Medtronic	
Study Title	A Pivotal Study to Evaluate Guardian™ Connect CGM System Performance in China
NCT Number	NCT03710083
Document Description	Clinical Biostatistical Plan (Version 1.0)
Document Date	14-DEC-2019

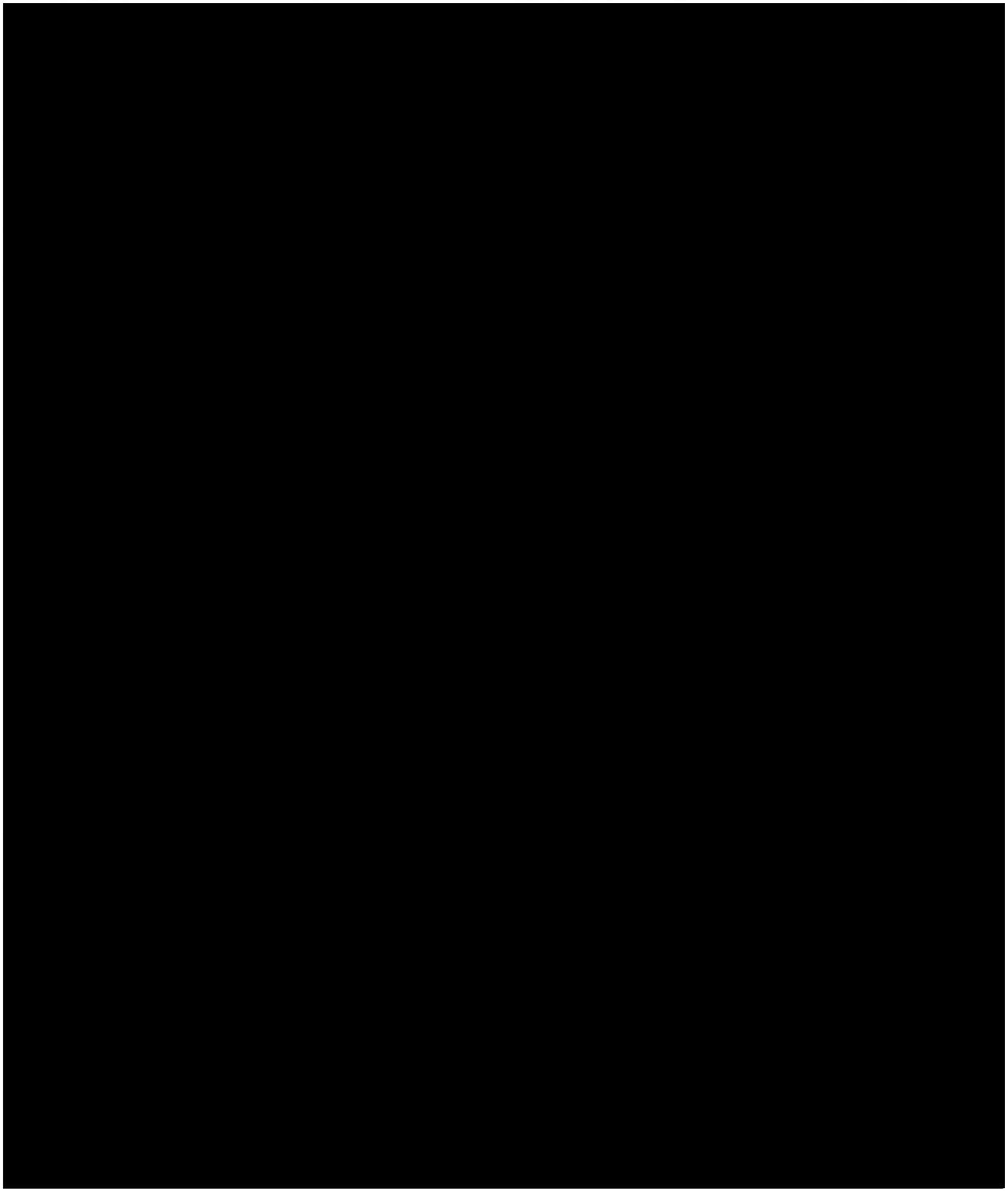
A Pivotal Study to Evaluate Guardian™ Connect CGM System Performance in China

Statistical Analysis Plan

V1.0

Statistical Analysis:	Medical Research & Biometrics Center, National Center for Cardiovascular Disease, CHINA
Sponsor:	Medtronic (Shanghai) Management Co., Ltd.
Issue Date:	14-December-2019

国家心血管病中心 医学统计部



Contents

1. Version History

Version	Summary of Changes	Author(s)/Title
V1.0	Not applicable	[redacted]

2. List of Abbreviations and Definitions of Terms

Term	Definition
A1C	Glycosylated hemoglobin
AE	Adverse Event
ARD	Absolute Relative Difference
BG	Blood Glucose
BMI	Body Mass Index
CEC	Clinical Events Committee
CFDA	China Food and Drug Administration
CGM	Continuous Glucose Monitoring
CGMS	Continuous Glucose Monitoring System
CIP	Clinical Investigation Plan
CTA	Clinical Trial Approval
DKA	Diabetic Ketoacidosis
EC	Ethics Committee
eCRF	Electronic Case Report Form
EGA	Error Grid Analysis
EOS	End of Study
FST	Frequent Sample Testing
Hct	Hematocrit
ICF	Informed Consent Form
ISIG	Interstitial Signal
IV	Intravenous
MARD	Mean Absolute Relative Difference
NMPA	National Medical Products Administration

Term	Definition
OC-RDC	Oracle Clinical Remote Data Capture
QC	Quality Control
SAE	Serious Adverse Event
SGV	Sensor Glucose Value
SID	Subject Identification
SMBG	Self-Monitoring of Blood Glucose
SQ	Subcutaneous
UADE	Unanticipated Adverse Device Effect
USB	Universal Serial Bus
YSI™*	Yellow Springs Instrument

CONTOUR™* is a registered trademark of Ascensia Diabetes Care.

YSI™* is a trademark of Xylem Inc. or one of its subsidiaries.

3. Introduction

Medtronic Continuous Glucose Monitoring Systems (CGMS) measure subcutaneous (SQ) glucose continuously over various ranges of time. Medtronic's newest generation SQ glucose sensor (the Guardian™ Sensor (3)) glucose sensor was approved by the United States Food and Drug Administration (FDA) in September 2016 as part of the MiniMed™ 670G System. The Guardian Sensor (3) is designed to be used with compatible Medtronic monitors, transmitters, or recorders to help users manage their diabetes.

The Guardian™ Sensor (3) glucose sensor design is similar to the Sof-sensor™ glucose sensor which is approved by National Medical Products Administration (NMPA, formerly CFDA). The Guardian™ Sensor (3) glucose sensor connects to a glucose sensor transmitter (Guardian™ Connect transmitter). A personal mobile device running the Guardian™ Connect app is the monitoring device used in the Guardian™ Connect System. Although the Guardian™ Connect System has been previously evaluated

during clinical investigation performed in the United States, testing with Chinese subjects is necessary to support NMPA approval of the Guardian™ Connect System.

For purposes of this study, subjects will wear two Guardian™ Sensor (3) glucose sensors. Subjects will manage their diabetes independent of the Guardian™ Sensor (3) glucose sensor values. During this time, venous blood glucose (BG) concentrations will be measured periodically by a reference method; these values will be compared to sensor glucose values (SGVs) in order to determine sensor accuracy.

The study is designed to demonstrate the performance and safety of the Guardian™ Sensor (3) glucose sensor when inserted in the abdomen used in Chinese subjects age 14 – 75 years. The data gathered is intended to support over 170 hours (7 days) of use with the Guardian™ Connect System. In addition, the study will collect data that will be used to model the accuracy of the Guardian™ Sensor (3) when used with the MiniMed™ 600G/700G series insulin pump systems and the Envision system.

Accuracy will be calculated based on comparing calibrated glucose sensor values to a "gold standard" (Yellow Springs Instrument [YSI™*™] plasma glucose values) in subjects during YSI™* frequent sample testing (FST). YSI™* glucose analyzers have been recognized standards for the measurement of BG and will be utilized across the investigational centers for the tests.

4. Study Objectives

The primary objective of the study is to demonstrate the accuracy of Guardian™ Sensor (3) when used over a period of 7 days with the Guardian™ Connect System.

5. Investigation Plan

5.1 Study Design

This study is a multi-center, prospective single-arm design without controls. Up to 72 subjects will be enrolled in order to have approximately 60 subjects complete the study. Three investigational centers in China will be used during this study.

During the study, each subject will be randomly assigned to one day of the Yellow Springs Instrument (YSI™*) frequent sample testing (FST) (Day 1, 3-5, or 7). Subjects will wear two Guardian™ Sensor (3)s each connected to a Guardian™ Connect transmitter for approximately 7 days (one which will be paired to the Guardian Connect app and the other will function as a glucose recorder.).

The Guardian Sensor (3)s will be worn in the abdomen area and self-inserted by the subject on same side or opposite sides.

On the evening prior to FST, subjects will be asked to fast for approximately 12 hours and adjust their insulin and medications according to routine care (for example as they would do for fasting lipid panel). Subjects may fast for shorter period of time **based on investigator discretion.**

The subject should be in fasting status upon arrival at hospital to start FST process. **The feeding protocol may be modified based on investigator discretion.**

The duration of FST will be approximately 7 hours.

During the study, subjects will continue with their current diabetes regimen independent of the study devices. Subjects will be instructed by the investigational center that they are not to use the study devices (except for the CONTOUR™* study meter) for the management of their diabetes. The CONTOUR™* study meter may be used for treatment decisions and calibration of Guardian™ Sensor (3).

Calibration at Home: Calibration at home will be required to be 3-4 times spread throughout the day for the Guardian™ Connect app.

Calibration should also be performed if the Guardian™ Connect app prompts for calibration (i.e., Smart calibration feature). NOTE: Subject will be instructed to minimize the delay time between prompt and calibration.

Fingerstick Testing: Fingerstick testing will be recommended to be a minimum of 4 times spread throughout the day. Subjects should test prior to and following meals and at bedtime.

5.2 Primary Endpoint

1a) 20/20% consistency evaluation with the reference value which is expressed by the agreement rate.

Sensor values from primary sensor will be compared to YSI™* plasma glucose values during YSI FSTs to analyze the agreement rate. Consistency rate = A + B:

A = (the number of pairs of the relative deviation of sensor value from primary sensor minus YSI™* value is within ± 20%/total number of pairs) * 100% (when the blood glucose concentration is greater than 4.4 mmol/L (80 mg/dL), percentage of results with relative deviations from the reference value within ±20%);

B = the number of pairs of the deviation of sensor value from primary sensor minus YSI™* value is within ± 20 mg/dL /total number of pairs) * 100% (when the blood glucose concentration is less than or equal to 4.4 mmol/L (80 mg / dL), percentage of results with deviations from the reference value within ±1.1 mmol/L (20mg/dL).

1b) Sensor values from primary sensor will be compared to YSI™* plasma glucose values during YSI™* FSTs. A mean rate in Zone A+B of Consensus Error Grid between Guardian Sensor (3) values and YSI™* plasma glucose values during YSI™* FST days

1c) Sensor values from primary sensor will be compared to YSI™* plasma glucose values during YSI™* FSTs. A mean rate in Zone A+B of Clarke Error Grid between Guardian Sensor (3) values and YSI™* plasma glucose values during YSI™* FST days

1d) Sensor values from primary sensor will be compared to YSI™* plasma glucose values during YSI™* FSTs. A mean absolute relative difference (MARD) between Guardian Sensor (3) values and YSI™* plasma glucose values during YSI™* FST days

The MARD is calculated by the following formula:

$$ARD_k = 100\% \frac{\left| y_{CGM}(t_k) - y_{ref}(t_k) \right|}{y_{ref}(t_k)}$$

$$MARD = \frac{1}{N} \sum_{k=1}^{N} ARD_k$$

5.3 Inclusion and Exclusion Criteria

Inclusion Criteria:

- Subject is 14 - 75 years of age at time of screening
- Subject has a clinical diagnosis of type 1 or 2 diabetes as determined via medical record or source documentation by an individual qualified to make a medical diagnosis
- Subject has adequate venous access as assessed by investigator or appropriate staff
- Subject is willing to follow the study procedures and willing to come to study visits.

Exclusion Criteria:

1. Subject will not tolerate tape adhesive in the area of Guardian™ Sensor (3) placement as assessed by qualified individual.
2. Subject has any unresolved adverse skin condition in the area of study device or device placement (e.g., psoriasis, rash, Staphylococcus infection)
3. Subject is actively participating in an investigational study (drug or device) wherein they have received treatment from an investigational study (drug or device) in the last 2 weeks
4. Subject is female and has a positive pregnancy screening test
5. Females of child bearing age and who are sexually active should be excluded if they are not using a

form of contraception deemed reliable by investigator

6. Subject is female and plans to become pregnant during the course of the study
7. Subject has a hematocrit (Hct) lower than the normal reference range
8. Subject may not be on the research staff of those performing this study

5.4 Study Timeline

The subject's participation from study start to completion is approximately 1 – 3 weeks (including replacement sensor wear and repeat in-clinic procedures).

Additional rescheduled visits could occur if Guardian™ Sensor (3)s dislodge and new Guardian™ Sensor (3)s must be re-inserted (See Replacement Sensors, Section 9.6)

- Visit 1: Consent and Screening
- Visit 2: Random Assignment
    - Visit 1 and 2 can be combined, however all eligibility criteria on Visit 1 should be met, including review of Hct prior to Visit 2.
    - Visit 1 and Visit 2 should be no more than 2 weeks apart.
- Visit 3: Study & Device Training
    - Two Guardian™ Sensor (3)s each connected to a Guardian™ Connect transmitter
    - Dispense Patient Log Sheet if CONTOUR™* study meter BG software and cable is not accessible at time of the study start
- Visit 4: YSI™* FST
    - Study and device training (Day 1 subjects who have not already received study and device training from Visit 3)
    - Subjects will undergo one YSI™* FST on any one of the following sensor wear days:
        - Day 1
        - Day 3 – 5
        - Day 7
- Visit 5: End of Study (EOS) Visit
    - Guardian™ Connect transmitters removed **(Please note that the subject should target**

wearing the device for 170 hours or longer from time of insertion)

- Investigational center staff will:
  - Download the Guardian™ Connect transmitters data from GST Download Utility Software following completion of the study devices wear
  - Download subject's CONTOUR™* study meter BG values
  - Collect subject's Patient Log Sheet if CONTOUR™* study meter BG software and cable is not accessible at time of the study start
  - Provide CONTOUR™* study meter and Guardian™ Connect transmitter data to sponsor
- Return study devices
- Subject complete questionnaires

6. General Statistical Considerations

6.1 Study Hypothesis

[redacted] **Statistical testing**

At 1a), 1b), and 1c) evaluation criteria, one proportion Z test will be used for the analysis of the primary endpoint. The 97.5% lower confidence limit of the agreement rate, the mean rate in Zone A+B of Consensus Error Grid, and the mean rate in Zone A+B of Clarke Error Grid will be tested against corresponding threshold, respectively.

At 1d) evaluation criteria, one sample T test will be used and the 97.5% upper confidence limit of the MARD will be tested against corresponding threshold.

The study will be considered as a success only when the four evaluation criteria (1a), 1b), 1c), and 1d)) meet the pre-defined thresholds.

6.2 Sample Size Determination

The sample size selected is based on the primary effectiveness endpoints, which are 1a) a within 20% agreement rate, 1b) a mean rate in Zone A+B of Consensus Error Grid , 1c) a mean rate in Zone A+B of Clarke Error Grid, and 1d) a MARD.

The primary endpoints of 1a) a within 20% agreement rate, 1b) a mean rate in Zone A+B of Consensus Error Grid, and 1c) a mean rate in Zone A+B of Clarke Error Grid are event rates, the estimated formula for sample size is as below:

$$n=\frac{\left[\mu_{1-\alpha}\sqrt{p_0(1-p_0)}+\mu_{1-\beta}\sqrt{p_T(1-p_T)}\right]^2}{(p_T-p_0)^2}$$

p_T is the efficacy level of the test group, p_0 is the target value level, μ indicates the quantile to which the standard normal distribution corresponds. When α level is 0.025 (one sided), $\mu_{1-\alpha}$ is 1.96; when β is 0.2, $\mu_{1-\beta}$ is 0.845.

Based on previous studies, assuming that for 1a) a within 20% agreement rate p_T is 0.65, p_0 is 0.6, 1,200 measuring points will provide power greater than 80% to demonstrate that 1a) is great than 0.6. Assuming each subject provides the results of at least 20 measuring points during YSI FST, including 60 subjects will meet sample size requirements. Considering 20% drop out, 72 subjects need to be enrolled.

Based on previous studies, assuming that for 1b) a mean rate in Zone A+B of Consensus Error Grid or 1c) a mean rate in Zone A+B of Clarke Error Grid p_T is 0.95, p_0 is 0.9, 1,200 measuring points will provide power greater than 80% to demonstrate that 1b) or 1c) is great than 0.9. Assuming each subject provides the results of at least 20 measuring points during YSI FST, including 60 subjects will meet sample size requirements. Considering 20% drop out, 72 subjects need to be enrolled.

The primary endpoint of 1d) a MARD is a continuous variable, the estimated formula for sample size is as below:

$$n=\frac{(\mu_{1-\alpha}+\mu_{1-\beta})^2\sigma^2}{(x_T-x_0)^2}$$

x_T is the MARD of the test group, x_0 is the MARD target value level, μ indicates the quantile to which the standard normal distribution corresponds. When α level is 0.025 (one sided), $\mu_{1-\alpha}$ is 1.96; when β is 0.2, $\mu_{1-\beta}$ is 0.845.

Bases on previous studies, assuming that for 1d) a MARD x_T is 18, x_0 is 20, σ is 5, 60 subjects will provide power greater than or equal to 80% to demonstrate that MARD value at subject level is less than 20%. Considering 20% drop out, 72 subjects need to be enrolled.

In conclusion, to meet the sample size requirements of four 1a), 1b), 1c), and 1d), it is determined that a total of 72 subjects are included.

6.3 Analysis Population

Full analysis set (FAS$^{\#1}$): a set of subjects who have at least one Guardian Sensor inserted.

Full analysis set (FAS$^{\#2}$): a set of subjects who have at least one Guardian Sensor inserted and have at least one paired Guardian Sensor and YSI^{TM*} measurement.

Per-Protocol Set (PPS) refer to the subgroup population of FAS$^{\#2}$ who have no major protocol deviations (subjects that violate the inclusion criteria, meet the exclusion criteria, etc.).

Safety Set (SS) refer to all enrolled subjects.

Primary efficacy analysis will be analyzed based on both the FAS$^{\#2}$ and PPS. All the baseline demographic data will be analyzed based on the FAS$^{\#1}$. Safety evaluation will be analyzed based on SS.

6.4 Analysis Population Set Determination Details

(1) No Guardian Sensor inserted: Subjects obtain randomly assigned FST, but have no Guardian Sensor inserted.

(2) No paired Guardian Sensor and YSI ^{TM*} measurement: Subjects obtain randomly assigned FST and have a Guardian Sensor inserted, but have no paired Guardian Sensor and YSI measurement.

(3) Violation of Inclusion/Exclusion criteria: Subjects do not meet the inclusion criteria or meet the exclusion criteria pre-specified in the protocol, and this protocol deviation may severely affect the results of primary effective endpoint. Whether protocol deviation severely affect the results of primary effective endpoints will be judged together by sponsor, investigators and statistician through discussion;

(4) SS= Number of subjects enrolled;

FAS$^{\#1}$= Number of the subjects randomly assigned FST- Number of subjects who have no Guardian Sensor inserted;

FAS$^{\#2}$= Number of FAS$^{\#1}$ - Number of subjects who have no paired Guardian Sensor and YSI ^{TM*} measurement;

PPS= Number of FAS$^{\#2}$ - Number of subjects who have major protocol deviations (i.e. violate inclusion and exclusion criteria, etc.).

The listing of subjects who have no Guardian Sensor inserted or have major protocol deviations will be listed, including the center number, subject ID, randomly assigned FST, gender, age, type, details, FAS$^{\#1}$, FAS$^{\#2}$ and PPS.

6.5 Missing Values, Abnormal Value and Outliers

Data entry error or non-reasonable values will be cleaned before data analysis. Imputation of missing data will be only applied for sensitivity analysis. And strategies are as follows:

The missing paired YSI-SG is caused from the following cases but not limited to:

- Missing FST-YSI
- FST-YSI with deletion flag (from protocol deviations collected by CRF)
- Existing FST-YSI value without paired SG point

Three imputation methods will be used for sensitivity analysis are as follows:

(1) Worst Case

1) 20/20% agreement rate: The missing paired YSI-SG will be filled with a 20/20% agreement of "No".

2) Zone of Clarke and Consensus Error Grid: The missing paired YSI-SG will be filled with a "Zone E".

3) ARD (%): The missing paired YSI-SG will be filled with a maximum value from the study data.

(2) Best Case

1) 20/20% agreement rate: The missing paired YSI-SG will be filled with a 20/20% agreement of "Yes".

2) Zone of Clarke and Consensus Error Grid: The missing paired YSI-SG will be filled with a "Zone A".

3) ARD (%): The missing paired YSI-SG will be filled with a minimum value from the study data.

(3) Randomly Pick from the Same Subject

1) The missing paired YSI-SG will be filled with a 20/20% agreement of random selection from the same subject.

2) The missing paired YSI-SG will be filled with a Zone of Clarke and Consensus Error Grid of random selection from the same subject.

3) The missing paired YSI-SG will be filled with an ARD (%) of random selection from the same subject.

For the dates of adverse events and medications collected during the trial, if the day or month or year is filled with "UK" or "NA" , they will not be processed.

6.6 General Considerations for Data Analysis

6.6.1 Data Collection

Analysis data source: The analysis of primary endpoints, other descriptive endpoints and simulation alert calculation will be based on both Oracle Remote Data Capture (RDC) database and external device dataset (collected sensor glucose data). The detailed specification of variables from external device data are as following:

- SG: Raw Sensor Glucose Reading
- PSGV: Predicted Sensor Glucose Reading
- SBG: Sensor Blood Glucose (values below 40 display as 40, values above 400 display as 400) – elaborated below
- ROC: Rate of Change of SG
- ISIG: Interstitial Signal
- FORCALBG: Used in Calibration

The value "511" in the SG and PSGV column indicates "null". The SBG column will display saturated values of 40 if the reading was displayed as "Low" (i.e. if SG < 40 mg/dL) or 400 if the reading was "High" (i.e. if SG > 400 mg/dL) while the SG column will display 511. SG value of 40 is indicated when SG and SBG columns are equal to 40, and same for 400. SG column is used to generate the numerical calculation for sensor performance. SBG is used for alert analysis and sensor life analysis. PSGV and ROC columns are used for alert analysis.

In addition, primary endpoints, other descriptive endpoints, and functional sensor life time will be based on real-time data collected from primary sensor. Precision analysis will be based on data collected from both sensors.

6.6.2 Pairing Scheme

1) All YSI™* collected will be presented. Descriptive analyses will include all YSI™* values, such as those less than (<) 40 mg/dL, or greater than (>) 400 mg/dL as long as Guardian™ Sensor (3) values are greater than or equal to (≥) 40 mg/dL or less than or equal to (≤) 400 mg/dL.

2) Reference glucose values (YSI™* values) will be paired with the closest Guardian™ Sensor (3) value between [0, 5) minutes.

3) The primary Guardian™ Sensor (3) will be used as a reference for precision analysis and the pairing scheme will be [0, 5) minutes.

6.6.3 YSI™* Retention

All YSI™* values will be captured and retained in OC-RDC database. However, if the absolute relative difference between Result A (black) and Result B (white) is greater than (>) 5%, the YSI™* values will not be included in the analysis dataset.

(1) At first, delete observations which were identified as deletion according to "deletion flag" collected in OC-RDC database;

(2) Calculate absolute relative difference between Result A (black) and Result B (white) by the formula:| Result A (black) - Result B (white)|/ Result A (black) *100%. If the ARD% is greater than 5%, then this observation will be deleted.

(3) The mean of Result A (black) and Result B (white) will be used in paired data analysis finally.

6.6.4 Sensor Data Retention

All collected sensor glucose data from the study will be analyzed. It includes, but is not limited to, sensor glucose data reprocessed with additional MeterBGs collected by patients.

6.7 Center (Site) Pooling

Data will be pooled for analysis.

6.8 Adjustments for Multiple Comparisons

No adjustments will be made.

6.9 Significance Level and Statistical Analysis Software

For the statistical analysis of the primary endpoint, the level of significance will be set at one-sided 0.025. The level of significance of tests performed on other indicators will be set at two-sided 0.05 level (expect special cases). SAS® 9.4 statistical software and R3.5.2 are applied for statistical analysis.

7. Indicators and Statistical Analysis Methods

7.1 Subject Disposition

The number of subjects enrolled in the study will be presented. The reasons for discontinuing prior to study completion will be summarized.

(1) Subjects Enrolled: Subjects who were enrolled in study.

(2) Screening Failure: The subject did not meet eligibility criteria on visit 1 and did not enter visit 2: random assignment. (Does the subject meet all inclusion criteria, no exclusion criteria, and randomization conditions collected in CRF was "No").

(3) Subjects Randomly Assigned FST: The subject entered visit 2: random assignment. (Does the subject meet all inclusion criteria, no exclusion criteria, and randomization conditions collected in CRF was "Yes").

(4) Subjects entering FST: The subject inserted Guardian Sensor successfully and entered visit 4: FST with YSI™*. (This item is determined based on the subjects' YSI™* information).

(5) Subjects completing the Study: Subjects completing Study: The subject completed the last visit. (This item is determined by the section study exit collected in CRF).

(6) Subjects discontinuing from the study: The subject did not complete the last visit. (This item is determined by the section study exit collected in CRF).

7.2 Subject Demographics and Baseline Characteristics

Subject demographics and characteristics, including age, gender, race, ethnicity, medical diagnosis, height, weight, BMI, CGM experience, diabetes type, pregnancy test result, baseline Hct and baseline A1C will be summarized by descriptive statistics (mean, standard deviation, minimum, median, and maximum) for continuous variables and by counts and percentages for categorical variables.

Age = (Informed Consent Date - Date of Birth) / 365.25;

BMI = Weight (kg) / (Height (m)) ^{2}.

7.3 Sensor Disposition

The number of Guardian™ Sensor (3) insertions and Guardian™ Sensor (3) removals for every subject enrolled in the study will be presented.

A descriptive analysis of Guardian™ Sensor (3) disposition including Guardian™ Sensor (3) dislodgement and reasons why it dislodged will be included in the Final Report. Guardian™ Sensor (3) insertion and removals will be characterized by the following:

- Duration of Guardian™ Sensor (3) wear by investigational center subject report
- The number and percentage of Guardian™ Sensor (3)s remaining in place at study end
- Reason for removal: for example, scheduled removal, companion sensor removal, early removal

The functional life of the primary Guardian™ Sensor (3) will also be characterized. The duration of Guardian™ Sensor (3) performance from the time of first valid Interstitial Signal (ISIG) to the time of

last available sensor value will be summarized by descriptive statistics (mean, standard deviation, minimum, median, and maximum) and described with Kaplan-Meier curves as well.

For the duration of Guardian™ Sensor analysis, both the physical sensor life time and functional sensor life time based on sensor level will be provided. The corresponding calculation methods are as following:

(1) Physical sensor life time (hours) = (Removal time- Insertion time) /3600, the insertion time and removal time collected in OC-RDC database.

(2) Functional sensor life time (hours): The analysis data collected in Guardian™ Sensor external database. Related calculation method is as following:

1) Find the first valid Interstitial Signal (ISIG) value (i.e. first ISIG after connection indicator), and its corresponding time will be considered as beginning time.

2) Find the last available Sensor value, and its corresponding time will be considered as ending time.

3) Functional sensor life time (hours) = (Ending time- Beginning time) /3600.

7.4 Treatment Characteristics

The total number of YSI™* collected and deleted will be calculated. In addition, the detailed reason for YSI™* deletion will also be provided.

7.5 Interim Analyses

Not Applicable.

7.6 Primary Endpoints

Detailed analyses for four primary endpoints are as following:

1a) For the 20/20% agreement rate analysis, the analyses based on measurement level will be provided. Both the approximate normal distribution method and Clopper-Pearson method will be used to estimate primary endpoint based on measurement level. The 97.5% lower confidence limit of the agreement rate will be tested against corresponding threshold.

1b) Sensor values from primary sensor will be compared to YSI™* plasma glucose values during YSI™* FSTs. A mean rate in Zone A+B of Consensus Error Grid between Guardian Sensor (3) values and YSI™* plasma glucose values during YSI™* FST days.

1c) Sensor values from primary sensor will be compared to YSI™* plasma glucose values during YSI™* FSTs. A mean rate in Zone A+B of Clarke Error Grid between Guardian Sensor (3) values and

YSI™* plasma glucose values during YSI™* FST days.

1d) Sensor values from primary sensor will be compared to YSI™* plasma glucose values during YSI™* FSTs. A mean absolute relative difference (MARD) between Guardian Sensor (3) values and YSI™* plasma glucose values during YSI™* FST days.

At 1b) and 1c) evaluation criteria, both the approximate normal distribution method and Clopper-Pearson method will be used to analysis primary endpoint based on measurement level. The 97.5% lower confidence limit of the mean rate in Zone A+B of Consensus Error Grid, and the mean rate in Zone A+B of Clarke Error Grid will be tested against corresponding threshold, respectively.

At 1d) evaluation criteria, one sample T test will be used based on subject level and the 97.5% upper confidence limit of the MARD will be tested against corresponding threshold.

The study will be considered as a success only when the four evaluation criteria (1a), 1b), 1c), and 1d)) meet the pre-defined thresholds.

The analysis of primary endpoint will be performed on both the FAS[#2] and PPS; and will be based on the available paired data only according to pairing scheme after considering YSI™* retention and sensor data retention principle.

7.7 Sensitivity analyses of primary endpoints

Sensitivity analyses will be performed on the FAS[#2].

(1) Missing data

To account for the effect of both missing data (e.g. missing FST-YSI™*, existing FST-YSI™* value without paired SG point, etc.), and FST-YSI™* with deletion flag (from protocol deviations collected by CRF) for primary endpoints evaluation, relevant sensitivity analyses for primary endpoints will also be provided.

In these analyses, the missing paired YSI-SG will be handled based on worst case, best case and randomly pick from the same subject methods, respectively. (See section 6.5). The 20/20% agreement rate, a mean rate in Zone A+B of Clarke and Consensus Error Grid based on measurement level will be analyzed after missing data being handled according to above different methods in sensitivity analysis.

(2) Site Effect

Investigative site effect will be evaluated using the generalized estimation equation (GEE) for 20/20% agreement rate, a mean rate in Zone A+B of Clarke and Consensus Error Grid and the linear regression

model for MARD. If site effect is detected, sensitivity analysis will be performed using the random-effects models to obtain combined effect for endpoints. If the random-effects model is not feasible then the estimation of fixed-effects model will be provided. Inverse variance method (weight=1/variance of site estimate) will be used in the analysis.

7.8 Other Descriptive Endpoints

7.7.1 Clarke Error Grid Analysis (EGA) of Paired Sensor and YSI™* and Reference Values

1) Description

Clarke EGA separates paired observations into five zones of clinical significance. The presence and severity of possible treatment error based on interstitial glucose assay evaluated by the sensor defines the five zones. Zone A represents the absence of treatment error, where the evaluation method and the reference method are within 20% of one another or in which both methods indicate hypoglycemia. Zone B represents cases where the two methods disagree by more than 20%, but do not lead to treatment error. Zones C, D, and E represent increasingly large and potentially harmful discrepancies between the evaluation and the reference method. If the method under evaluation has a high percentage (greater than (>) 90%) of its pairs in Zones A and B, then it is considered clinically acceptable [Clarke et al, 1987].

2) Statistical analysis

Summary statistics (N, %) for each of the zones, as well as combined Zones A and B, will be calculated.

In order to evaluate differing levels of accuracy at various YSI™* defined glucose levels, the number and percentage of paired observations falling into Zones A, B, A+B, C, D, and E will be provide form YSI™* glucose ranges of ≤80 mg/dL, greater than (>) 80-120 mg/dL, greater than (>)120-240 mg/dL, and greater than (>) 240 mg/dL.

All analysis performed using the Clarke Error Grid comparing the paired sensor and YSI™* reference glucose values will be duplicated using the Consensus Error Grid [Parkes et al, 2000].

7.7.2 Precision Analysis

Precision analysis will be performed for the two sensors worn by the same subject. This part of indicators will be summarized by descriptive statistics. The analysis method will be referred to section 7.2.

7.7.3 Other Accuracy Analyses

The ARD, the absolute relative differences between the sensor and YSI™* relative to the YSI™* reference will be calculated by FST days separately. Summary statistics will include its mean, standard deviation, min, median, and max.

The mean numerical bias, which is the difference between the sensor and YSI™* values, will be calculated by FST days. Summary statistics will include its mean, standard deviation, min, median, and max.

20% agreement rate will be summarized descriptively by FST days (Day 1, 3-5 and 7). A mean rate in Zone A+B of Clarke and Consensus Error Grid, a mean absolute relative difference (MARD) will also be summarized descriptively by FST days (Day 1, 3-5 and 7).

Above indicators will be summarized by descriptive statistics. The analysis method will be referred to section 7.2.

In addition, 20/20% agreement rate, a mean rate in Zone A+B of Clarke and Consensus Error Grid, a mean absolute relative difference (MARD) will be summarized descriptively by age subgroup (14-17 years old and 18-75 years old). The descriptive statistics and corresponding 95% confidence interval will be provided.

7.9 Safety Evaluation

For Safety analysis, no formal hypothesis testing will be performed. Descriptive analytics will be used to summarize safety events. Safety events which will be characterized include:

1) Skin assessment of Guardian™ Sensor (3) insertion sites
2) All adverse events (AEs) to include but not limited to:
    a) Device Related AE
    b) Procedure Related AE
    c) Serious Adverse Event (SAE)
    d) Serious Adverse Device Effects (SADE)
    e) UADE
    f) Severe Hypoglycemia
    g) Diabetic Ketoacidosis (DKA)

7.10 Device Deficiencies

Descriptive summary will be used to characterize device deficiencies.

7.11 Subject Feedback

Descriptive summary will be used to characterize study questionnaire results. The questionnaire will use a Likert scale rating to assess their Guardian™ Sensor (3) experience.

7.12 Simulation Alert Calculation

The simulation alert calculation for the device 640G Pump and GS3 (Guardian™ Sensor 3) APP will be provided based on data collected in OC-RDC database and external device database.

(1) Dataset preparation: The device dataset is prepared for alert calculation. The valid data should be after first valid ISIG. Any data before first valid ISIG should be excluded from analysis.

(2) Variables used for calculation: SBG, PSGV, ROC, and YSI™*/MBG where applicable.

1) **SBG (Saturated Sensor Blood Glucose)** is used for **threshold alert analysis**. Note: SBG is a saturated SG when SG <40, it sets to 40 or 400 when SG>400.

2) **PSGV (Predictive Sensor Glucose Reading)** is used for **hypo predictive alert analysis** of 640G Pump

3) **ROC (Rate of Change of SG)** is used for **hyper predictive alert analysis** of 640G Pump and **for both hypo and hyper predictive alert analyses** of GS3 APP. ROC-based predictive SG will be calculated with the formula: PSGV=SBG+ROC*30.

(3) The CGM Setting: For hypo range, the CGM alert settings will be set as <= 50/60/70/80/90/100 mg/dL, respectively. For hyper range, the CGM alert settings will be set as >=300/250/220/180 mg/dL, respectively. Reference value YSI™* and SBG/PSGV threshold will keep changing as alert setting changes.

(4) Calculation indicators and methods: Detection rate, missed detection rate, false alert rate and true alert rate will be calculated. For these four indicators, three types of simulation alert/event rates will be calculated. These three types are threshold, predictive and combined threshold and predictive, respectively. The detailed definition of above four indicators are below:

1) Detection Rate (% of Hypo/Hyper Events Correctly Detected): The device alarmed at the specified CGM settings within 30 minutes (or 15 minutes) before or after the reference blood glucose

value (YSI™* or MBG) goes at/below (for hypo) or at/above (for hyper) the specified CGM setting levels.

2) Missed Detection Rate (% of Hypo/Hyper Events Not Detected): The device did not alarm at the specified CGM settings within 30 minutes (or 15 minutes) before or after the reference blood glucose value (YSI™* or MBG) goes at/below (for hypo) or at/above (for hyper) the specified CGM setting levels.

3) False Alert (% False Alerts): There is no reference blood glucose value (YSI or MBG) goes below (for hypo) or above (for hyper) the specified CGM setting levels within 30 minutes (or 15 minutes) before or after the sensor alarmed at the specified alert settings.

4) True Alert (% of Alerts Verified by Hypo/Hyper Events): There is at least one reference blood glucose value (YSI or MBG) goes below (for hypo) or above (for hyper) the specified CGM setting levels within 30 minutes (or 15 minutes) before or after the sensor alarmed at the specified alert settings.

7.13 Changes to Planned Analysis

1. Analysis result of Continuous Error Grid will not be provided in statistical analysis report as there is no requirement in registration technical review guideline of National Medical Products Administration (NMPA).

2. Analysis of the pooled investigative study center factor for all 4 endpoints will be performed by the generalized estimation equation or the linear regression model, because Gail-Simon test is not applicable for the study design.

8. Validation Requirements

Level I or Level II validation are required for analysis output. Level I requires that the peer reviewer independently programs output and then compares the output with that generated by the original Statistical Programmer. Level II requires that the peer reviewer reviews the code; where appropriate, performs manual calculations or simple programming checks to verify the output.

9. Mock Tables, Listings and Figures

See SAP appendix tables, listings and figures.