A Multi-Centered, 2-Period, Randomized Crossover Study to Assess the Efficacy of Predictive Low Glucose Suspend versus Sensor-Augmented Pump Therapy in the Management of Type 1 Diabetes

Tandem PLGS Pivotal Study

Statistical Analysis Plan
Randomized Crossover Trial

Version 1.2
Version Date: 1/22/2018
Author: Zoey Li
Protocol Version: 5.0

Note: The table shells are included in a separate document.

Revision History

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

Version Number	Author	Senior Statistician	Effective Date	Study Stage	Revision Description
1.0	Zoey Li	Craig Kollman	1/11/2017	Enrollment	Original Version
1.1	Zoey Li	Craig Kollman	1/18/18	Enrollment	Added more protocol adherence metrics and system usability questionnaire
1.2	Zoey Li	Craig Kollman	2/2/18	Data closeout	Removed "procedural deviations" tabulations, edited basal insulin calculation wording, and re-defined a ketosis event.

Author: Zoey Li

Digitally signed by Zoey Li
DN: cn=Zoey Li ou=Stats
Reason: I am the author of this document
Location:
Date: 2018-05-15 15:12-04:00

Senior Statistician: Craig Kollman

Digitally signed by Craig Kollman
DN: cn=Craig Kollman ou=Stats
Reason: I am approving this document
Location:
Date: 2018-02-08 10:10-05:00

Principal Investigator:

John Lum

Digitally signed by John Lum
DN: cn=John Lum ou=APP
Reason: I am approving this document
Location:
Date: 2018-02-07 09:52-05:00

This document outlines the statistical analyses to be performed for the Tandem PLGS Pivotal study. The approach to sample size and statistical analyses for this study are summarized below.

1. Study Overview

This is a multicenter, randomized, two-period crossover study to assess the efficacy and safety of a Predictive Low-Glucose Suspend system (PLGS) in comparison with Sensor-Augmented Pump therapy (SAP) in T1D patients ≥ 6 years old. Enrollment will target 90 participants. All participants will receive both interventions, and the order of receiving them will be randomized based on a 1:1 ratio.

Following screening, the study includes 5-13 clinic visits and 4-10 phone/email/text contacts for subjects completing the study. The training phases immediately following screening (pre-randomization) will consist of 10-14 days for CGM training and 14-28 days for SAP training. Either CGM Training or both training phases may be skipped at an investigator's discretion based on device use status per protocol. Randomization will be preceded by a 10-day pilot period where data collected from this period will be evaluated for system usability and predetermined safety metrics before participants ≥ 12 years of age may be randomized into the crossover trial. Immediately after randomization, subjects will enter the two 3-week study periods (21 days each) and will test one intervention per study period, using either the PLGS or SAP system first. The interventions will not be separated by a washout period.

2. Statistical Hypotheses

- *Null Hypothesis*: There is no difference in the mean time spent below 70 mg/dL over the 21-day period between the two treatment groups.
- *Alternative Hypothesis*: There is a nonzero difference in the mean time spent below 70 mg/dL over the 21-day period between the two treatment groups.

3. Sample Size

The study is projected to include at least 90 subjects. Of these study participants, at least 45 will be < 18 years old, and at least 45 will be ≥ 18 years old. Of the 45 participants that are < 18 years old, at least 15 participants will be 12 to 17 years old, and at least 15 will be 6 to 11 years old. The 90 study participants will also include a minimum of 15 pump users, 15 multiple daily injection of insulin (MDI) users, 15 CGM naïve users (may be pump or MDI users), and 15 experienced CGM users (may also be pump or MDI users).

The subjects will be randomized to the order of the interventions based on a 1:1 ratio. Randomization will not be stratified.

4. Outcome Measures

4.1 Primary Efficacy Endpoint:

Percent of CGM glucose sensor values < 70 mg/dL

4.2 Secondary Efficacy Endpoints:

CGM Metrics

Overall Glucose Control

1) Mean of CGM glucose sensor values
2) Percent of CGM glucose sensor values between 70 and 180 mg/dL (not inclusive of the limits)

Hypoglycemia

3) Percent of CGM glucose sensor values < 60 mg/dL
4) Percent of CGM glucose sensor values < 50 mg/dL
5) Area over curve for CGM glucose sensor values < 70 mg/dL.
6) Low blood glucose index (LBGI).
7) Frequency of CGM-measured hypoglycemic events per week (see below)

Hyperglycemia

8) Percent of CGM glucose sensor values > 250 mg/dL
9) Percent of CGM glucose sensor values > 180 mg/dL
10) Area under curve for CGM glucose sensor values > 180 mg/dL.
11) High blood glucose index (HBGI)

4.3 Calculation of CGM Metrics

Each metric will be calculated for CGM sensor values over the entire 24-hour period.

Primary Outcome - % Time Below 70 mg/dL

The percentage of glucose sensor values <70 mg/dL will be calculated for each participant in each 3-week treatment period by pooling all glucose sensor readings that occur within each 3-week period. Available CGM data will be included in the calculation based on when the subject completes the study visits. If exact date-times of when patient left the clinic on study period initiation visits are available, then the study windows for CGM data inclusion are defined as:

	Includes All Available CGM data from:	
	Start Date	**Last Date**
Period 1	Date-time in which patient left clinic on Period 1 Initiation Visit	Earlier of: a) Date-time in which patient left clinic on Period 1 Initiation Visit + 21 full days, or b) Date-time in which patient left clinic on Period 2 Initiation Visit
Period 2	Date-time in which patient left clinic on Period 2 Initiation Visit	Earlier of: a) Date-time in which patient left clinic on Period 1 Initiation Visit + 21 full days, or b) Final Visit

Otherwise, in the case of a protocol deviation and the above information is not available, the study windows are defined as:

	Includes All Available CGM data from:	
	Start Date	**Last Date**
Period 1	Randomization Visit +1	Earlier of: a) Randomization Visit + 22 Days, or b) Period 2 Initiation Visit
Period 2	Period 2 Initiation Visit +1	Earlier of: a) Period 2 Initiation Visit + 22 Days, or b) Final Visit

All glucose sensor readings will be weighted equally in the pooled percentages regardless of how they distribute across weeks. Data will not be truncated due to protocol deviations.

Secondary CGM Metric Outcomes

All secondary CGM metrics will be calculated analogously as described above for the primary outcome.

PLGS and CGM Use

Number of hours of CGM/PLGS use per week will be calculated as the sum of the time intervals between readings, excluding calibration readings. If an actual time interval is greater than the device's default time interval, then that time interval will be set to the device's default time interval. Number of hours of CGM and PLGS use will be expressed as a percentage of the total number of hours in a week.

Total percentage of time PLGS and CGM is used per week will be calculated by dividing the total number of hours of PLGS readings by the maximum possible number of hours within the study period visit dates described above. Number of days of PLGS and CGM use per week is calculated as the count

of days during the week where there was at least one PLGS/CGM reading. Data from both study periods will be included in this analysis. No missing data will be imputed.

Hypoglycemic Events

A CGM-measured hypoglycemic event will be defined as:

- at least 2 sensor values < 54 mg/dl
- that are 15 or more minutes apart
- no intervening values > 54 mg/dl
- Event ends when:
  - at least 2 sensor values > 70 mg/dl
  - that are 30 or more minutes apart
  - no intervening values < 70 mg/dl
- Participant becomes eligible for a new event as soon as the above criteria for ending the previous event have been met.

5. Description of Statistical Methods

5.1 General Approach

All analyses will compare the 3-week period of PLGS to the 3-week period of SAP and will follow the intention-to-treat principle with each period analyzed according to the treatment assigned by randomization regardless of actual PLGS or SAP utilization. All participants who have at least one CGM reading in each 3-week period will be analyzed for the primary and secondary outcomes. All p-values will be two-sided.

Standard residual diagnostics will be performed for all analyses. If values are highly skewed, then a transformation or nonparametric methods will be used instead. Previous experience suggests that a transformation may be necessary for the hypoglycemic and hyperglycemic outcomes. We do not expect that a transformation will be necessary for some overall glucose control metrics (such as mean glucose).

5.2 Analysis Cohorts

- All randomized participants who have at least one CGM reading in both 3-week periods will be analyzed for the Intention-to-Treat (ITT) Analysis with each day included in the treatment group assigned by randomization.
- Safety outcomes will be reported for all enrolled participants, irrespective of whether the study was completed.
- A per-protocol analysis will be restricted to participants with ≥ 200 hours of CGM data in each 3-week period. This analysis will only involve the primary outcome, and will only be done if at least 10% of subjects are excluded based on this criterion.

6. Analysis of the Primary Efficacy Endpoint

6.1 Included Subjects

As noted in section 5.2, the primary analysis for this study will include all randomized subjects with at least one CGM sensor reading in each of the two treatment periods. No additional minimum amount of CGM use is required otherwise.

6.2 Missing Data

There will be no imputation of missing CGM data for the primary analysis in this study.

6.3 Statistical Methods

Mean ± SD (standard deviation) or summary statistics appropriate to the distribution will be reported for % time below 70 mg/dL for each treatment arm. A repeated measures regression model with an unstructured covariance structure will be fit for percent time below 70 mg/dL to compare the two treatments. The model will adjust for period as a covariate.

Residual values from the above model will be examined for an approximately normal distribution. It is likely that the residuals themselves will have a skewed distribution, but that the paired differences from the same subject may follow an approximate bell-shaped curve. However, if the paired differences have a skewed distribution, then an appropriate transformation or a nonparametric analysis based on ranks will be performed. Significance will be assessed at the 5% level.

A carryover effect is not expected, but it will be assessed with a separate model built with the inclusion of a period by treatment interaction. However, the study is not expected to have sufficient power in making definitive conclusions about the period by treatment interaction.

7. Analysis of the Secondary Endpoints

The same subjects will be included as in the primary analysis. Missing data will not be imputed for any of the secondary analyses in this study.

7.1 Statistical Methods

For all secondary outcomes, summary statistics appropriate to the distribution will be tabulated by treatment group. Analysis of all secondary endpoints will parallel the primary analysis. A ranked normal score transformation will be applied to all highly skewed secondary outcomes.

7.2 Secondary Analyses by Time of Day

Each of the secondary outcomes will be calculated separately for sensor values during daytime (6:00 AM to just before 10:00 PM) and sensor values during nighttime (10:00 PM to just before 6:00 AM).

The primary outcome, percentage of glucose sensor values < 70 mg/dL, will also be divided into daytime and nighttime periods as an additional secondary outcome.

For all secondary outcomes separated by time of day, summary statistics appropriate to the distribution will be tabulated by treatment group. The same model described above for the primary and secondary analyses will be fit with the inclusion of a treatment by time of day interaction. The p-value for the interaction term will be reported. These analyses will be conducted to determine whether a similar trend to the overall treatment effect is seen in the different times of day. An appropriate transformation or a nonparametric analysis based on ranks will be applied to all highly skewed secondary outcomes.

The study is not expected to have sufficient statistical power for definitive conclusions in the secondary analyses by time of day, and statistical power will be low to formally assess a treatment by time-of-day interaction. Interpretation of the analyses by time of day will depend on whether the overall analysis demonstrates a significant treatment effect. In the absence of any significant treatment effects in the overall analyses, assessment of secondary analyses by time of day will be considered exploratory and used to suggest hypotheses for further investigation in future studies.

8. Safety Analyses

Safety outcomes will be tabulated by participant for all adverse events from enrollment to the final study visit or final adverse event, whichever comes later.

8.1 Definitions

Reportable adverse events for this protocol include severe hypoglycemia (SH), diabetic ketoacidosis (DKA), and any other untoward medical occurrence that meets criteria for an adverse event (AE) or serious adverse event (SAE). Device deficiencies that could have led to an adverse device effect will also be reported.

Hypoglycemic events will be considered severe if the participant required assistance from another person to actively administer carbohydrate, glucagon, or engage in other resuscitative actions. This means that the subject is impaired cognitively to the point that he/she is unable to treat him- or herself, is unable to verbalize his or her needs, is incoherent, disoriented, and/or combative, or experiences seizure or coma. If plasma glucose measurements are not available during such an event, neurological recovery attributable to the restoration of plasma glucose to normal is considered sufficient evidence that the event was induced by a low plasma glucose concentration.

Definite DKA is defined by the Diabetes Control and Complications Trial (DCCT) as having all of the following:

- Symptoms such as polyuria, polydipsia, nausea, or vomiting

- Serum ketones > 1.5 mmol/L or large/moderate urine ketones
- Either arterial blood pH < 7.30 or venous pH < 7.24 or serum bicarbonate < 15
- Treatment provided in a healthcare facility

8.2 Adverse Events Summary

All episodes of SH, DKA, unanticipated adverse device effects, and any other reportable adverse events will be listed by treatment group. For each event, the following information will be reported:

- Onset date of the event
- Resolution date of the event
- Duration of the event
- Phase of the trial (run-in phase, pilot period, PLGS phase, or SAP phase)
- Description of the event
- Intensity of the event
- Seriousness of the event
- Whether the event was related to the study treatment
- Whether the event was related to the study procedure
- Whether the event was related to the study device
- Whether the event required treatment
- Outcome of the event

Any adverse events that occurred before the randomization visit (i.e., during the run-in or pilot phase) will be tabulated in separate pre-randomization tables in the same manner as the post-randomization adverse events.

8.3 Comparison of Safety Outcomes between Treatment Groups

Due to the short duration of the study, it is not anticipated that there will be enough DKA, SH, ketosis (a calendar day with at least one instance of blood ketone level > 1.0 mmol/L), unanticipated adverse device effects, or any other reported adverse events for a formal statistical comparison of the treatment arms. If there are enough events, the following analyses will be performed.

For each of the following safety outcomes, mean ± SD or summary statistics appropriate to the distribution will be tabulated by treatment group:

- Number of subjects with any DKA events
- Number of episodes of DKA events per subject and incidence rate per 100 person years
- Number of subjects with any SH events
- Number of episodes of SH events per subject and incidence rate per 100 person years
- Number of subjects with any ketosis events
- Number of episodes of ketosis events per subject and incidence rate per 100 person years

- Number of subjects with any unanticipated device effects
- Number of unanticipated device effects per subject and incidence rate per 100 person years
- Number of subjects with any other reported adverse events
- Number of other reported adverse effects per subject and incidence rate per 100 person years

All of the above safety outcomes will be tabulated for all subjects (including dropouts and withdrawals), regardless of whether CGM data are available or whether the device was operational. Any adverse events that occurred before the randomization visit (i.e., during the run-in or pilot phase) will not be included in the rate calculations or treatment group comparisons listed above.

The number of person-years for the calculation of incidence rates will be defined as the number of person-years from the randomization visit to the last visit date or adverse event date, whichever is later.

For each of DKA, SH, ketosis, unanticipated adverse device effects, and all other reported adverse events, if there are enough events, rates will be compared using a repeated measures Poisson regression model adjusting for period.

9. Adherence and Retention Analyses

9.1 Amount of PLGS and CGM Usage

The amount of PLGS and CGM use in hours per week and days per week will be tabulated for each treatment arm (PLGS use will only be for the PLGS arm).

Summary statistics appropriate to the distribution and range will be reported by treatment group for the days per week and percentage of hours per week of CGM and PLGS use. The percentage of time per week of CGM use will be compared between treatments using the same model described above for the primary outcome.

Amount of PLGS and CGM usage will further be tabulated by daytime and nighttime (as defined above), but no formal statistical comparison will be made.

9.2 Protocol Adherence

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

- Number of protocol deviations by severity with descriptions of each deviation instance
- Number of unscheduled visits and reasons for each unscheduled visit instance
- Number of device issues and descriptions of each instance

Tabulations and analyses will be divided into pre- and post-randomization stages where possible. For post-randomization stages, tabulations and analyses will be further divided into treatment groups. In addition, the following flowcharts will be produced:

- Flow chart accounting for all subjects at all visits post-randomization to assess visit completion rates
- A flow chart accounting for the number of subjects enrolled, the number of subjects who completed the run-in phase, the number of dropouts pre- and post-randomization, and the number of subjects eligible to be included in the primary analysis

10. Baseline Descriptive Statistics

Baseline demographic and clinical characteristics of the study cohort will be summarized in a table in addition to baseline CGM metrics. Baseline CGM data will be taken as follows from the home use, and SAP training phases (details above in Section 1):

- If subject was required to go through SAP Training, then baseline CGM data from the SAP Training period will be used.
- If subject was allowed to skip both CGM and SAP Training, then baseline CGM data from 28 days of home use prior to enrollment will be used.

The following descriptive statistics will be tabulated overall and by randomization group:

- Age in years
- Gender
- Race
- Type I diabetes duration in years
- BMI (if available)
- HbA1c (% mmol/L)
- Baseline CGM Metrics
- Current insulin modality
- Past amount of pump use for patients using pump at enrollment
- Daily insulin units (total, basal, and bolus)
- CGM use status at enrollment
- Days of CGM use in the past month at enrollment
- Daily insulin units
- Time range of most recent severe hypoglycemic event
- Number of severe hypoglycemic events in the last 12 months
- Number of severe hypoglycemic events involving seizure/coma over lifetime

- Number of severe hypoglycemic events involving seizure/coma in the last 12 months
- Time range of most recent DKA event
- Number of DKA events in the last 12 months
- Number of glucose tests per day
- Other non-insulin blood sugar control medications taken

For continuous variables, summary statistics appropriate to the distribution will be given. For discrete variables, number and percentage will be reported for each category.

11. Planned Interim Analyses

The purpose of the interim analysis is to re-estimate sample size in the case that the overall rate of hypoglycemia (% time below 70 mg/dL) in the control (SAP) arm is significantly less than what was assumed in the original sample size calculation. Only CGM data from baseline and the SAP arm will be used. The analyses will be masked to the PLGS data. No assessment of efficacy with the PLGS arm is to be made, and thus there will be no inflation of type 1 error rate.

With the planned sample size of 90, the original calculation estimated statistical power >99% to detect a 33% relative reduction assuming the control arm hypoglycemia is 4.0% using a two-tailed test with the type 1 error rate at 5%. The goal of this interim analysis is to assess whether the 4.0% assumption used in the original calculation was approximately correct, and if not whether the sample size needs to be modified. The interim analysis will be conducted when the 90th subject is enrolled into the screening phase. All partial data from all non-dropped subjects with either baseline or post-randomization CGM data will be included. Baseline CGM data will be taken as specified in Section 10, and all post-randomization data up to the final study period date will be included.

Study Stage Completed[a]	**Rough Projection for # Subjects at Interim Analysis**
CGM data from pre-randomization run-in[b]	N=90
Completed ≥1 weeks of Period #1	N=45 (~22 or 23 randomized to control-first)
Completed Period #1	N=22 (~11 randomized to control-first)
Completed ≥1 weeks into Period #2	N=0 (but possible if slower recruitment)
Completed Period #2	N=0 (but possible if slower recruitment)

[a] Each row denotes a subset of subjects in the row above it.
[b] Taken as specified in the table in section 10

A repeated measures model controlling for study stage will be run with 1-3 data values per subject depending on which stage of the study (as categorized in the table above) was completed by the subject at the time of the interim analysis. This model will be used to estimate the amount of hypoglycemia in the control arm utilizing information from both the baseline and the control arm of the RCT phase. Any

data from a participant's PLGS study period will be excluded. The assumption is that the amount of hypoglycemia from the baseline will be reasonably well-correlated with the amount in the control arm during the RCT phase.

Hypoglycemia, or percent of time glucose is below 70 mg/dL, from each of the study stages in the table will be calculated as specified in Section 4.3. If the overall estimated time below 70 mg/dL is <4%, then we will take the adjusted sample size to be:

*52*4% / estimated hypoglycemia rate in control arm*

where the estimated rate in the denominator is calculated from the model described above. If the estimated rate of hypoglycemia is so low that even 150 subjects would provide little power, then we will stay with the originally planned 90.

12. Subgroup Analyses

Subgroups will be defined by:

- Age (grouped by < 18 years and ≥ 18 years)
- Type 1 diabetes duration (split into 2 groups with cutpoint approximately at the median)
- Baseline HbA1c (grouped by < 8.0% and ≥ 8.0%)

Mean ± SD or summary statistics appropriate to the distribution will be reported for % time below 70 mg/dL for each treatment arm in each subgroup. These analyses are considered exploratory and are not being evaluated as independent hypothesis tests.

13. Multiple Comparisons/Multiplicity

The primary analysis involves a single treatment arm comparison for a single outcome measure so no correction for multiple comparisons will be performed.

For the secondary analyses, the false discovery rate (FDR) will be controlled using the adaptive Benjamini-Hochberg procedure for multiple comparisons. The FDR corrections will be done separately for the following categories:

- Secondary Efficacy Outcomes and CGM Usage
- Secondary Efficacy Outcomes by Time of Day
- Miscellaneous (CGM Usage and Insulin Analyses)

Safety analyses will not be included in the multiple comparison corrections.

14. Exploratory Analyses

No additional exploratory analyses are planned.

15. Additional Analyses

The following outcomes will be tabulated separately over a 24-hour period, daytime, and nighttime where applicable:

- Insulin delivery per treatment group, including total insulin units per day, basal insulin units per day, and bolus insulin units per day.
- PLGS group only: Frequency of insulin suspension events and duration of events per subject-day, including individual suspensions and cumulative suspension time
- PLGS group only: CGM glucose nadir during suspension events and peak within 2 hours after events.

Mean ± SD or summary statistics appropriate to the distribution will be reported for each of these outcomes. Insulin delivery will be compared between treatments using the same model described above for the primary outcome.

Total basal insulin is recorded as an hourly cumulative running total of the amount of basal insulin delivered for each day. Total basal insulin will be calculated by a summation of the increments between the cumulative counts for each day, not including days where the insulin pump was reset or days where a time change caused basal readings to overlap. Total bolus will be calculated by summing all non-missing bolus delivery amounts recorded by the insulin pump for each period.

App-Based Monitoring and Intervention Questionnaire

Visits in which participant took carbs or insulin to correct blood glucose in response to a commercially-available CGM monitoring app will be tabulated by treatment group. The questions are answered at every visit past the first study period's initiation visit, and include any carbs or insulin taken since the last visit or contact. Responses are a binary yes/no. Summary statistics tabulated will include number and percentage of a) participants taking carbs and b) participants taking insulin at each visit/contact. The analysis cohort for all of these additional analyses are the same as the primary analysis. Data from both study periods (defined in section 4.3) will be included in this analysis. No missing data will be imputed.

System Usability Questionnaire

This 10-item questionnaire is given at the end of every participant's PLGS study period to determine system usability of the PLGS feature:

Question
1. I think that I would like to use this system frequently
2. I found the system unnecessarily complex
3. I thought the system was easy to use
4. I think that I would need the support of a technical person to be able to use this system
5. I found the various functions in this system were well integrated
6. I thought there was too much inconsistency in this system
7. I would imagine that most people would learn to use this system very quickly
8. I found the system very cumbersome to use
9. I felt very confident using the system
10. I needed to learn a lot of things before I could get going with this system

Responses are ranked on a 5-point scale from 1 (strongly disagree) to 5 (strongly agree). For item 1, 3, 5, 7, and 9, each participant's response is scored by taking the response ranking and subtracting 1. For items 2, 4, 6, 8, and 10, each participant's response is scored by subtracting the response ranking from 5. Higher scores denote better perceived usability. Composite scores are calculated by taking the sum of the individual item scores from each participant and multiplying by 2.5. Possible composite scores range from 0 to 100, with higher scores denoting better perceived usability.

Summary statistics appropriate to the distribution will be tabulated for each individual item and for the composite score.