime, and nighttime will be compared between treatment arms using the same repeated measures least squares regression model described above for the co-primary outcomes.

The insulin outcomes listed in Section 4 under "Additional Exploratory Endpoints" will be tabulated separately by treatment arm with no formal statistical comparison. For these outcomes, change will be calculated from the start of the period.

7.3 Secondary HbA1c Outcomes

For HbA1c, a longitudinal model adjusting for period and pre-study 670G system use as fixed effects will be used to compare treatments. The model will include three time points: (1) baseline, (2) period 1 follow-up, (3) period 2 follow-up.

The binary HbA1c outcomes listed in Section 4 under "Additional Exploratory Endpoints" will be tabulated separately by treatment arm with no formal statistical comparison. For these outcomes, improvement in HbA1c will be assessed from the start of the period.

7.4 Secondary Clinical Outcomes

For BMI, the same model described in section 7.3 for HbA1c will be used for a treatment arm comparison. The model will additionally adjust for HbA1c at randomization, age, and gender as fixed effects.

7.5 Boxplots of Monthly Trends

For each of the following secondary outcomes computed over the full 24 hours of the day, monthly trends will be evaluated using boxplots:

- Percentage of time spent with sensor glucose levels between 70 and 180 mg/dL
- Mean of sensor glucose levels
- Total Daily Insulin

For each of these secondary outcomes, we expect to have 12 weeks' worth of data for each subject for each treatment arm. This amounts to approximately three months' worth of data. Each period will be partitioned into three disjoint months using the same approach described in section 6.2.3. A single value for each of these three months will be calculated for each of these three secondary outcomes for each treatment arm by pooling the CGM readings corresponding to each month. These calculated values will be used to generate boxplots displayed by treatment arm and period for each of the three months over time.

7.6 Analysis Windows

Only HbA1c and BMI measurements obtained within the following analysis windows at each time point included in the analyses:

- 12 weeks: ±14 days from the 84^{th} day of the period

8. Safety Analyses

8.1 Definitions

Reportable adverse events for this protocol include all serious adverse events (SAE), adverse device effects, adverse events occurring in association with a study procedure, severe hypoglycemia (SH), and diabetic ketoacidosis (DKA) as defined in the protocol.

8.2 Adverse Events Summary

All episodes of SH and of DKA along with any other reportable adverse events will be listed by treatment arm from the time of enrollment to the last study visit. Separate listings will be provided for pre-randomization events and for events occurring during each treatment arm.

8.3 Comparison of Safety Outcomes between Treatment Arms

Each of the following safety outcomes will be tabulated by treatment arm:

- Number of adverse events/participants with at least one event
  - Number of events meeting definition of ADE
- Number of serious adverse events/participants with at least one event
  - Number of SAEs meeting definition of UADE
- Listing of hospitalizations and reasons for the hospitalization
- Severe Hypoglycemic Events
  - Number of severe hypoglycemic events as defined in the protocol/number of participants with an event
  - Number of severe hypoglycemic events associated with seizure or loss of consciousness/number of participants with at least one event
- Diabetic Ketoacidosis
  - Number of diabetic ketoacidosis events, as defined in the protocol/number of participants with an event

Of the outcomes listed above, only the ones involving SH or DKA events will be compared between treatment arms, if there are at least five events across the two treatment arms. The others will only be tabulated by treatment arm.

All of the above safety outcomes will be tabulated for all subjects (including dropouts and withdrawals) for their entire period of study participation. Any adverse events that occurred prior to the start of period 1 (as defined above) will not be included in the treatment arm comparisons listed above. Each period will inclusively consist of all days in between the start of the period and the 12-week visit date.

If there are enough SH and DKA events associated with each treatment arm to allow for formal statistical modeling, then treatment arm comparisons will be performed for the SH and DKA outcomes listed above. For binary outcomes, repeated measures logistic regression models adjusting for period, pre-study 670G system use, HbA1c at randomization, and whether the subject experienced any events in the 12 months prior to enrollment will be used to compare treatment groups. If the model does not converge, a mid-p binomial test will be considered. For counts, the treatment groups will be compared using repeated measures Poisson regression models adjusting for period, pre-study 670G system use, HbA1c at randomization, and whether the subject experienced any events in the 12 months prior to enrollment.

9. Adherence and Retention Analyses

The following tabulations will be performed by treatment arm to assess protocol adherence for this study:

- Number of protocol deviations by severity with brief descriptions listed for each treatment arm
- Number of and reasons for unscheduled visits and phone calls for each treatment arm

Additionally, the following two flowcharts will be created:

- Flowchart displaying all dropouts prior to and post randomization in addition to treatment arm completion to assess dropout rates
- Flowchart displaying the number of subjects completing and missing each visit by treatment arm to assess visit completion rates

10. Baseline Descriptive Statistics

Baseline demographic and clinical characteristics of the cohort of all randomized subjects will be summarized in a table. Descriptive statistics will be tabulated overall and by randomization group. For continuous variables, summary statistics appropriate to the distribution will be given. For discrete variables, number and percentage will be reported for each category.

The following baseline CGM metrics will be included in a separate table:

- Percentage of time in the target range (70-180 mg/dL)
- Mean glucose
- Coefficient of Variation
- Percentage of time below 54 mg/dL and below 70 mg/dL
- Percentage of time above 180 mg/dL and above 250 mg/dL
- Percentage of time above 180 mg/dL during the daytime (6:00 AM – 11:59 PM)
- Percentage of time above 180 mg/dL during the nighttime (12:00 AM – 5:59 AM)

These baseline metrics will be calculated in the manner described in sections 4.1 and 4.2. Additionally, each of these baseline metrics will be calculated for the post-meal periods as described in section 4.2.1, and will be reported in another table.

11. Planned Interim Analyses

No formal interim analyses or stopping guidelines are planned for this study.

The DSMB will review data for safety per the DSMB Standard Operating Procedures approximately every six months. The data to be reviewed will include information regarding adverse events, device issues, and protocol adherence/deviations as they may relate to participant safety.

12. Subgroup Analyses

The treatment effect in subgroups based on baseline factors will be assessed for each co-primary outcome. These analyses will be considered exploratory and will be conducted to determine whether a similar trend to the overall treatment effect is seen in these subgroups. Assessment of subgroup analyses will be interpreted with caution, particularly if the primary analysis does not show significant results. The general approach for these subgroup analyses will be to add an

interaction term for the subgroup factor by treatment into the model used for the primary analysis.

The following baseline factors will be assessed:

- Age: as a continuous variable and in 2 age groups: 14-<21 and 21-<30 years old
- Baseline HbA1c: as a continuous variable and in 2 groups: HbA1c ≤8.5% and ≥8.6%
- Insulin method/CGM: 670G user, injection user with CGM, injection user without CGM, pump user with CGM, and pump user without CGM
- T1D duration as a continuous variable
- Gender
- Race/ethnicity
- Random C-peptide level

Cut points for dichotomizing T1D duration and random C-peptide level variables will be set to be the medians for display purposes.

In addition to the assessment for interaction, the co-primary analyses and all secondary analyses described in sections 6 and 7 will be conducted separately in the two age groups and two HbA1c groups. Analyses will only be replicated within these subgroups for efficacy outcomes where p-values are being calculated. Exploratory endpoints will not be calculated within these subgroups.

13. Multiple Comparisons/Multiplicity

For the co-primary analyses, the intervention will be deemed effective only if the null hypotheses for both co-primary outcomes are rejected. Therefore, the type I error rate is not inflated and there will be no correction for multiple comparisons.

For the secondary analyses, the false discovery rate will be controlled using the adaptive Two Stage Benjamini-Hochberg procedure. False discovery rate corrections will be applied within each of the following categories:

- Secondary CGM Analyses
    - Over the full 24 hours of the day
    - During the daytime, nighttime, or post-meal periods
- Secondary Insulin Analyses, HbA1c Analyses, and Clinical Metrics
- Questionnaires
- Subgroup Analyses
    - Based on baseline factors
        - For the hyperglycemic outcome
        - For the hypoglycemic outcome
    - Within the 14-<21 age group
        - CGM Metrics
            - Over the full 24 hours of the day
            - During the daytime, nighttime, or post-meal periods
        - Insulin Metrics, HbA1c Metrics, and Clinical Metrics
    - Within the 21-<30 age group
        - CGM Metrics
            - Over the full 24 hours of the day

- During the daytime, nighttime, or post-meal periods
    - Insulin Metrics, HbA1c Metrics, or Clinical Metrics
- Within the HbA1c ≤8.5% group
    - CGM Metrics
        - Over the full 24 hours of the day
        - During the daytime, nighttime, or post-meal periods
    - Insulin Metrics, HbA1c Metrics, or Clinical Metrics
- Within the HbA1c ≥8.6% group
    - CGM Metrics
        - Over the full 24 hours of the day
        - During the daytime, nighttime, or post-meal periods
    - Insulin Metrics, HbA1c Metrics, or Clinical Metrics

A table containing the number of p-values broken down into each of these categories will also be provided.

For safety analyses, the per-protocol analysis, tests for carryover effects, and intervention compliance analyses, no adjustments for multiple comparisons will be made. For the replication of the primary analyses omitting the first four weeks of each period (overall and within the age and HbA1c groups), the p-values will be included in the appropriate daytime, nighttime, and post-meal CGM category.

14. Additional Tabulations and Analyses

14.1 Device Issues

All reported device issues will be tabulated regardless of whether they were associated with adverse events. Separate listings will be provided for pre-randomization events and for events occurring during each treatment arm. For each device issue, the following information will be listed:

- Onset date
- Type of device issue
- Description of the device issue

14.2 Amount of CGM and Auto Mode Use

Mean ± SD or summary statistics appropriate to the distribution will be tabulated for the percentage of time using the auto mode feature (over the entire period and with the CL system active) and the CGM for each treatment arm. All CL system discontinuations will be listed in a table by treatment arm.

The percentage of time using the CGM will be calculated by dividing the total number of available CGM readings by the maximum possible number of readings based on the length of the study period described in section 4.1.

The percentage of time using auto mode over the entire study period will be computed by summing the gaps in between segments when auto mode is active and dividing by the total length of the period as defined in section 4.1. For a particular segment, if auto mode is active at the start of the segment, then the gap will be included in the numerator. However, if it is only active at the end of the segment, the gap will not be included in the numerator. Gaps greater than

90 minutes long where auto mode is active at the start and end of the gap will not be counted in the numerator because it is possible that the system may not have been on during those periods.

Additionally, the percentage of time using auto mode while using the CL system will be calculated. In these calculations, the numerator will be calculated analogously as described above. However, the denominator will be the sum of the gaps in between all segments (not the total possible amount of minutes in the entire period).

For example, suppose the CL system provides the following sequence: Auto Mode Active [0:00:00], Auto Mode Active [0:05:00], Auto Mode Active [0:10:00], Auto Mode Inactive [0:14:00], Auto Mode Active [0:25:00], Auto Mode Active [0:30:00]}. In this case, the denominator in the calculation is 30 minutes. During this time, auto mode was inactive from 0:14:00 to 0:25:00, and it was active for the other 19 minutes. Therefore, the percentage of time using auto mode while using the CL system would be 100*(19/30) = 63%.

The percentages of time using CGM and the time auto mode is active (over the entire period and while using the CL system) will be compared between treatments using the same model described above for the primary outcome.

Monthly values for CGM use and auto mode use (over the entire period and while using the CL system) will also be calculated, and summary statistics will be reported for each of the three months. The three months will be defined as in section 6.2.3. Within each month, CGM use and auto mode use will be calculated analogously as described above. Treatment arm comparisons will not be done for CGM use and auto mode use for any of the three months.

Scatter plots of CGM use and auto mode use by time in range (70-180 mg/dL), time > 180 mg/dL, and time < 54 mg/dL will be produced. Summary statistics for these metrics will also be reported by CGM use and auto mode use.

14.3 Analysis of Quality of Life Data

Each of the following questionnaires will be administered to all participants who are adults or adolescents at baseline and at the end of each period:

- Diabetes Distress Scale
    - Emotional Burden Subscale
    - Physician Related Distress Subscale
    - Regimen Related Distress Subscale
    - Interpersonal Distress Subscale
- Glucose Monitoring Satisfaction Survey
    - Openness Subscale
    - Emotional Burden Subscale
    - Behavioral Burden Subscale
    - Trust Subscale
- Hypoglycemia Confidence Survey
- Diabetes Technology Attitudes Survey
- INSPIRE Survey (separate versions for adolescents and adults)

The electronic data capture system for this study will not allow the subject to submit any questionnaires until all items are completed. However, it is still possible to have missing data, because each question will provide participants with the option to not answer. Analysis will be limited to subjects who submit a questionnaire (no imputation).

For each full questionnaire and subscales, total scores (or mean scores depending on the scoring instructions for the survey) will be calculated and summary statistics appropriate to the distribution will be tabulated at each time point by treatment arm. For questionnaires and subscales where the total score is calculated, only subjects completing the entire survey (or entire subscale) will be included in the tabulations of summary statistics. Since it is possible for a subject to complete all questions in a particular subscale but not in the whole survey, it is possible that the N will differ between the full questionnaire and the subscales. The distribution of responses for each individual question at each time point will also be tabulated. Total scores for the full questionnaires and subscales will also be compared between treatments using the same model described above for the primary outcome.

14.4 Pump Settings

The relationship between the following pump settings and time in range (70-180 mg/dL), time > 180 mg/dL and time < 54 mg/dL will be examined:

1) Active Insulin Time (AIT)
2) Set point
3) Insulin-Carbohydrate Ratio (ICR)

A participant should have a value of Insulin-Carbohydrate Ratio (ICR) for each auto-bolus delivered. The average of all values will be calculated for each participant by period. An average of Active Insulin Time will be calculated for each participant and will be weighted by the amount of time spent at each value. Mean ± SD or summary statistics appropriate to the distribution will be reported for these metrics by treatment group. Summary statistics for the CGM metrics will be reported by treatment group for categories of AIT and ICR.

Set point is the daily sensor glucose target while in auto mode. The only two possible values are 100 mg/dL and 120 mg/dL. The percentage of time a participant has spent with a set point of 100 mg/dL will be calculated for each participant. Scatter plots of percent time with a set point of 100 mg/dL and time in range, time above 180 mg/dL and time below 54 mg/dL will be shown.