Statistical Analysis Plan

for

Prepivotal Evaluation of the Safety and Effectiveness of the Omnipod Horizon™ Automated Glucose Control System in Patients with Type 1 Diabetes

SAP: version 1, dated 23 October 2019

Prepivotal Protocol: version 2.0, dated 23 October 2019

Table of Contents

1 Purpose 4
2 Study Design 4
2.1 Study Design Overview 4
2.2 Investigational Device 4
2.3 Study Schedule 4
2.4 Target Blood Glucose Challenges 5
2.5 Study Success 5
2.6 Randomization 5
2.7 Point of Enrollment 5
3 Study Objectives and Endpoints 5
3.1 Safety Objectives and Endpoints 5
3.2 Primary Effectiveness Objective and Endpoint 6
3.3 Secondary Effectiveness Objective and Endpoints 6
4 Analysis of Primary Effectiveness Endpoints 6
4.1 Sample Size 7
4.2 Primary Analysis 7
4.3 Subgroup Analyses 7
5 Analysis of Safety 7
5.1 Medical Monitor 8
5.2 Relationship to Study Device and Procedure 8
5.3 Severity 8
5.4 Analysis of Adverse Events 9
5.5 Evaluation of Device Deficiencies 9
6 Analysis of Secondary Effectiveness Endpoints 9
6.1 Analysis of Percentage of Time in Range Endpoints 10
6.2 Analysis of Percentage of Time in Hybrid Closed-Loop 10
6.3 Analysis of Glucose Measures 10
6.4 Analysis of Insulin Requirements 11
6.5 Analysis of Secondary Effectiveness Endpoints 11
7 General Statistical Considerations 11

7.1 Analysis Sets 11
7.1.1 ITT (Intention to Treat) Analysis Set 12
7.1.2 mITT (modified Intention to Treat) Analysis Set 12
7.1.3 Per Protocol 12
7.2 Control of Systematic Bias 12
7.3 Pooling Data across Centers 13
7.4 Calculation of Percentage of Time in Range 13
7.5 Handling of Missing Data 13
7.6 Other Data Summaries 13
7.7 Statistical Software 14
7.8 Periodic Review of Controller Data 14

1 PURPOSE

The purposes of this Statistical Analysis Plan (SAP) is to describe the statistical methodology that will be used to confirm the safety and effectiveness of the Omnipod Horizon™ Automated Glucose Control System in patients with type 1 diabetes. Should there be any discrepancy between the study protocol and this SAP, the content of the SAP shall prevail.

2 STUDY DESIGN

2.1 STUDY DESIGN OVERVIEW

This is a single-arm, multi-center, prospective clinical study. The study will be conducted in 6-8 clinical study sites in the United States. The study is expected to enroll up to 48 subjects in order to obtain 36 evaluable subjects with 18 subjects in each of two age cohorts (6-13.9 years and 14-70 years of age).

2.2 INVESTIGATIONAL DEVICE

The Omnipod Horizon™ Automated Glucose Control System ("Horizon™ System") is comprised of three primary components:

- Omnipod Horizon™ tubeless, insulin delivery alternate controller enabled (ACE) pump (Pod)
- Omnipod Horizon™ Personal Diabetes Manager (PDM) which is a Samsung J3 locked down android device that operates the Omnipod Horizon™ App
- Dexcom G6 - Continuous Glucose Monitoring (CGM) system

In addition, the following non-investigational, commercially available devices will be used during the study:

- Contour® Next One blood glucose meter (Ascensia Diabetes Care, 5 Wood Hollow Road, Parsippany, NJ 07054 USA)
- Precision Xtra ketone meter (Abbott Diabetes Care Inc., 1360 South Loop Road, Alameda, CA 94502 USA)

2.3 STUDY SCHEDULE

The study will consist of 14-day standard therapy (ST) phase and a 14-day hybrid closed-loop (HCL) phase. Following the ST phase, subjects will be divided into 2 groups. Group 1 subjects (N=16; 8 subjects in each age cohort) will participate in the HCL phase for the first 2 days while in the supervised hotel/rental house environment and then transition to the 12-day outpatient setting. The Group 2 subjects (n=20; 10 subjects in each age cohort) will participate in the HCL phase in a 14-day outpatient setting.

The Horizon™ System has two modes. In manual mode, the system will function equivalently to

the Omnipod DASH System. This includes delivering insulin at programmed basal rates and bolus amounts with the option to set temporary basal profiles. In automated mode, the system will support the use of multiple target glucose values. The programmed basal rates, glucose targets and bolus calculator settings will inform the MPC algorithm for insulin dosing parameters.

2.4 TARGET BLOOD GLUCOSE CHALLENGES

All subjects take part in the prescribed target blood glucose (BG) challenges for approximately 72 consecutive hours each:

- Target BG of 130 mg/dL during HCL days 1-3
- Target BG of 140 mg/dL during HCL days 4-6
- Target BG of 150 mg/dL during HCL days 7-9
- Target BG of subject's choosing of 110-150 mg/DL during HCL days 10-14

2.5 STUDY SUCCESS

There are no pre-specified criteria that are evaluated to consider the study a success. An independent Data Safety Monitoring Board (DSMB) will evaluate the study progress with respect to safety and determine whether the enrollment of pivotal subjects may proceed.

2.6 RANDOMIZATION

This is a single-arm, multicenter, prospective clinical study where all eligible subjects will use the investigation device. Subjects will not be randomized.

2.7 POINT OF ENROLLMENT

A subject is enrolled in the study upon placement of the first study CGM. Subjects that do not meet the eligibility criteria will not continue in the study and will be considered screen failures.

3 STUDY OBJECTIVES AND ENDPOINTS

3.1 SAFETY OBJECTIVES AND ENDPOINTS

The primary safety objective is to evaluate the safety of the Omnipod Horizon™ Automated Glucose Control System in patients with type 1 diabetes. The safety objective will be evaluated by summarizing the following endpoints during HCL:

- Primary safety endpoint:
  - Proportion of subjects with serious device-related adverse events
- Secondary safety endpoints:
  - Proportion of subjects with severe hypoglycemia
  - Proportion of subjects with diabetic ketoacidosis (DKA)

3.2 PRIMARY EFFECTIVENESS OBJECTIVE AND ENDPOINT

The primary effectiveness objective is to evaluate the effectiveness of the Omnipod Horizon™ Automated Glucose Control System. The primary effectiveness objective will be evaluated by summarizing the following:

- Percentage of time in range 70-180 mg/dL during hybrid closed-loop (HCL) compared to standard therapy (ST) for
    - Target BG challenge days (approximately HCL days 1-9)
    - Non-challenge days (approximately HCL days 10-14)
    - Overall (HCL days 1-14)

3.3 SECONDARY EFFECTIVENESS OBJECTIVE AND ENDPOINTS

The secondary effectiveness objective is to evaluate additional glycemic measures of effectiveness of the Omnipod Horizon™ Automated Glucose Control System. The secondary objective will be evaluated based on the following per subject endpoints:

- Glucose metrics from system CGM during the hybrid closed-loop phase will be compared to the standard therapy phase during the day, overnight, and overall:
    - Mean glucose
    - Percentage of time in range 70-180 mg/dL
    - Percentage of time in range 70-140 mg/dL
    - Percentage of time in range > 180 mg/dL
    - Percentage of time in range ≥ 250 mg/dL
    - Percentage of time in range ≥ 300 mg/dL
    - Percentage of time in range < 70 mg/dL
    - Percentage of time in range < 54 mg/dL
    - Standard deviation
    - Coefficient of variation
- Glucose management indicator (GMI) based on overall mean glucose
- Percentage of time in hybrid closed-loop as proportion of overall device usage time
- Insulin requirements during the hybrid closed-loop phase will be compared to the standard therapy phase:
    - Total daily insulin (TDI) (units, units/kg)
    - Total daily basal insulin (units, units/kg)
    - Total daily bolus insulin (units, units/kg)

4 ANALYSIS OF PRIMARY EFFECTIVENESS ENDPOINTS

The primary effectiveness objective will be evaluated by summarizing the following:

- Percentage of time in range (TIR) 70-180 mg/dL during hybrid closed-loop (HCL) compared to standard therapy (ST) for
    - Target BG challenge days (approximately HCL days 1-9)
    - Non-challenge days (approximately HCL days 10-14)
    - Overall (HCL days 1-14)

4.1 SAMPLE SIZE

This is a single-arm, multi-center, prospective study. The sample size is not hypothesis-driven and was chosen to provide adequate information on the investigational device's safety and performance. As such, the study plans to enroll up to 48 subjects to obtain a minimum of 36 evaluable subjects equally distributed across two age cohorts (6-13.9 years and 14-70 years).

4.2 PRIMARY ANALYSIS

There are no pre-specified hypotheses associated with the primary effectiveness endpoints. Summary statistics will be presented by age cohort for all endpoints, stratified by time points of interest (e.g., day, night, overall). The primary effectiveness endpoints will be summarized for modified Intention to Treat (mITT) and Per Protocol (PP) analysis sets, where the percentage of TIR during ST phase will be compared to the percentage of TIR during HCL BG challenge days, HCL non-challenge days, and over the entire HCL.

Since the target BG challenges require a specific setpoint, the start of the challenge will be the first record at the given setpoint, and the end of the challenge will be last record at the same setpoint.

All statistical comparisons will be conducted at a two-sided significance level of 5% using a paired t-test. If the assumptions for parametric test are grossly violated, a non-parametric method such as Wilcoxon signed rank test may be used. Since the results of endpoint analyses will not be used to support clinical claims, no adjustment for multiplicity will be performed.

4.3 SUBGROUP ANALYSES

As described in this document, results will be stratified by phase of the study (ST versus HCL), age group (subjects aged 6-13.9 years versus subjects aged 14-70 years) and day versus night (daytime: 6AM to <12AM; nighttime as 12AM to <6AM). There are no plans to perform analyses separately for Group 1 and Group 2. Results for other subgroups may be presented as appropriate.

5 ANALYSIS OF SAFETY

All adverse events reported during the study will be reviewed and adequately reported to comply with applicable regulations. All reportable AEs will be assessed by the investigator who will determine whether the event is related to the study procedures or related to the study device, and

whether the event meets any of the criteria for seriousness. The event will be considered serious if the event meets at least one criterion for seriousness.

5.1 MEDICAL MONITOR

An independent Medical Monitor will be responsible for individual and timely review of adverse events. The Medical Monitor will also adjudicate all serious adverse events (SAEs), including events of severe hypoglycemia and diabetic ketoacidosis, as well as all device-related adverse events for seriousness, severity, relationship to study device and procedure, whether the event is anticipated or unanticipated, and event categorization. The adjudication decision by Medical Monitor will be used for the final classification of adverse events for regulatory reports, product labeling, and publications or presentation.

5.2 RELATIONSHIP TO STUDY DEVICE AND PROCEDURE

The causal relationship to the study procedures and the investigational device for each adverse event will be rated as follows:

- Unrelated: The event is not related to the procedures or the investigational device.
- Possibly Related: The temporal sequence is such that the relationship is not unlikely or there is no contradicting evidence that can reasonably explain the subject's condition. There is a possibility of any relation between the event and the procedures or the investigational device.
- Related: The temporal sequence is relevant or the event abates upon completion of the procedure/ investigational device, or the event cannot be reasonably explained by the subject's condition or comorbidities. The event is related or most likely associated with the procedures or the investigational device.

For the purposes of dichotomizing the causal relationship in safety summaries, events that are "Related" or "Possibly Related" will be considered related.

5.3 SEVERITY

The severity of the adverse event will be rated based upon the following grades:

- Mild: Usually transient, requires no special treatment, and does not interfere with the participant's daily activities
- Moderate: Usually causes a low level of inconvenience, discomfort or concern to the participant and may interfere with daily activities, but is usually ameliorated by simple therapeutic measures and participant is able to continue in study

- Severe: Interrupts a participant's usual daily activities, causes severe discomfort, may cause discontinuation of study device, and generally requires systemic drug therapy or other treatment

5.4 ANALYSIS OF ADVERSE EVENTS

Safety summaries and analyses will be based on all subjects that are enrolled in the study (i.e., Intention to Treat analysis set). All adverse events reported over the course of the study will be summarized and tabulated by study phase (standard therapy or HCL), event category, seriousness, severity, and relationship to the study procedures and the investigational device. Except where indicated, a subject reporting the same adverse event more than once will be counted once when calculating the number and percentage of subjects with that particular event. Adverse Event outcomes will be assessed descriptively as opposed to inferentially.

Adverse events leading to death or to discontinuation from the study will be listed separately. A listing of all adverse events will be provided. No formal tests of hypotheses are proposed for the safety endpoints.

5.5 EVALUATION OF DEVICE DEFICIENCIES

Device deficiencies will be tabulated and listed in a manner similar to the methods described for adverse events. Any device deficiency leading to an adverse event or study termination will be listed separately.

6 ANALYSIS OF SECONDARY EFFECTIVENESS ENDPOINTS

The secondary effectiveness objective is to evaluate additional glycemic measures of effectiveness of the Omnipod Horizon™ Automated Glucose Control System. The secondary effectiveness objective will be evaluated based on the following per subject endpoints:

- Glucose metrics from system CGM during the hybrid closed-loop phase will be compared to the standard therapy phase during the day, overnight, and overall:
  - Mean glucose
  - Percentage of time in range 70-180 mg/dL
  - Percentage of time in range 70-140 mg/dL
  - Percentage of time in range > 180 mg/dL
  - Percentage of time in range ≥ 250 mg/dL
  - Percentage of time in range ≥ 300 mg/dL
  - Percentage of time in range < 70 mg/dL
  - Percentage of time in range < 54 mg/dL
  - Standard deviation
  - Coefficient of variation

- Glucose management indicator (GMI) based on overall mean glucose
- Percentage of time in hybrid closed-loop as proportion of overall device usage time
- Insulin requirements during the hybrid closed-loop phase will be compared to the standard therapy phase:
    - Total daily insulin (TDI) (units, units/kg)
    - Total daily basal insulin (units, units/kg)
    - Total daily bolus insulin (units, units/kg)

6.1 ANALYSIS OF PERCENTAGE OF TIME IN RANGE ENDPOINTS

The analysis of secondary endpoints that summarize the percentage of time in range (TIR) will follow the analysis set forth for the primary effectiveness endpoint, with the range value updated as appropriate.

6.2 ANALYSIS OF PERCENTAGE OF TIME IN HYBRID CLOSED-LOOP

The percentage of time in hybrid closed-loop will be calculated as:

$$100 \times \frac{sum\ of\ gaps\ between\ consecutive\ CGM\ records}{difference\ between\ the\ earliest\ and\ the\ latest\ CGM\ records} = \%\ time\ in\ HCL$$

A gap is calculated as the difference in time (either minutes or seconds) between two consecutive CGM records as captured by the Horizon™ System. The gaps are then summed to obtain total time in HCL. As records are generally taken in about every 5 minutes and to ensure that only confirmed HCL records are included, gaps greater than 7 minutes will be excluded from analysis. For the denominator, the chronologically earliest and the latest CGM records during HCL as captured by the Horizon™ System will be used.

As opposed to the time in range endpoints, records with no glucose value in the Horizon™ System device output (such as due to an error or device deficiency during which the device does not record glucose readings) will be included in analysis. CGM records reported prior to start of HCL or after subject's discontinuation from HCL will be excluded from analysis. This endpoint will also be reported separately for the manual and automated modes.

6.3 ANALYSIS OF GLUCOSE MEASURES

Mean glucose, glucose measurement indicator (GMI), standard deviation (SD) of glucose and coefficient of variation (CV) of glucose will be evaluated per subject based on CGM output. The GMI and CV are calculated as follows:

$$GMI\ (\%) = 3.31 + 0.02392 \times (mean\ glucose\ in\ mg/dL)$$

$$CV(\%) = \frac{SD\ of\ glucose}{mean\ glucose}$$

6.4 ANALYSIS OF INSULIN REQUIREMENTS

Insulin requirements (TDI, total daily basal insulin and total daily bolus insulin) during the hybrid closed-loop phase will be compared to the standard therapy phase. The insulin requirements will be collected at Visit 2 (start of ST phase) and compared to the data collected by the Horizon™ System during HCL:

- Total daily insulin: sum of basal and bolus insulin reported over the course of HCL, adjusted for 24-hour period
- Total daily basal insulin: sum of basal insulin reported over the course of HCL, adjusted for 24-hour period
- Total daily bolus insulin: sum of bolus insulin reported over the course of HCL, adjusted for 24-hour period. This will include only bolus insulin administered and recorded by the Horizon™ System.

Data will be summarized both in insulin units and units/kg to provide a weight-adjusted comparison.

6.5 ANALYSIS OF SECONDARY EFFECTIVENESS ENDPOINTS

There are no pre-specified hypotheses associated with the secondary effectiveness endpoints. Summary statistics will be presented by age cohort for all endpoints, stratified by time points of interest (e.g., day, night, overall). All statistical comparisons will be conducted at a two-sided significance level of 5%. If the assumptions for parametric tests are grossly violated, a non-parametric method such as Wilcoxon signed rank test may be used. Since the results of endpoint analyses will not be used to support clinical claims, no adjustment for multiplicity will be performed.

As described in this document, results will be stratified by phase of the study (ST versus HCL), age group (subjects aged 6-13.9 years versus subjects aged 14-70 years) and day versus night (daytime: 6AM to <12AM; nighttime as 12AM to <6AM). Results for other subgroups may be presented as appropriate.

7 GENERAL STATISTICAL CONSIDERATIONS

7.1 ANALYSIS SETS

The following analysis sets are planned for the study.

7.1.1 ITT (Intention to Treat) Analysis Set

The ITT analysis set includes all subjects that are enrolled in the study. All safety analyses (other than primary analysis for primary safety endpoints) will be based on the ITT analysis set.

7.1.2 mITT (modified Intention to Treat) Analysis Set

The mITT analysis set is a subset of the ITT analysis set. The mITT analysis set will consist of subjects who have entered the hybrid closed-loop phase of the study successfully. The mITT analysis set will be used as the primary analysis set for all primary and secondary endpoints (both safety and effectiveness), and other clinical outcome data.

7.1.3 Per Protocol

The Per-Protocol (PP) analysis set is a subset of the mITT analysis set. Subjects will be included in the PP analysis set if they have a minimum of 80% system use during the hybrid closed-loop phase inclusive of manual and automated modes over a duration of 14 days and have completed the study without major protocol deviations. The PP analysis set will be used as supportive analysis for the effectiveness endpoints. The following will be considered major protocol deviations:

- Major inclusion/exclusion criterion deviation
- Significant protocol non-compliance that may confound the study objective data (e.g., use of prohibited medications)

The list of subjects excluded from the PP analysis set will be determined prior to analysis. If the PP analysis set does not differ from the mITT analysis set, separate analyses will not be presented.

7.2 CONTROL OF SYSTEMATIC BIAS

Several measures are incorporated into the study design to help minimize study bias as follows:

1) This is a multi-center trial to help ensure that investigator or site or subject enrollment bias is minimized. Selection of subjects will be made from the Investigator's usual subject load. Consecutively eligible subjects should be enrolled into the study.
2) This document specifies appropriate statistical methodology to ensure that bias is minimized.
3) The effectiveness measures will be based on the direct output from the device, ensuring objective reporting.
4) An independent Medical Monitor will adjudicate all SAEs and device related adverse events; the Medical Monitor's assessment of adverse events will be used for regulatory reports, product labeling, and publications or presentation.

7.3 POOLING DATA ACROSS CENTERS

Up to 8 clinical sites will enroll subjects into the study. The analyses will be presented using data pooled across study centers. A formal assessment of the poolability of subjects across sites will not be performed.

7.4 CALCULATION OF PERCENTAGE OF TIME IN RANGE

Several effectiveness endpoints involve calculation of percentage of time in a specific glycemic range. These endpoints will be based on the direct output from a device (either a CGM or The Omnipod Horizon™ Automated Glucose Control System). The percentage of time in range (TIR) will be calculated as:

$$100 \times \frac{\#\ of\ CGM\ records\ in\ range}{\#\ of\ evaluable\ CGM\ records} = TIR\%$$

The following CGM records will be excluded from analysis, and therefore, from calculation of TIR endpoints:

- No glucose value is provided in the device output, such as due to an error or device deficiency during which the device does not record glucose readings
- CGM records reported prior to start of standard therapy phase or after subject's discontinuation from HCL

7.5 HANDLING OF MISSING DATA

All practical monitoring and follow-up steps will be taken to ensure complete and accurate data collection. It is anticipated that the rate of missing data would be very low due to the limited duration of the study. All analyses will be based on available data only; no imputation for missing data is planned.

7.6 OTHER DATA SUMMARIES

The distribution of each baseline characteristic or demographic parameter of interest (such as age, gender, medical history, etc.) will be presented. Data on all enrolled subjects will be presented. If a considerable number of subjects do not enter the HCL phase of the study, these results may also be presented for the mITT and/or PP analysis sets. Continuous variables will be summarized using count, mean, median, standard deviation, and range. Categorical variables will be summarized using counts and percentages.

Statistical testing will be performed to assess significant change from Baseline in various measures. Any testing performed will be based on one-sample t-test for continuous variables, or chi-square test or Fisher's exact test for categorical variables, using a two-sided significance level of 5%. If the

observed data is found not to follow a normal distribution, non-parametric methods may be employed (such as Wilcoxon rank sum test as appropriate).

7.7 STATISTICAL SOFTWARE

The statistical software package SAS® 9.4 or later will be used for all the data derivations, summarization, data listings and statistical analyses. Additional statistical software may be used for graphics or validation purposes as appropriate.

7.8 PERIODIC REVIEW OF CONTROLLER DATA

Following the upload of controller data into the database, the data will be made available to Sponsor representatives for periodic reviews. Reviews of the controller data will be conducted to identify any anomalies and to confirm the algorithm appears to be functioning as expected. These reviews, including any findings, will be documented. Unless potential safety concerns are identified, the results of such reviews will not be shared with study sites or subjects, and will not affect the data analyses.