Insulin Only Bionic Pancreas

Bridging Trial

Adult RCT Phase

Statistical Analyses Plan

Version 0.4

September 21, 2018

Based on Protocol Version 8.0

Note: The table shells are included in a separate document

Version History

Version	Author	Approvers	Effective Date	Revision Description	Study Stage
1.0	Dan Raghinaru/ Tonya Riddlesworth	Craig Kollman Katrina Ruedy Edward Damiano Steven Russell	9/xx/2018	Original Version	The Adult RCT Phase started the enrollment on 7/xx/2018. Still enrolling

Lead Statistician and Author:

_____

Senior Statistician Approver:

_____

JAEB PI Approver:

_____

Sponsor Approver:

_____

Study PI Approver:

_____

1. Study Overview

This document outlines the statistical analyses to be performed for the IOBP Bridging Trial – Adult RCT Period - to be presented to DSMB and maybe published.

The IOBP Bridging Trial consists of four different periods in this order:

- Test Run,
- Adult RCT,
- Pediatric Transitional,
- Pediatric RCT.

The Test-Run Period will be completed and the data will be reviewed to verify safety prior to beginning the Adult RCT Period.

The Adult RCT Period will consist of a multi-center, three-period, random-order, cross-over, feasibility study in 36 adult participants ≥ 18 years old with T1D (18 participants at MGH and 18 participants at Stanford). All 18 adult participants at MGH will have a Senseonics Eversense CGM sensor implanted prior to the initiation of the Adult RCT Period by the clinical PI at MGH (Dr. Steven Russell) or by a study physician under his supervision. These subjects will use the Senseonics Eversense CGM as the input to the bionic pancreas instead of the Dexcom G5. All 18 adult participants at Stanford will use the Dexcom G5 CGM as the input to the bionic pancreas. All other aspects will remain the same across the two sites.
Insulin therapy for each participant will be administered (i) in one period using the iLet in the insulin-only configuration with the iLet pigtail adapter and the iLet ready-to-fill insulin cartridge, the Contact Detach infusion set, and the insulin analog that they use for their usual care (either Humalog or Novolog), (ii) in another period using the iLet in the insulin-only configuration using the iLet pigtail adapter, the Contact Detach infusion set, and faster insulin aspart (Fiasp) in PumpCart, where the pharmacokinetic (PK) parameter for tmax used by the insulin-dosing algorithm will be set to the same value as is used for Humalog and Novolog (65 minutes), and (iii) in a third period using the participant's own usual care (UC), where each participant will wear a CGM. All three experimental periods will be followed by round-the-clock, remote, telemetric monitoring for hyperglycemia (> 300 mg/dl for ≥ 90 minutes) and hypoglycemia (< 50 mg/dl for ≥ 15 minutes). The three experimental periods will each have a duration of 7 days. Washout periods of approximately 7 days in duration will follow the first and second experimental periods. The washout period could be as short as 5 days or as long as 9 days, depending on scheduling of study visits. Of the 36 participants, 9 at MGH and 9 at Stanford will use MDI therapy for their usual diabetes management; the remaining 18 will use insulin pump therapy. The participants at Stanford will wear the Dexcom G5 CGM and the participants at MGH will wear the Senseonics Eversense CGM. This will allow us to collect further comparative accuracy data on both the Senseonics Eversense and the Dexcom G5 in the outpatient setting using the Ascensia Contour Next One meter as the reference. Four participants at MGH will have a Senseonics Eversense CGM sensor implanted and will participate in a 3-day supervised Test Run. This Senseonics Eversense Test Run Period will be completed and data will be reviewed to verify safety prior to beginning the Adult RCT Period at MGH using the Senseonics Eversense CGM sensor as the input to the iLet. The cohort of 18 participants at MGH and 18 participants at Stanford might run simultaneously, might overlap, or might not overlap in time.

Major eligibility criteria include:

- Clinical diagnosis of type 1 diabetes for at least 12 months
- HbA1c level < 11.0%
    - At least 13 participants must have an HbA1c level of 8–11%
    - At least 13 participants must have an HbA1c level < 8%
- Using an FDA-approved insulin therapy (including Fiasp on MDI therapy for adults, Humalog or Novolog with CSII or MDI therapy; glulisine users will be excluded)
    - 18 participants must use MDI therapy to manage their diabetes
    - 18 participants must use CSII therapy to manage their diabetes
- Age ≥ 18 years old
    - Approximately 9 participants in a young adult group (18–24 years old)
    - Approximately 27 participants in an adult group (≥ 25 years old)
- At least 3 daily SMBG or 2 if used with a CGM

The following table gives an overview of the schedule of study visits and key procedures during each one of the three 7-day randomized treatment periods [BP with analog insulin (1), BP with rapid insulin (2), and UC (3)]:

			Each Treatment Period			
	Screening	**Senseonics Eversense Sensor Insertion**	**0**	**Day 3 or Day 4**	**Day 7**	**Senseonics Eversense Sensor Removal**
Informed Consent	**X**					
Eligibility assessment	**X**		**X**			
Bloodwork: eGFR and HbA1c	**X**					
EKG	**X**					
Body weight	**X**		**X**		**X**	
Urine pregnancy test	**X**					
Physical exam	**X**					
Insertion of Dexcom CGM sensor			**X**			
Insertion of Senseonics Eversense CGM sensor		**X**				
Removal of Senseonics Eversense CGM sensor						**X**
Psychosocial questionnaires			**X**		**X**	
Data download					**X**	
Adverse event querying			**X**	**X**	**X**	

2. Statistical Hypotheses

The primary outcome for this study is CGM-measured mean glucose when compared during (2) BP with rapid insulin and UC (3) treatment periods.
The hypotheses are:

1. *Null Hypothesis*: There is no difference in mean CGM-measured glucose during UC and BP with rapid insulin treatment periods
2. *Alternative Hypothesis*: The mean CGM-measured glucose is different during UC and BP with rapid insulin treatment periods.

The intervention will be considered effective if the above null hypothesis is rejected using a significance level of 0.05 (i.e. $p<0.05$).

3. Sample Size

Thirty-six subjects will be enrolled and randomized at two sites. This is a convenience sample; however assuming an effective SD of 25 mg/dl for mean glucose, the half width of the associated 95% confidence interval are approximately ±10 mg/dl for N=36 and ±15 mg/dl for N=18. Results from the control period of the JDRF CGM RCT were used to estimate SD for mean glucose in the proposed study.

4. Outcome Measures

4.1 Primary Efficacy Outcome

- CGM-measured mean glucose

4.2 Secondary Efficacy Outcome

4.2.1 Secondary Efficacy Endpoint Included in Hierarchical Analysis

- CGM-measured time <54 mg/dl

4.2.2 Other Secondary Efficacy Outcomes

The following endpoints are considered exploratory.

- CGM metrics related to hypoglycemia
  - % <70 mg/dL
  - % <60 mg/dL
  - low blood glucose index
- CGM metrics related to overall control
  - % time in target range 70-180 mg/dL
  - % time in target range 70-120 mg/dL
  - glucose variability measured with CV (coefficient of variation)
  - glucose variability measured with MODD (mean of daily difference)
- CGM metrics related to hyperglycemia
  - % >180 mg/dL
  - % >250 mg/dL
  - high blood glucose index

- Total daily dose of insulin

- Questionnaires - The following table provides a description of the questionnaires to be administered during each one of the three 7-day randomized treatment periods [BP with analog insulin (1), BP with rapid insulin (2), and UC (3)]:

Questionnaire's Name	Day 0	Day 7	Description
Diabetes Treatment Satisfaction Questionnaire - Status	**X**		Measures patient satisfaction with diabetes treatment. 8 items (6 for assessing treatment satisfaction and two for frequency of hypoglycemia and hyperglycemia). Scale 0-6. Higher scores denote higher satisfaction or higher frequency of hypoglycemia/hyperglycemia.
Diabetes Treatment Satisfaction Questionnaire - Change		**X**	Similar with the above (Status); except scale -3 to 3.
Diabetes Distress Scale	**X**	**X**	Measures Adult and Partner of Adult diabetes distress. Adult: 17 items (4 subscales and total score) and scale 1-6. Partner: 21 items (4 subscales and total score) and scale 0-4. Higher score denotes more of a problem or more distress.
Hypoglycemia Fear Survey	**X**	**X**	23 items (2 subscales and total score). Scale 0-4. Higher score denotes more fear.
Hypoglycemia Confidence Scale - Status	**X**	**X**	Measures confidence status. 8 items and a 4-level scale.
Hypoglycemia Confidence Scale - Change		**X**	Measures confidence change. 8 items and a 5-level scale.
INSPIRE Survey	**X**	**X**	Measures experience with automated insulin delivery systems. Adult and partner versions. 31 items. 5-point Likert scale.
Bionic Pancreas User Opinion Survey*		**X**	Measures user's experience with BP. 43 items. Two subscales (benefits and difficulties) and total score. 5-point Likert scale.
Daily At-Home Questionnaires			Administered at the end of each day – expected 6 times for each subject and randomized period combination. 17 items – most of them with sub-items (depending on the subject's answer).

*to be administered only during the two BP periods

4.3 Calculation of CGM Metrics:

- All sensor data excluding the first 48 hours will be included in the calculation of glycemic metrics.
- Only subjects with at least 24hr of CGM data in any two of the three treatments periods will be included in CGM analyses. It is considered that 24hr of data is the minimum and a representative amount of data for any treatment period.
- If there are 12-<24hr of CGM data from days 3-7 in each one of the three treatment periods and there are more CGM data during days 1-2, then the data from days 3-7 will be supplemented with

data from days 1-2 (going backwards) until 24hr of data is reached. Otherwise the data are considered missing for that treatment period.

- Metrics will be calculated giving equal weight to each CGM point from each treatment period.

5. Description of Statistical Methods

5.1. General Approach

Analyses will involve pairwise comparisons of the three treatment arms:

1. Bionic Pancreas (BP) with analog insulin
2. Bionic Pancreas (BP) with rapid insulin
3. Usual Care (UC)

All primary, hierarchical, and secondary analyses comparing the three treatment periods will follow the intention-to-treat (ITT) principle with each period analyzed according to the treatment assigned by randomization. All randomized participants with at least 24hr of CGM readings during at least two of the three treatment periods will be included in the primary, hierarchical, and secondary analyses of CGM metrics. There will be no imputation of missing data.

All p-values will be two-sided.

Standard residual diagnostics will be performed for all analyses. If values are highly skewed, then a rank transformation will be used instead for the primary and secondary outcomes. From prior experience, the values for time <54 mg/dl will have a skewed distribution and for mean glucose a bell-shaped distribution.

5.2. Analyses Cohorts

Primary and Secondary Analyses:

- For CGM-based and insulin outcomes, all randomized participants and treatment periods will be analyzed according to the ITT principle and the 24hr rule as described above.
- For questionnaires, all questionnaires with at least one item answered will be included in analyses. Total and subscale score may be calculated or not depending on specific instructions or rules (see section 7.2).

Safety Analyses:

- Safety outcomes will be reported for all enrolled participants, irrespective of whether the safety event happened during the randomized treatment periods or in between (i.e. during wash-out periods).

Sensitivity Analyses:

- Primary and secondary CGM-based outcomes will be replicated using:
  - all CGM data during the randomized treatment periods (i.e. days 1-7), and
  - only the first 24hr of CGM data from each treatment period following randomization.
- Additional models for primary and secondary outcomes will be run to test for any carry-over effects, by adding a treatment by order interaction term.

6. Primary Analysis

The primary analysis will be a comparison of CGM-measured mean glucose between (2) BP with rapid insulin versus (3) UC treatment periods.

Plots and summary statistics appropriate to the distribution for mean glucose will be reported.

A repeated measures linear model with mean glucose as the dependent variable will be fit to the data including all three treatment periods. The model will adjust for treatment period and account for correlated data from the same subject using an unstructured covariance matrix.

The p-value for the primary analysis will be calculated using a contrast in the model described above between the (2) BP with rapid insulin (3) UC treatment arms. From prior experience, the values for mean glucose will follow a bell-shaped distribution. Residual values will be examined for an approximate normal distribution. If values are highly skewed then rank transformation will be used instead.

7. Secondary Analyses

Plots and summary statistics appropriate to the distribution will be given for both mean glucose and time below 54 mg/dl for all three treatment arms.

For mean glucose, contrasts from the same model described above will also be used to compare:

- (2) BP with analog insulin versus (3) UC, and
- (2) BP with rapid versus (1) BP with analog insulin

Both of these are considered secondary analyses.

A similar model will be fit with time below 54 mg/dl as the dependent variable with contrasts for all three pairwise comparisons of the treatment arms. It is anticipated that the distribution of this metric will be skewed. In this case, a transformation will be used as described above.

7.1 Hierarchical Analyses

To preserve the overall type 1 error for the 6 comparisons (2 selected key secondary endpoints times 3 treatment arm pairs), a hierarchical testing procedure will be used. If the primary analysis described above results in a statistically significant outcome ($p < 0.05$), then testing will proceed to the next comparison in the following order:

- Mean glucose for (2) BP with rapid insulin vs. (3) UC (primary analysis)
- Time <54 mg/dl for (2) BP with rapid insulin vs. (3) UC
- Mean glucose for (1) BP with analog insulin vs. (3) UC
- Time <54 mg/dl for (1) BP with analog insulin vs. (3) UC
- Mean glucose for (2) BP with rapid vs. (1) BP with analog insulin
- Time <54 mg/dl for (2) BP with rapid vs. (1) BP with analog insulin

This process continues iteratively moving to the next variable down on the list until a non-significant result ($p \geq 0.05$) is observed, or all six comparisons have been tested. If a non-significant result is

encountered, then formal statistical hypothesis testing is terminated and any comparisons below on the list are not formally tested.

For example, in the hypothetical scenario depicted in the table below, the first three comparisons both have a significant result so testing continues to the fourth row [Time <54 mg/dl for (1) BP with analog insulin vs. (3) UC]. The result is not significant for that fourth row (p = 0.06) so testing stops. No formal hypothesis test is conducted for the fifth and sixth comparisons on the list in this example scenario.

Table 1. Example application of the hierarchical procedure to control type 1 error.

HIERARCHICAL ORDER	COMPARISON	CONTRAST P-VALUE	SIGNIFICANT?	ACTION
1st	Mean glucose for (2) BP with rapid insulin vs. (3) UC (primary outcome)	0.001	Yes	Test next comparison
2nd	Time <54 mg/dl for (2) BP with rapid insulin vs. (3) UC	0.02	Yes	Test next comparison
3rd	Mean glucose for (1) BP with analog insulin vs. (3) UC	0.007	Yes	Test next comparison
4th	Time <54 mg/dl for (1) BP with analog insulin vs. (3) UC	0.06	No	Stop formal testing
5th	Mean glucose for (2) BP with rapid vs. (1) BP with analog insulin	Not tested	Unknown	N/A
6th	Time <54 mg/dl for (2) BP with rapid vs. (1) BP with analog insulin	Not tested	Unknown	N/A

Regardless of the results of the hierarchical testing, summary statistics appropriate to the distribution will be tabulated.

7.2 Other Endpoint Analyses

CGM-Measured Outcomes

The analyses for the secondary CGM-measured outcomes will parallel the descriptive statistics and regression model mentioned above for the primary outcome without the hierarchical procedure. Analyses will also be done limiting to daytime (6am-12mn) and nighttime (12mn-6am) data. Only subjects with at least 12hr of CGM data in any two of the three treatments periods will be included in daytime/nighttime CGM analyses.

Questionnaires

For each question, the percent of all answers will be reported across all subjects by treatment periods; and if applicable, by beginning and ending of each treatment period. For questions with additional numeric or ordinal scales, mean score will be reported in addition to the percent.

Some questionnaires will be administered at the beginning and ending of each treatment period, and the results presented as such. If a subject completed the beginning but not the ending questionnaires, the results will be presented as such (i.e. with different underlying number of subjects).

For some questionnaires total and subscale scores will be calculated and mean ± SD values or percentiles appropriate to the distribution will be reported across all subjects by treatment periods.

All questionnaires will be administered online (through REDCap) and subjects can skip specific questionnaires or questions within a questionnaire. 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 questions, then the following "75% rule" will be used. For all questionnaires, 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 in order to a corresponding total or subscale score.

The Daily At-Home Questionnaire will be administered at the end of each day (for a maximum of 6 times for each one of the three treatment periods). Each question in this questionnaire will be first summarized (percent or mean) and presented at period level and then summarized across all subjects by treatment period as mentioned above.

No formal hypothesis tests will be done for questionnaires.

Sensor Accuracy

Each reference glucose measurement from the blood glucose meter will be paired with the closest sensor reading within ± 5 minutes (choosing the CGM that immediately precedes the BGM value, rather than the other way around). Each reference measurement will be paired with only reading from sensor and each sensor reading will only be paired with one meter reference value. Any meter measurements used to calibrate the sensor will be excluded from the accuracy analysis. Any data available during the wash-out periods will also be included.

Difference, absolute relative difference, and International Organization for Standardization criteria (ISO) will be calculated for each sensor-reference pair. Descriptive summary statistics will be reported separately for Eversence and DexCom according to their distribution overall and by different meter ranges:

- <70 mg/dl
- 70-180 mg/dl
- >180 mg/dl
- >250 mg/dl

Insulin Outcomes

Total daily insulin analyses during the BP treatment period will include data from days 3-7 (i.e. exclude the first 48 hours of data) and will be displayed using boxplots for each day. Summary statistics for the total daily insulin reported on the CRF on Days 0 and 7 will also be given for both treatment periods.

8. Safety Analyses

All subjects will be included in these analyses and all their safety events will be reported.
The circumstances of all reportable cases of the following will be summarized and tabulated by randomized treatment periods and wash-out intervals:

- Severe hypoglycemia (as defined in Section 8.1 of the protocol)
- Diabetic ketoacidosis (as defined in Section 8.1 of the protocol)
- All other adverse events
- Ketone events defined as ketone level >1.0 mmol/l
- CGM-measured hypo- and hyperglycemia events

Due to the relatively short follow-up duration for this study, it is unlikely there will be enough such events for a formal statistical comparison. However, if there are enough events, a Poisson regression that accounts for correlated data from the same subject will be used to compare the number of events between the three treatment periods. The amount of follow up will be included as an offset covariate to compare the rates. Influence diagnostics will be reviewed and if outliers are present then the model will be run without them and the results compared.

The CGM-measured (days 1-7) hypo- and hyperglycemia events will use all available CGM data and are defined as follows:

- hypoglycemia - defined as at least 15 consecutive minutes <50 mg/dl
- hyperglycemia - defined as at least 90 consecutive minutes >300 mg/dl

9. Intervention Adherence

A listing of the following outcomes as well as summary statistics appropriate to the distribution will be given to assess intervention adherence for the study:

- The daily frequency of downloaded SMBG use
- BP use (using days 3-7 and days 1-7), # hours BP was “on”, # hours BP was in Closed Loop mode and % of time BP was in Closed Loop mode.

10. Adherence and Retention Analyses

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

- Flow chart accounting for all subjects at all scheduled visits
- A listing of protocol deviations
- A listing of unscheduled visits
- A listing of subjects who dropped out of the study and reasons

11. Baseline Descriptive Statistics

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

- Age
- HbA1c
- Gender
- Race/ethnicity
- Income, education, and insurance status
- Clinical Center
- Diabetes duration
- BMI
- Total daily insulin
- Previous Pump and CGM use

12. Device Issues

Device issues reported on the CRF by the sites will be tabulated according to treatment period.

13. Planned Interim Analyses

No formal interim analyses are planned for this study.

The DSMB will review data collected.

14. Subgroup Analyses

This study was not powered for subgroup analyses. These are considered exploratory and results will be interpreted with caution. Subgroup analyses will only be performed for those pairwise treatment comparisons (see above) that were statistically significant at $p < 0.05$.

These analyses will include, but will not be limited to, evaluating potential differences between sites, including differences between the RCT Period at the MGH site using the Senseonics Eversense as the input to the iLet and the RCT Period at the Stanford site using the Dexcom G5 CGM as the input to the iLet. The general approach for these analyses will be to add an interaction term for the subgroup factor by treatment into the models used for the analyses described above. Continuous variables will be modelled as such in the subgroup analyses to calculate p-values for interaction, but summary statistics will be displayed as discrete subgroups in tables.

Summary statistics appropriate to the distribution will be tabulated in each subgroup by treatment period.

The following baseline factors will be assessed:

- Age

- Baseline HbA1c
- Study Center/Sensor
- Insulin method: pump versus injection
- T1D duration

15. Multiple Comparison/Multiplicity

For the 6 key comparisons listed in Section 7.1, strong control of the familywise error rate (FWER) will be accomplished using the hierarchical procedure described above.

For all other secondary analyses which are considered exploratory, the false discovery rate will be controlled using the adaptive Benjamini-Hochberg procedure. The method will be applied separately for the following four categories:

- Secondary CGM-measured and insulin for BP-rapid vs. UC comparisons,
- Secondary CGM-measured and insulin for BP-analog vs. UC comparisons,
- Secondary CGM-measured and insulin for BP-rapid vs. BP-analog comparisons, Subgroup analyses mentioned in section 14 (if applicable)

No multiple comparisons adjustment will be made for sensitivity analyses or safety analyses.

16. Additional Analyses

Twenty-four hour boxplots by the three treatment periods will be generated for the following CGM-based outcomes:

- % <54 mg/dl
- Mean glucose
- % >180 mg/dl
- % 70-180 mg/dl
- Coefficient of variation

The purpose of the above boxplots will be to see the trend of each outcome within a 24-hour day for each treatment period and across all subjects.