Medtronic

Study Title	ADvanced Hybrid Closed Loop study in Adult Population with Type 1 Diabetes (ADAPT)
NCT Number	NCT04235504
Document Description	Clinical Biostatistical Plan (Version 2.0)
Document Date	08-JUN-2021

Medtronic **Statistical Analysis Plan**	
Clinical Investigation Plan Title	**ADAPT** **AD**vanced Hybrid Closed Loop study in **A**dult **P**opulation with **T**ype 1 Diabetes: *A Prospective, Open-label, Multi-center, Adaptive, Confirmatory and Randomized Controlled Study*
Clinical Investigation Plan Identifier	327
Clinical Investigation Plan Version	H dated 17 SEP 2020
Sponsor/Local Sponsor	Medtronic International Trading Sàrl. ("Medtronic") Route du Molliau 31 1131 Tolochenaz, Switzerland
Document Version	Version B dated 8 JUNE 2021

Table of Contents

1. Version History

Version	Summary of Changes	Author(s)/Title
A	• Initial Release	[redacted], Senior Biostatistician
B	• Superiority is added in section 4.1. to clarify that superiority test will be performed for the primary analysis to comply with the new standard ISO14155:2020. • The specific criteria for exclusion of subjects from per protocol set is added in section 7.1.3.2. • The definition of the day-time and night-time is changed from (06:01 to 23:59) and (00:00 to 06:00) to (06:00 to 23:59) and (00:00 to 05:59). • A miniumum of 10 days sensor data was planned to evaluate CGM related endpoints, however we changed to a minimum of 7 days as the number of subjects with sensor data < 10 days was higher than expected at Baseline. Using data from the SMILE study, we evaluated if there was difference between ≥10 and ≥7 days and similar results were observed, indicating not significant difference. This is updated in section 7.4. • The definition of Serious Adverse Event, Unanticipated Serious Adverse Device Effect (USADE), Device deficiency in section 7.9.3 is updated to the new standard ISO14155:2020.	[redacted], Senior Biostatistician

2. List of Abbreviations and Definitions of Terms

Abbreviation	Definition
AE	Adverse Event
ADAPT	**AD**vanced Hybrid Closed Loop study in **A**dult **P**opulation with **T**ype 1 Diabetes
AHCL	Advanced Hybrid Closed Loop
AUC	Area Under Curve
BG	Blood Glucose
BMI	Body Mass Index
CIP	Clinical Investigation Plan
CGM	Continuous Glucose Monitoring
DKA	Diabetic Ketoacidosis
DMC	Data Monitoring Committee
DQoL	Diabetes Quality of Life questionnaire
DTSQ	Diabetes Treatment Satisfaction Questionnaire
DTSQc	Change Version of DTSQ
DTSQs	Status Version of DTSQ
eCRF	Electronic Case Report Form
EMEA	Europe Middle East Africa
EOS	End of Study
FGM	Flash Glucose Monitoring
HbA1c	Glycosylated hemoglobin
HFS	Hypoglycemia Fear Survey
ICG	International Consensus Group
ICU	Intensive Care Unit
ITT	Intent to Treat
BOCF	Baseline Observation Carry Forward
MAGE	Mean Amplitude of Glycemic Excursions
MDI	Multiple Daily Injections
MNHE	Mean Number of Hypoglycemic Events
OC	Oracle Clinical
PIC	Patient Information and Informed Consent Form
PP	Per Protocol
RT	Real-Time
SAE	Serious Adverse Event
SADE	Serious Adverse Device Events
SAP	Statistical analysis plan
SG	Sensor Glucose
SMBG	Self-Monitoring of Blood Glucose
T1DM	Type 1 Diabetes Mellitus
TIR	Time in Range
TLFs	Tables, Listings and Figures
USADE	Unanticipated Serious Adverse Device Effect

3. Introduction

In patients with insulin dependent diabetes mellitus, glycemic control is influenced by numerous factors, such as insulin dosage, insulin absorption, timing, physiological/ lifestyle factors such as exercise, food intake, hormones and illness. These factors may contribute to significant variability in insulin requirements, which makes self-management of type 1 diabetes challenging.

Patients who are using continuous glucose monitoring, including sensor-augmented pump therapy, experience improvements in glycemic control. Advanced features of sensor-augmented pump therapy are now being used in clinical practice; these include automatic suspension of insulin delivery when a pre-set glucose threshold is reached (suspend on low) or is predicted to be reached (suspend before low). Both approaches have shown that a significant reduction in the risk and burden of hypoglycemia can be achieved, especially in patients who are prone to experiencing hypoglycemia.

Parallel to these approaches to mitigate the risk of hypoglycemia, more progressive advancements in technology can link insulin delivery directly to glucose levels. Closed-loop insulin delivery is different from conventional pump therapy and low glucose management technology, because it uses a control algorithm to automatically adjust insulin delivery based on subcutaneous sensor data to improve diabetes management. Manual meal-time announcement and prandial insulin boluses still need to be carried out by patients in order to overcome the delay in insulin action of currently available insulin analogues administered subcutaneously. The 'hybrid' closed-loop approach is in contrast to a 'fully' closed-loop approach, in which user input to the control algorithm related to meals would no longer be required.

One arm of the study will assess the MiniMed™ 670G running in advanced hybrid closed loop (AHCL) mode. The MiniMed™ 670G hybrid closed loop system is currently in commercial distribution in the United States and in an increasing number of European countries. Real world data from the MiniMed™ 670G in the United States has documented safety and efficacy in adults and children. There have, however, been further advancements to the hybrid close loop proportional integrative derivative (PID) algorithm model that seek to improve functionality and efficacy based on retrospective analysis of data from the MiniMed™ 670G insulin pump, using a modified proportional integrative derivative model, with insulin feedback and additional safety features. These enhancements intended to maximize time spent in hybrid closed-loop operation, in order to further improve glucose control and overall user satisfaction. Advancement have been implemented in the new system with the algorithm version 4.0, such as automatic correction bolusing, sensor glucose based meal bolusing, automatic calibrations of Blood Glucose (BG) measurements transmitted to the pump and a variable target for automatic basal deliveries, all of which are intended to contribute to a reduction of unnecessary Auto Mode exits, which will consequently increase time in euglycemic range and overall user satisfaction.

Several studies are currently ongoing on AHCL MiniMed™ 670G System to demonstrate the safety and effectiveness on the AHCL MiniMed™ 670G System version 4.0 . Nevertheless, additional clinical evidence

is required to support the available efficacy and safety data with additional long-term data of the AHCL system in comparison with the current standard of care for Type 1 Diabetes.

Across the world, the majority of Type 1 Diabetes patients requiring insulin are treated with Multiple Daily Injection (MDI) therapy. While a growing population have access to diabetes technology through Flash Glucose Monitoring (FGM) or Continuous Glucose Monitoring (CGM) in the EMEA region, a large proportion of insulin requiring patients with Type 1 Diabetes on MDI therapy with FGM or CGM still have sub-optimally controlled diabetes management with a HbA1c > 8.0 % (64 mmol/mol). Additional long-term efficacy and cost-effective data will be needed to further support market access and adoption of AHCL therapy for those patients who are not well controlled with the current standard of care for Type 1 diabetes patients requiring insulin.

This Statistical Analysis Plan (SAP) will be used to support the final report and analysis of the study phase of the ADAPT study. The Statistical Analysis Plan has been designed to document, before data is analyzed, the planned analyses for the final report. This SAP does not limit the analysis in reports, and additional analyses of the study data beyond this plan might be conducted. However, this document provides the basis for the statistical sections of the final report. Analyses not planned in the SAP and incorporated into the final report will be referred to as "not-prespecified". Further, in case any analyses will be done differently than planned in the CIP or SAP, an explanation will be provided in the final report. A separate SAP will be prepared for the continuation phase of the study.

4. Study Objectives

4.1 Primary Objective

The primary objective of the study is to evaluate the superiority on glycemic control of the AHCL system versus MDI + FGM with patients with Type 1 Diabetes.

4.2 Exploratory Objective

A second objective of the study is to conduct an exploratory analysis on the efficacy of the AHCL system versus MDI + RT-CGM with patients with Type 1 Diabetes.

5. Investigation Plan

This is a pre-market, multi-center, prospective, open label, adaptative, confirmatory, randomized controlled trial in patients with Type 1 diabetes. As mentioned in section 4, the purpose of the study is to evaluate efficacy of the Advanced Hybrid Closed Loop (AHCL) in comparison with Multiple Daily Injection (MDI) therapy with Flash Glucose Monitoring (FGM) or Continuous Glucose Monitoring (CGM).

The study consists of two separate cohorts, Cohort A with subjects on MDI + FGM (confirmatory part of the study) and Cohort B with subjects on MDI + Real-Time CGM (exploratory part of the study). As shown in Figure 1, the study has three phases, a run-in phase, a study phase and a continuation phase. The purpose of the run-in phase is to collect CGM baseline data while subjects are on their current

therapy. During the 6-month study phase, subjects will be randomized to continue with their current therapy or to start using the AHCL system.

Cohort A (confirmatory): Subjects on **MDI + FGM** will be randomized into:

- Treatment Arm: Start AHCL (and stop FGM at visit 6A)
- Control Arm: Continue MDI + FGM

Cohort B (exploratory): Subjects on **MDI + Real-Time CGM** will be randomized into:

- Treatment Arm: Start AHCL (and stop CGM at visit 6A)
- Control Arm: Continue MDI + CGM

Blinded CGM data will be collected at 3 and 6 months after randomization for subjects in the MDI therapy with FGM or CGM. For the duration of the 6-month continuation phase, all enrolled subjects will be using the AHCL system. Due to the uncertainty about the magnitude of the treatment effect and the standard deviation, the sample size will be re-estimated by an independent DMC in an interim assessment based on conditional power after a minimum of 32 subjects in cohort A (approximately 16 subjects in the treatment arm and 16 subjects in the control arm) have been randomized and completed the study phase of 6 months. Early stopping is only considered in case of futility. No sample size decrease is considered for efficacy. Details of the interim analysis and adaptive design are available in the DMC Analysis Plan.

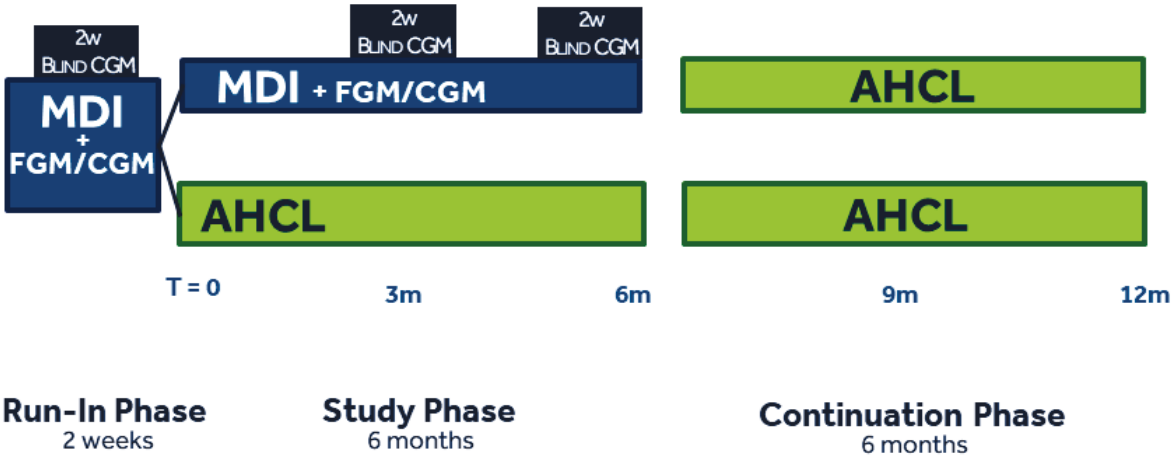

Figure 1. Study Design Overview

6. Determination of Sample Size

For Cohort A (MDI+FGM), the sample size calculation was performed based on the following assumptions: alpha=0.05, power=80%, 0.7 standard deviation and 0.5 reduction in mean change of HbA1c in the treatment arm, as compared to the reduction in control arm, which is the minimum clinically meaningful difference. The standard deviation was based on the *Eurythmics study* (Hermanides, 2011) comparing SAP therapy (Treatment arm) vs MDI (control arm) with a pooled standard deviation of 0.83.

Based on these assumptions, the minimum sample size required is 64 subjects (32 in each arm) in Cohort A comparing AHCL with MDI + FGM subjects.

The following drop-out assumptions have been taken into account, based on experience with previous studies: At screening: 10%; After run-in: 5%; During 6-month follow-up: 7.5%

Incorporating these drop-out rates, a total of 84 subjects need to be screened in Cohort A (AHCL vs MDI+FGM) in order to have 74 subjects starting the run-in phase, 70 subjects randomized to start the treatment phase and 64 subjects to complete the study phase (6 months).

For the exploratory analysis in Cohort B, comparing MDI + CGM versus AHCL, no sample size calculation was performed and 30 subjects will be required for analysis. Incorporating the expected dropout rates, a total of 40 subjects need to be screened, in order to have 36 subjects starting the run-in phase, 34 subjects randomized to start the treatment phase and 30 subjects to complete the study. Subjects will be randomized in a 1:1 ratio to the 2 arms.

7. Statistical Methods

The methods in this section are applicable to cohort A except the summary statistics of percentage of sensor readings in MDI + RT-CGM control arm in section 7.8.3.10. These methods are also applicable to cohort B where all the endpoints (except endpoints that are specific to cohort A) will be assessed in an exploratory manner including the difference in the mean HbA1c change (6 months - baseline) between the AHCL and the MDI + RT-CGM arms on ITT basis. However, the following will not be conducted for cohort B:

1. Center pooling test
2. Sensitivity analysis for missing data for the primary and secondary analysis
3. Interim analyses for sample size re-assessment
4. Adjustments for Multiple Comparisons

7.1 Study Subjects

7.1.1 Disposition of Subjects

A subject is enrolled in the study when he/she signs and dates the Patient Informed Consent. Summary of number of subjects enrolled, entered the run-in phase, randomized and completed (by study arm), screen failures and drop-outs will be reported. List of subjects that failed to meet randomization criteria, screen failures and drop-out with the corresponding reasons will also be reported.

A consort flow diagram similar to figure 2 will be created to describe patient disposition.

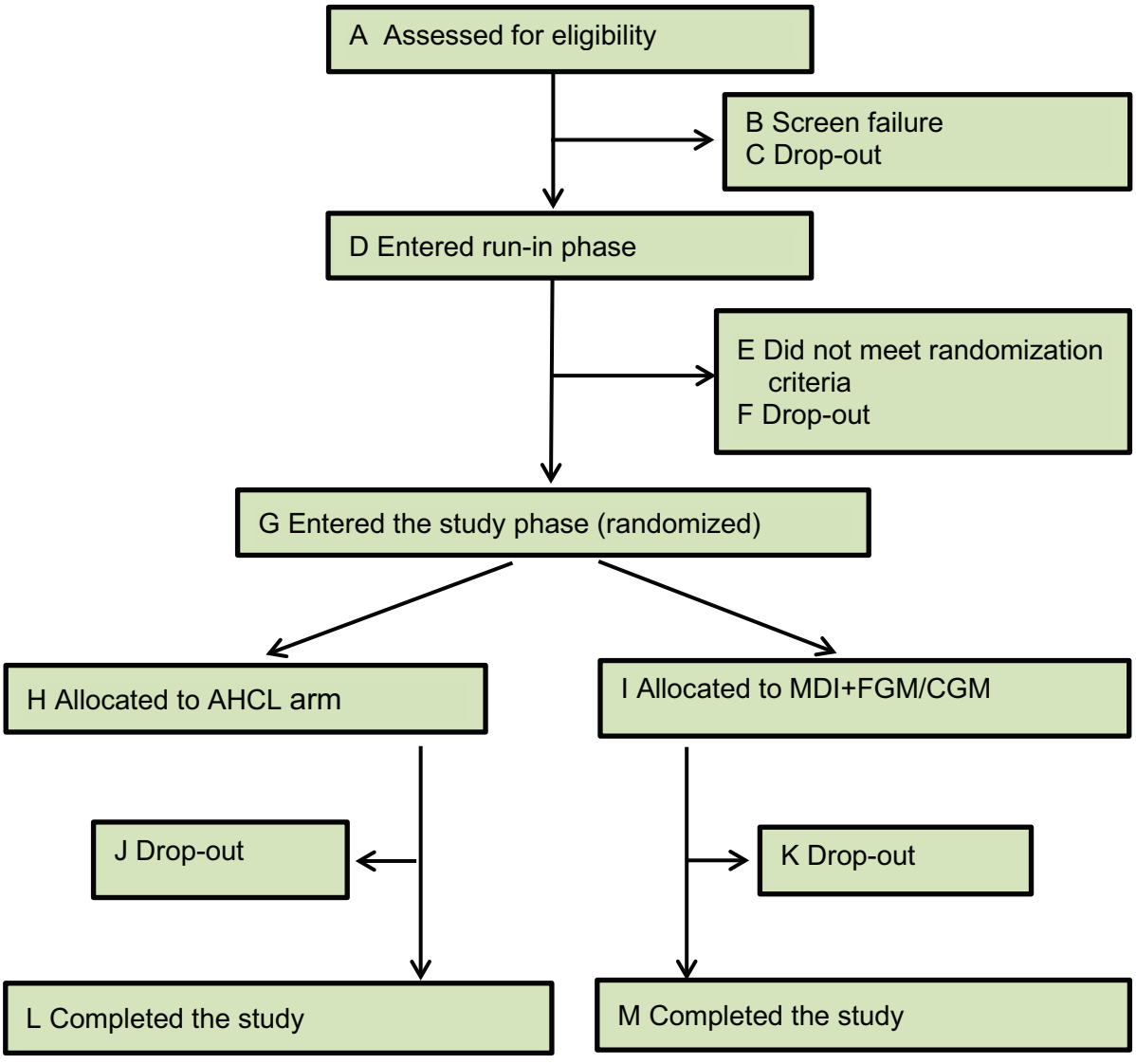

Figure 2. Consort Diagram of Subject Disposition

7.1.2 Clinical Investigation Plan (CIP) Deviations

Deviations from the clinical investigation plan will be collected as deviations on the Study Deviation eCRF. Deviations will be summarized in the final report in a table by category. The number of deviations per category, the number and percentage of subjects with a deviation in each category will be reported. Listing of deviations will also be reported.

7.1.3 Analysis Sets

7.1.3.1 Efficacy analysis set

For the primary, secondary and ancillary endpoints, efficacy analyses will be performed in the Intent to Treat (ITT) basis. The ITT set will be composed of all randomized subjects. These subjects will be assessed and analyzed as members of the intended randomized arm, irrespective of their compliance to the planned course of treatment or deviations from protocol.

7.1.3.2 Per Protocol set

For sensitivity purposes, efficacy analysis of the primary endpoint will also be performed on the Per Protocol set. The Per Protocol (PP) set is composed of randomized subjects with no major study deviation. Below is the list of criteria used to exclude subjects from per protocol analysis set. Subjects with at least one of the listed criteria will be excluded from PP.

- **Only for Treatment Arm Subjects**
  - Auto Mode use during the 6 months study phase is less than 75%.
  - Mean number of boluses per day during the 6 months study phase is less than 1.5.
  - Sensor use during the 6 months study phase is less than 70%.
- **Only for Control Arm Subjects**
  - Percentage of FGM sensor reading is less than 70 %.
- **For Both Control and Treatment Arms**
  - Subject did not complete the 6 months study phase.
  - Subject did not meet screening criteria but continued the study.
  - Subject did not meet randomization criteria but continued the study.
  - Subject randomized to wrong cohort.

7.1.3.3 Safety set

Safety analysis will include all subjects with signed inform consent and safety data will be presented by phase (run-in phase and study phase). Safety reporting will be done separately for the run-in period and the study phase period. Any AE and SAE occurring from the date of Visit 1 to before the date of Visit 5 will be reported as happening in the run-in phase. Any AE and SAE occurring from the date of visit 5 to the date of Visit 9 will be reported as happening in the study phase and will be reported according to the randomization arm that the subject was assigned.

For comparison of safety in the two arms, the safety set is composed of all randomized subjects, and assessed and analyzed as members of the intended randomized arm.

7.2 General Methodology

Summary statistics for Continuous variables will be represented by number of subjects(n), mean, median, standard deviation, minimum and maximum and categorical variables will be represented by counts and percentages.

P-values for hypothesis testing will be evaluated based on two-sided testing using significance level of 0.05. Confidence intervals will be reported as two-sided 95% confidence intervals. Normality will be verified for appropriate statistical methodology.

The templates for Tables, Listings and Figures (TLFs) will be available in the TLFs document.

7.3 Center Pooling

The study is expected to be conducted in three countries. The same clinical investigational plan and training is followed and standardized data collection methodology and electronic case report forms (eCRF) are used in all sites. Descriptive summary of the primary endpoint will be reported by country. Available data will be used and in case of missing data, no imputation will be performed. The variation in primary endpoint between the countries will be evaluated using qualitative treatment by country interaction by Gail Simon test. In the absence of significance for this test, the primary analysis will be evaluated using the overall treatment effect estimate.

7.4 Handling of Missing, Unused, and Spurious Data and Dropouts

Primary endpoints and Secondary endpoints will be collected at baseline, end of 3 months and end of 6 months. The main analysis is based on ITT (Intention to Treat) using a random effect model that uses available data and accounts for missing at random. In addition, sensitivity analyses will be performed.

For the primary analysis (change of A1c), if A1c at 3 months or 6 months is not available in subjects that drop-out of the study, the following sensitivity analyses will be carried out:

- Per Protocol Dataset
- Baseline Observation Carry Forward (BOCF)
- Data will be imputed by selected A1c data from the same arm considering baseline characteristics such as age, gender and duration of diabetes

For per protocol analysis, if A1c is collected out of window (± 1.5 months from scheduled visit) A1c will be considered missing for that particular period/visit.

For the secondary endpoints during the two weeks period of SG , a minimum number of 2016 of SG values (288 SG values/day × 7 days out of 14 days) is required to evaluate the glycemic outcome at each period

(baseline, 3 months and 6 months). For those subjects with less than 2016 SG during the two weeks of sensor wear, the glycemic outcome will be considered missing at that period in the main analysis. The following sensitivity analysis will be carried out to mitigate the impact of data loss:

- CGM data gap filled within the same subject: Missing CGM data in one arm will be imputed by randomly replaced CGM data in the same subject during the same period of sensor wear.
- CGM data gap filled within the arm, ignoring baseline characteristics: Missing CGM data will be imputed by a randomly selected CGM data from the same arm, ignoring baseline characters during the same period of sensor wear.
- CGM data gap filled within the arm, considering baseline characteristics: Missing CGM data will be imputed by selected CGM data from the same arm, considering baseline characters such as age, gender and duration of diabetes during the same period of sensor wear.

As a sensitivity analysis, secondary endpoints will also be evaluated using a minimum of 10 days of sensor data instead of 7 days.

In case of dates collected at baseline with missing day and/or month the next procedure will be applied. A missing day and month will be imputed using the month 'July 1' and a missing day will be imputed with the day '15'. This applies only to dates related to medical history and baseline information with missing month and/or day.

For tables and listings of safety data a conservative/worst case scenario approach will be taken in case of partially missing dates. For date of discharge with missing day, the day will be set to last day of that month. For date of admission with missing day, the day will be set to first day of that month. If that is before day of start of study phase (within the same month), then day will be set equal to day of start of study phase. Otherwise, if it is before day of start of run-in (within the same month), then it will be set to day of start of run-in. Thus, duration is set to be as long as possible.

7.5 Adjustments for Multiple Comparisons

The following hierarchical test procedure reflects the relative importance of the endpoints and controls for multiplicity.

Fixed sequential testing of primary and selected secondary endpoints

For the following endpoints, the procedure test hierarchically the ordered hypotheses in sequence at level α=0.05 until a first hypothesis is non-rejected.

Primary endpoint

1. Change in HbA1c

Change in HbA1c will be tested for superiority as described in section 7.9.1 and a p-value < 0.05 will be considered statistically significant. If p-value < 0.05, continue to next test, else stop.

Secondary endpoints

2. Percentage of time spent in hyperglycemic range with SG > 250 mg/dL

   Non-inferiority test with non-inferiority margin of 6%, if p-value < 0.05 reject null hypothesis and continue, else stop

3. Percentage of time spent in hyperglycemic range with SG > 250 mg/dL

   Superiority test, if p-value<0.05 reject null hypothesis and continue, else stop

4. Percentage of time spent in hyperglycemic range with SG > 180 mg/dL

   Non-inferiority test with non-inferiority margin of 6%, if p-value < 0.05 reject null hypothesis and continue, else stop

5. Percentage of time spent in hyperglycemic range with SG > 180 mg/dL

   Superiority test, if p-value<0.05 reject null hypothesis and continue, else stop

6. Percentage of time spent within range 70 - 180 mg/dL

   Non-inferiority test with non-inferiority margin of 6%, if p-value<0.05 reject null hypothesis and continue, else stop

7. Percentage of time spent within range 70 - 180 mg/dL

   Superiority test, if p-value<0.05 reject null hypothesis and continue, else stop

8. Percentage of time spent in hypoglycemic range with SG < 54 mg/dL

   Non-inferiority test with non-inferiority margin of 2 %, if p-value < 0.05 reject null hypothesis and continue, else stop

9. Percentage of time spent in hypoglycemic range with SG < 70 mg/dL

   Non-inferiority test with non-inferiority margin of 5 %, if p-value < 0.05 reject null hypothesis and continue, else stop

Exploratory analysis

Superiority test for percentage of time spent in hypoglycemic range with SG < 54 mg/dL, < 70 mg/dL and analyses on ancillary endpoints and safety endpoints may be performed, and p-values will be reported but may not be claimed.

7.6 Demographic and Other Baseline Characteristics

A summary of basic subject demographics, medical history and other baseline characteristics will be reported using appropriate summary statistics (number of subjects (n), mean, standard deviation, median, minimum and maximum for continuous variables; frequency and percentages for categorical variables). Baseline summary for all subjects and by treatment group will be presented.

Demographics, medical history and other baseline characteristics will be obtained from eCRF and baseline sensor data. Baseline variables to be summarized include, but are not limited to: age, sex, weight, BMI, country of enrollment, HbA1c, Diabetes history, Creatinine Clearance value and total insulin dose per day.

7.7 Interim Analyses

An interim analysis for sample size re-estimation will be based on the conditional power approach by Li et al. (2002) (Li, Shin, Xie, & Lu, 2002) and a method by Chen et al. (2004) (Chen, DeMets, & Lan, 2004) as extended by Mehta and Pocock (2011) (Mehta & Pocock, 2011). All the details regarding the interim analysis can be found in the Statistical Analysis Plan of the Data Monitoring Committee.

7.8 Evaluation of Objectives

7.8.1 Analysis of primary endpoint

The goal is to compare the HbA1c change from baseline to end of study (6 month) between the two study arms, MDI+FGM and AHCL. The between group difference in the mean HbA1c change is of primary interest and will be compared using the following hypotheses:

Null-hypothesis: H_o: $\Delta HbA1c_{AHCL} = \Delta HbA1c_{MDI\ therapy + FGM}$

Alternative hypothesis: H_A: $\Delta HbA1c_{AHCL} \neq \Delta HbA1c_{MDI\ therapy + FGM}$

Where $\Delta HbA1c_{AHCL}$ is the mean change in HbA1c (6 months [Visit 8]- baseline [Visit 1]) in the AHCL group and $\Delta HbA1c_{MDI\ therapy + FGM}$ is the mean change in HbA1c (6 months [Visit 8]) - baseline [Visit 1] in the MDI + FGM group. The null hypothesis will be tested against the alternative hypothesis using a linear mixed model. Superiority of AHCL will be concluded if $\Delta HbA1c_{AHCL} < \Delta HbA1c_{MDI\ therapy + FGM}$.

7.8.2 Analysis of secondary endpoints

The blinded (for the control arm) or unblinded (for the treatment arm) sensor data collected during the 2 periods of 2 weeks will be used to calculate the endpoints.

7.8.2.1 Percentage of Time in Range (70-180 mg/dL), in Hyperglycemia (> 180 mg/dL, > 250 mg/dL) and Hypoglycemia (< 54 mg/dL, < 70 mg/dL)

Percentage of time spent within range with sensor glucose (SG) between 70 - 180 mg/dL (3.9-10.0 mmol/L) , in hyperglycemic range with SG > 180 mg/dL (> 10.0 mmol/L), > 250 mg/dL (13.9 mmol/L) and in hypoglycemic range with SG < 54 mg/dL (3.0 mmol/L), < 70 mg/dL (3.9 mmol/L) will be computed using the sensor data collected during each of the 2 periods of 2 weeks. The last 4032 sensor readings (2 weeks) from each period will be used for analysis.

The percentage of time spent within range (70-180 mg/dL) for subject i at period j will be calculated as:

$$\frac{\text{Total number of sensor reading with SG between } 70-180 \text{ mg/dL for subject } i \text{ at period } j}{\text{Total number of sensor readings for subject } i \text{ at period } j}$$

The percentage of time spent in hyperglycemic range > 180 mg/dL or > 250 mg/dL for subject i at period j will be calculated as:

$$\frac{\text{Total number of sensor reading with SG} > 180 \text{ mg/dL or} > 250 \text{ mg/dL for subject i at period j}}{\text{Total number of sensor readings for subject i at period j}}$$

Similarly, the percentage of time spent in hypoglycemic range < 54 mg/dL or < 70 mg/dL for subject i at period j will be calculated as:

$$\frac{\text{Total number of sensor reading with SG} < 54 \text{ mg/dL or} < 70 \text{ mg/dL for subject i at period j}}{\text{Total number of sensor readings for subject i at period j}}$$

The between group difference in these secondary endpoints will be compared for the overall 3 and 6 months using a linear mixed model adjusting by baseline value of the endpoints. In case of partially missing baseline measurements, the method proposed by White and Thompson will be used (Adjusting for partially missing baseline measurements in randomized trials, 2005, Statistics in Medicine).

7.8.3 Ancillary Endpoints

Analysis for endpoints with repeated measures will be using a random effect model (similar analysis as in section 7.8.2.1) that uses available data and accounts for missing data and no imputation of missing value will be performed. Analysis for endpoints with single measure will be using analysis of variance (ANOVA) and multiple imputation will be used for missing values. For CGM related endpoints, the 2 weeks blinded (for the control arm) and unblinded (for the treatment arm) sensor data collected during the 2 periods will be used.

7.8.3.1 Percentage of Time in Ranges

Percentage of time spent in 70 - 140 mg/dL (3.9 -7.8 mmol/L) range will be analyzed and reported. The percentage of time will be calculated in a similar way as in section 7.8.2.1. The percentage of time in 70 - 140 mg/dL and 70 - 180 mg/dL during the day-time and night-time will also be calculated, analyzed and reported.

7.8.3.2 Percentage of Time and Area Under the Curve (AUC) in Hypoglycemic and Hyperglycemic range for several threshold

Percentage of time spent and mean daily area under the curve (AUC) in hyperglycemic range with SG > 140 mg/dL (7.8 mmol/L), >350 mg/dL (19,4 mmol/L) and AUC in hypoglycemic range with SG < 54 mg/dL (3.0 mmol/L), <70 mg/dL (3.9 mmol/L) and in hyperglycemic range with SG >180 (10 mmol/L) , >250 mg/dL (13,9 mmol/L) will be analyzed and reported. The percentage of time will be calculated in a similar way as in section 7.8.2.1. Mean daily AUC below/above a threshold for subject i at period j will be calculated as:

$$\frac{\text{total AUC below/above a threshold for subject i in period j}}{\text{duration of sensor wear in days (number days with CGM data) for subject i at period j}}$$

Where the duration of sensor wear in days for subject i at period j is the number of sensor readings (capped at 4032) for subject i at period j divided by 288.

Each of these endpoints as well as percentage of time spent in hypoglycemic range with SG < 54 mg/dL (3.0 mmol/L), <70 mg/dL (3.9 mmol/L) and in hyperglycemic range with SG >180 (10 mmol/L), >250 mg/dL (13,9 mmol/L) during the day-time (06:00 to 23:59) and night-time (00:00 to 05:59) will also be calculated, analyzed and reported.

7.8.3.3 Number of Biochemical Hypoglycemic events

A biochemical hypoglycemic event with SG < 54 mg/dL is defined as sensor glucose values less than 54 mg/dL (3.0 mmol/L) for 15 or more consecutive minutes. A biochemical hypoglycemic event with SG < 70 mg/dL is defined as sensor glucose values less than 70 mg/dL (3.9 mmol/L) for more than 20 consecutive minutes. When the time between two successive events is less than 30 minutes, they will be combined and counted as one event. The mean number of biochemical hypoglycemic events with SG < 54 mg/dL (3.0 mmol/L), < 70 mg/dL (3.9 mmol/L) per subject per week based on 15 and 20 consecutive minutes will be computed using the sensor data collected during each of the 2 periods of 2 weeks. The last 4032 sensor readings (2 weeks) from each period will be used for analysis. The mean number of hypoglycemic events (MNHE) per week at each period for each subject will be calculated as:

$$\frac{\text{number of biochemical hypoglycemic event for subject i at period j}}{\text{duration of sensor wear for subject } i \text{ at period } j \text{ in weeks}}$$

Where the duration of sensor wear for subject i at period j is the number of sensor readings (capped at 4032) for subject i at period j divided by 2016.

Mean number of biochemical hypoglycemic events with SG < 54 mg/dL (3.0 mmol/L) based on 20 minutes definition and < 70 mg/dL (3.9 mmol/L) per week based on 15 minutes definition will be analyzed and reported as additional analyses. Analysis on mean number of biochemical hypoglycemic events with SG < 54 mg/dL (3.0 mmol/L) (based on 15 minutes) will be using non-inferiority test with non-inferiority margin of 35% of the observed mean in the control group. Each of these endpoints during the day-time and night-time will also be calculated, analyzed and reported.

7.8.3.4 Mean of Sensor glucose Values

Mean number of sensor glucose (SG) in mg/dL will be analyzed and reported. The mean SG for a subject at period j will be calculated as the overall mean of the SG values (maximum of 4032 readings closer to the visit) during the period. The mean SG during the day-time and night-time will also be calculated, analyzed and reported.

7.8.3.5 Coefficient of Variation of Sensor Glucose Values (CV)

The coefficient of variation of sensor glucose values (CV) will be analyzed and reported. The CV for subject i at period j will be calculated as:

$$\frac{\text{standard deviation of sensor glucose values for subject i at period j}}{\text{mean of sensor glucose values for subject i at period j}}$$

The mean of sensor glucose values for subject i at period j will be calculated as described in section 7.8.3.4. The standard deviation of sensor glucose values for subject i at period j will be the standard deviation of all sensor readings (with maximum of 4032 recent readings) during period j.

7.8.3.6 Amplitude of Glycemic Excursions (MAGE)

The excursion amplitudes of the glucose values measured by mean amplitude of glycemic excursions (MAGE) will be calculated for each period using the blinded 2 weeks sensor glucose data. Analysis will be performed and reported.

7.8.3.7 Percentage of Time Spent in Auto Mode and in Manual Mode

The percentage of time spent in auto mode and manual mode for subjects in the AHCL arm will be calculated as:

$$\frac{\text{time spent in auto/manual mode}}{\text{total time in AHCL therapy}} \text{ x } 100\%$$

Descriptive statistics for the percentage of time in auto mode and manual mode overall as well as during the day-time and night-time will be reported.

7.8.3.8 Percentage of Sensor Use for AHCL Arm

Descriptive statistics for the percentage of sensor use during the 6 months (visit 6A to visit 8) study phase will be reported for AHCL group.

7.8.3.9 Number of Scans and Percentage of Sensor Readings for MDI + FGM Control Arm

Descriptive statistics for the mean number of scans per day and percentage of sensor reading for MDI + FGM control group will be reported. The mean number of scans per day and the percentage of sensor readings for MDI+ FGM will be collected from CRF data. This sensor reading is from summary report of FGM.

7.8.3.10 Percentage of Sensor Readings for MDI + RT-CGM Control Arm

Descriptive statistics for the percentage of sensor reading for MDI + CGM control group will be reported. The percentage of sensor readings for MDI+ CGM will be collected from CRF data. This sensor reading is from summary report of RT-CGM not from the blinded sensor wear.

7.8.3.11 Change in HbA1c level by subgroups

Descriptive statistics for the change in HbA1c level from baseline to end of the 6 months follow-up will be reported by study arm and by:

- Age group
  - Grouping A. group 1: age <65, group 2: Age ≥ 65
  - Grouping B. group 1: age < median age, group 2 ≥ median age.
- Duration of diabetes
  - Grouping A. group 1: <10 years, group 2 ≥ 10 years
  - Grouping B. group 1: <20 years, group 2 ≥ 20 years
  - Grouping C. group 1: <30 years, group 2 ≥ 30 years

7.8.3.12 Change in Weight and BMI

The change in weight and BMI from baseline (Visit 1) to end of 6 months follow-up (Visit 8) will be calculated, analyzed and reported.

7.8.3.13 Change in Total Daily Insulin Dose

The change in total daily insulin dose from baseline to end of study phase will be computed, analyzed and reported. The total daily insulin dose will be based on self-reported data (from CRF) when in MDI therapy and will be based on pump data uploaded in CareLink Clinical when in AHCL therapy.

7.8.3.14 Mean number of SMBG

The mean number of self-monitored blood glucose (SMBG) readings and mean of SMBG values will be calculated in AHCL arm from the pump data uploaded in CareLink Clinical and will be reported using summary statistics for quantitative variables for the following window: visit 5A to visit 7A, visit 7A to visit 8, visit 5A to visit 8.

7.8.3.15 Diabetes related Consequences

Summary of diabetes-related number and mean duration of hospitalizations, number and mean duration of intensive care unit (ICU) care, number of emergency room admissions, number of events requiring ambulance assistance will be reported during the run-in phase and study phase (by study arm) categorized by reason of diagnosis.

7.8.3.16 Lost days from school or work

Summary number of lost days from school or work will be reported during the run-in phase and study phase by study arm.

7.8.3.17 Questionnaires' scores

The Hypoglycemia Fear Survey (HFS) score, Diabetes Treatment Satisfaction Questionnaire (status version, DTSQs and change version DTSQc) scores and Diabetes Quality of Life Questionnaire (DQoL) will be calculated in each study arm and reported using descriptive statistics for quantitative variables.

Hypoglycemia Fear Survey (HFS)

Hypoglycemia can lead to various aversive symptomatic, affective, cognitive, physiological, and social consequences, which in turn can lead to the development of possible phobic avoidance behaviors associated with hypoglycemia. The hypoglycemia fear survey (HFS) is a psychometric instrument designed to quantify this fear. The instrument has internal consistency and test-retest stability and varies with elevated glycosylated hemoglobin. The HFS has in most translations two subscales, the behavior subscale and the worry subscale and has a recollection period of 6 months.

The two subscales of the HFS are scored as: behavior subscale (first 15 items) items are added together, the 18 Worry subscale items are added together. Additionally, a total score, adding all 33 items, will be computed.
Descriptive statistics for the HFS score (behavior, worry and total) will be reported for Visit 2 (baseline), and visit 8 (end of 6 months of follow-up).

DTSQs and DTSQc

The Diabetes Treatment Satisfaction Questionnaire (DTSQ) has been specifically designed to measure satisfaction with diabetes treatment regimens in people with diabetes. The DTSQ [status version (DTSQs)] is an eight-item questionnaire, in which six questions (items 1, 4, 5, 6, 7 and 8) assess treatment satisfaction and the other two assess perceived frequency of hyper- and hypoglycemia.

Each item is scored from 6 (very satisfied) to 0 (very dissatisfied) such that the Treatment Satisfaction score can range from 36 (very satisfied) to 0 (very dissatisfied) and the perceived frequency of hyper- and hypoglycemia scores range from 6 (most of the time) to 0 (none of the time).

Although the DTSQs has proved highly sensitive to change, in many studies where patients are very satisfied with treatment used at baseline, the DTSQs cannot show improvements when they switch to a new treatment, even though they might be even more satisfied with the new treatment. To overcome the limitation of the DTSQs, a change version (DTSQc) has also been developed, which asks participants to rate how their current treatment compared with their previous treatment.

The DTSQc instrument contains the same 8 items as the DTSQs version. The difference lies in the wording of the response options and instructions, which, in the DTSQc, direct the respondent to compare their experience of treatment before the study began. All items are rated from +3 to -3.

The DTSQc instructions and response options differ from those of the DTSQs to produce measures of relative change in satisfaction rather than measures of absolute satisfaction:

Treatment satisfaction (change): items 1, 4, 5, 6, 7 and 8 are summed to produce a Treatment Satisfaction (change) score (range: +18 to -18). The higher the score, the greater the improvement in satisfaction with treatment; the lower the score, the greater the deterioration in satisfaction with treatment. A score of 0 represents no change.

A major advantage of the DTSQs and DTSQc is that they have been developed to be suitable for people with type 1 or type 2 diabetes using a wide range of treatments, including various methods of insulin delivery, oral medications and diet alone, and is, therefore, appropriate for use before and after patients switch between very different treatment regimens.

Descriptive statistics for DTSQs will be reported for Visit 2 (baseline) and visit 8 (end of 6 months of follow-up) and for DTSQc will be reported for visit 8.

Diabetes Quality of Life Questionnaire (DQoL)

DQoL has been widely used to measure quality of life among diabetes patients. The instrument has four scales: satisfaction with treatment, impact of treatment, worries about future effects of diabetes, and worries about social and vocational issues. The instrument also includes a generic health item that does not contribute to the scales. A score will be obtained for each dimension. A higher score represents higher quality of life.

Descriptive statistics for the DQoL score will be reported for Visit 2 (baseline) and visit 8 (end of 6 months of follow-up).

7.9 Safety Evaluation

7.9.1 Number of severe hypoglycemic events

Severe Hypoglycemia is an event requiring assistance of another person due to altered consciousness to actively administer carbohydrate, glucagon, or other resuscitative actions. This means that the subject was impaired cognitively to the point that he/she was unable to treat him or her self, was unable to verbalize his or her needs, and was incoherent, disoriented and/or combative. These episodes may be associated with sufficient neuroglycopenia to induce seizure or coma. Plasma glucose measurements may not be available during such an event, but 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.

Total number of severe hypoglycemic events reported in the eCRF, during Run-In (Visit 1 to before Visit 5/Visit 5A for Control/Treatment) and Study Phase will be reported. Total number of severe hypoglycemic events during the Study Phase (from date of Visit 5 /Visit 5A to date of Visit 9) will be reported by treatment arm.

The number of severe hypoglycemic events per year will be computed for each subject based on the entire study phase duration (happening after Visit 5 to and including Visit 9). If a subject i is followed up for x_i years and the number of observed severe hypoglycemic events during that follow up is m_i, the number of severe hypoglycemic events per year (SHEyear) for patient i will be estimated as:

$$\frac{m_i}{x_i}$$

Annualized crude incidence rates will be expressed as number of severe hypoglycemic events per 100 patients' year and will be calculated as:

$$\frac{\sum_{i=1}^{n} m_i}{\sum_{i=1}^{n} x_i} \times 100$$

where i=1, 2, …, n and n total number of subjects.
The number of severe hypoglycemic events per 100 patients' year during the study phase will be reported by study arm.

7.9.2 Number of diabetic ketoacidosis events

A diabetic ketoacidosis event is defined as an event of blood glucose greater than 250 mg/dL(13.9 mmol/L) arterial pH less than 7.3, bicarbonate less than 15mEq/l, moderate ketonuria or ketonemia, and requiring treatment within a health care facility.

Total number of diabetic ketoacidosis events will be reported for run-in and study phase categorized by study arm. The number of diabetic ketoacidosis events per 100 patients' year during the study phase will be calculated in a similar way as in section 7.9.1 and will be reported by study arm.

7.9.3 Number of Serious Adverse Events, Serious Adverse Device Effects, Unanticipated Serious Adverse Device Effects and Device Deficiencies

Serious Adverse Event (SAE) is adverse event that led to any of the following:

- **a)** death,
- **b)** serious deterioration in the health of the subject, users or other persons as defined by one or more of the following:
    1. a life-threatening illness or injury, or
    2. a permanent impairment of a body structure or a body function, including chronic disease, or
    3. in-patient or prolonged hospitalization, or
    4. medical or surgical intervention to prevent life-threatening illness or injury or permanent impairment to a body structure or a body function,
- **c)** fetal distress, fetal death or a congenital abnormality or birth defect including physical or mental impairment.

NOTE 1: Planned hospitalization for a pre-existing condition, or a procedure required by the CIP, without serious deterioration in health, is not considered a serious adverse event.

For the purpose of this study, the term "life-threatening" in the definition of "serious" refers to an event in which the patient was at risk of death at the time of the event; it does not refer to an event which hypothetically might have caused death if it were more severe.

Serious Adverse Device Effect (SADE) is adverse device effect that has resulted in any of the consequences characteristic of a serious adverse event.

Unanticipated Serious Adverse Device Effect (USADE) is serious adverse device effect which by its nature, incidence, severity or outcome has not been identified in the current e risk assessment.

Device deficiency is inadequacy of a medical device with respect to its identity, quality, durability, reliability, usability, safety or performance. It includes malfunctions, use errors, and inadequacy in the information supplied by the manufacturer including labeling. This definition includes device deficiencies related to the investigational medical device or the comparator.

The number of serious adverse events, serious adverse device effects, unanticipated serious adverse device effects, device deficiencies (DD), DD with SADE potential will be reported for run-in and study phase categorized by study arm. The listing of these events will also be reported.

7.10 Health Outcomes Analyses

Not applicable

7.11 Changes to Planned Analysis

This SAP will be executed in full but does not limit the analysis in reports, and additional analysis of the study data beyond this plan is expected. Analyses beyond the SAP will be identified as such and referred to as not pre-specified. Any deviation from the analyses described in this statistical analysis plan and a justification for making the change, will be described in the clinical study report.

8. Validation Requirements

Level I validation is required for Statistical and SAS programming of primary and secondary endpoints. Level I requires that the peer reviewer independently programs output and then compares the output with that generated by the original Statistical Programmer.

For the other endpoints, baseline characteristics and other summary outputs, minimally a Level II validation will be used. Level II requires that the peer reviewer reviews the code; where appropriate, performs manual calculations or simple programming checks to verify the output.

9. References

Chen, Y., DeMets, D., & Lan, K. (2004). Increasing the sample size when the unblinded interim result is promising. *Statistics in Medicine*, 1023-1038.

Danne, 2014. The PILGRIM Study: In Silico Modeling of a Predictive Low Glucose Management System and Feasibility in Youth with Type 1 Diabetes During Exercise. *Diabetes Technol Ther.*

Danne, 2017. International Consensus on Use of Continuous Glucose Monitoring. *Diabetes Care.*

ICHOM, 2019. *Type 1 and Type 2 Diabetes in Adults Data Collection Reference Guide,* s.l.: s.n.

Li, G., Shin, w., Xie, T., & Lu, J. (2002). A sample size adjustment procedure for clinical trials based on conditional power. Biostatistics, 277-287.

Mehta, C., & Pocock, S. (2011). Adaptive increase in sample size when interim results are promising: A practical guide with examples. Statistics in Medicine, 3267-3284.

White, I., Thompson, S. (2005). Adjusting for partially missing baseline measurements in randomized trials. Statistics in Medicine, 24: 993-1007.