Sponsor Name:	Theracos Sub, LLC
Protocol Number and Title:	A Phase 3, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate Bexagliflozin in Subjects with Type 2 Diabetes Mellitus Who Are not Adequately Controlled by Metformin Alone
Protocol Version and Date:	Version 1.0, 26 July 2017
INC Research Project Code:	1007780
Author(s):	Lu Wang, Biostatistician II Fang Zhu, Principal Biostatistician
SAP Version:	Version 2.0
SAP Version Date:	15-Feb-2019

Revision History

Version #	Date (dd-mmm-yyyy)	Document Owner	Revision Summary
Version 0.1	13-Feb-2018	Lu Wang	Initial Release Version
Version 0.2	23-Feb-2018	Lu Wang	Update to Address Sponsor Comments
Version 1.0	28-Feb-2018	Lu Wang	Update to Address Sponsor Comments
Version 1.1	15-Jun-2018	Lu Wang	Update to Address FDA Comments
Version 1.2	06-Jul-2018	Lu Wang	Update to Address Sponsor Comments
Version 2.0	15-Feb-2019	Fang Zhu	Update per dry run comments: 1. Correct WHODrug and MedDRA versions 2. Use “high glycemic”, instead of “open-label”. Correct language to include high glycemic group, e.g. change first dose of double-blind study medication to first dose of study medication during treatment period. 3. Add ITT analysis set to high glycemic group and updated safety analysis set definition to keep both groups be consistent. 4. Added LOCF for HbA_{1c} high glycemic group 5. Correct section 12. LOCF will be conducted. Tipping point analysis is used to replace jump to reference MI.

I confirm that I have reviewed this document and agree with the content.

APPROVALS	
INC Research	
Lead Biostatistician Fang Zhu Principal Biostatistician	18-Feb-2019 **Date** (dd-Mmm-yyyy)
Senior Reviewing Biostatistician Mike Ou Director, Biostatistics	18-Feb-2019 **Date** (dd-Mmm-yyyy)
Theracos Sub, LLC	
Sponsor Contact Yuan-Di Halvorsen Program Manager for Theracos Sub, LLC	15-Feb-2019 **Date** (dd-Mmm-yyyy)

TABLE OF CONTENTS

1. GLOSSARY OF ABBREVIATIONS

Abbreviation	Description
AE	Adverse Event
ANCOVA	Analysis of Covariance
ATC	Anatomical Therapeutic Chemical
BMI	Body Mass index
BP	Blood Pressure
CEC	Cardiovascular Endpoint Committee
CI	Confidence Interval
CRF	Case Report Form
DKA	Diabetic Ketoacidosis
dL	Deciliter
DSMB	Data and Safety Monitoring Board
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
eGFR	Estimating Glomerular Filtration Rate
FPG	Fasting Plasma Glucose
GEE	Generalized Estimating Equation
GMI	Genital Mycotic Infection
HbA_{1c}	Hemoglobin A_{1c}
ICF	Informed Consent Form
ICH	International Conference on Harmonization
ITT	Intention to Treat
IWRS	Interactive Web Randomization System
LDL-C	Low Density Lipoprotein Cholesterol
LOCF	Last Observation Carried Forward
LS	Least Squares
MACE	Major Adverse Cardiovascular Event

Abbreviation	Description
MAR	Missing at Random
MCAR	Missing Completely at Random
MCMC	Monte Carlo Markov Chain
MedDRA	Medical Dictionary for Regulatory Activities
mg	Milligram
mL	Milliliter
MMRM	Mixed Model Repeated Measures
MNAR	Missing not at Random
N/A	Not Applicable
Na	Sodium
PK	Parmacokinetics
PP	Per Protocol
PR	Time from the beginning of the P wave to the beginning of the QRS complex in electrocardiogram
PT	Preferred Term
Q1	1st Quartile
Q3	3rd Quartile
QRS	Graphical deflections corresponding to ventricular depolarization in a typical electrocardiogram
QTc	Time between the start of the Q wave and the end of the T wave in the ECG, corrected for heart rate
RR	Time between the start of one R wave and the start of the next R wave in the ECG
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SBP	Systolic Blood Pressure
SD	Standard Deviation
SE	Standard Error
SI	Standard International System of Units
SMBG	Self-Monitored Blood Glucose

Abbreviation	Description
SOC	System Organ Class
SOP	Standard Operating Procedure
T2DM	Type 2 Diabetes Mellitus
TEAE	Treatment-emergent Adverse Event
TLF	Table, Listing And Figure
UACR	Urine Albumin To Creatinine Ratio
UTI	Urinary Tract Infection
WBC	White Blood Cell
WHO-DD	World Health Organization Drug Dictionary

2. PURPOSE

The purpose of this statistical analysis plan (SAP) is to ensure that the data listings, summary tables and figures which will be produced, and the statistical methodologies that will be used, are complete and appropriate to allow valid conclusions regarding the study objectives.

2.1. RESPONSIBILITIES

Theracos has designed the study protocol and is responsible for the conduct of the study. INC Research is responsible for the development and validation of a clinical database using MediData RAVE platform.

INC Research will perform the statistical analyses and are responsible for the production and quality control of all tables, figures and listings (TLFs).

Adverse events that have met the seriousness criteria defined in the protocol are reported on the serious adverse event (SAE) forms using the MediData RAVE platform. An SAE case consists of the information reported in the SAE forms, subject characteristics documented in the case report forms, and additional source data such as hospital discharge summaries. Each SAE case is recorded in a validated ARGUS database which is managed by Covance. Any discrepancies in critical data fields of each SAE will be reconciled between the ARGUS and THR-1442-C-419 clinical database prior to database lock. The SAE coding, analyses and summaries are based on the final study data recorded in the clinical database. Detailed serious adverse event follow-up data will be reported from the ARGUS database and are not included in this SAP.

Theracos will perform review of all tables, figures and listings before the finalization.

2.2. TIMINGS OF ANALYSES

The final analysis of safety and efficacy is planned after all subjects complete the planned 24 weeks of study treatment and the subsequent follow-up period or terminate early from the study.

3. STUDY OBJECTIVES

3.1. PRIMARY OBJECTIVE

To determine the placebo-corrected change in hemoglobin A1c (HbA_{1c}) from baseline to week 24 in adults with type 2 diabetes mellitus (T2DM) whose disease is inadequately controlled by metformin alone.

3.2. SECONDARY OBJECTIVES

The secondary efficacy objectives of this study are:

- To assess the effect of bexagliflozin on the change in fasting plasma glucose (FPG) as a function of time
- To assess the effect of bexagliflozin on the change in systolic blood pressure (SBP) as a function of time
- To assess the effect of bexagliflozin on the proportion of subjects achieving HbA_{1c} of ≤ 7% as a function of time
- To assess the change in total body weight as a function of time in subjects with baseline body mass index (BMI) ≥25 kg/m^2
- To assess the change in HbA_{1c} as a function of time
- To assess the change in HbA_{1c} as a function of time among subjects who have baseline HbA_{1c} of > 10.5% and ≤ 12.0%

3.3. SAFETY OBJECTIVES

- To determine the frequency and severity of treatment emergent adverse events
- To determine the frequency and severity of treatment emergent adverse events of interest
- To record and evaluate changes in concomitant medication use that may affect subject safety
- To evaluate any potentially adverse changes in laboratory test values
- To assess changes in cardiac rhythm through 12-lead electrocardiogram (ECG)
- To evaluate vital signs
- To assess general health by physical examination

3.4. OTHER OBJECTIVES

The other objective is:

- To measure bexagliflozin plasma concentrations as a function of time from dosing (sparsely sampled)

3.5. BRIEF DESCRIPTION

THR-1442-C-419 is a phase 3, multi-center study to evaluate the efficacy and safety of orally administered bexagliflozin tablets, 20 mg or placebo, in subjects whose T2DM is not adequately controlled by metformin as the only therapy in addition to diet and exercise. Approximately 350 subjects who meet all the inclusion criteria, none of the exclusion criteria, and who consent to participate in the study, are eligible for study enrollment. All subjects must have taken metformin at an optimal or near-optimal stable dose for ≥ 8 weeks prior to screening and have received diet and exercise counseling.

Approximately 300 subjects with HbA_{1c} values ≥ 7.5% and ≤ 10.5% at screening (the double-blind treatment group) who successfully complete a 1-week run-in and who remain eligible will be randomized in a 1:1 ratio to receive once daily double-blind treatment of bexagliflozin tablets, 20 mg or placebo. Study subjects will continue receiving open-labeled metformin background medication during the entire study at a stable dose and frequency. The treatment period will be 24 weeks. The study will be conducted in an outpatient setting. Study subjects will have clinic visits at weeks 6, 12, 18, and 24 for safety and efficacy evaluation. A final follow up visit will be conducted at week 26.

Up to 50 subjects who have HbA_{1c} > 10.5% and ≤ 12.0% at screening will be assigned to the high glycemic group to receive open labeled bexagliflozin tablets, 20 mg, in addition to metformin. There will be no placebo control cohort for the high glycemic group. Subjects assigned to the high glycemic group will be asked to follow the same study procedures as subjects assigned to the two study arms.

Subjects with persistent hyperglycemia based on blood glucose levels may receive approved antidiabetic rescue medications.

A sparse sampling assessment of bexagliflozin pharmacokinetics (PK) in the study population will also be performed. A net enrollment of approximately 200 subjects is planned for this aspect of the study. Samples will be taken between weeks 6 and 12. The population PK study design and analysis plan will be described in a designated study protocol and the analysis will be reported separately.

3.6. SUBJECT SELECTION

The study population will include approximately 350 subjects whose T2DM is inadequately controlled by metformin and who meet all of the inclusion criteria and none of the exclusion criteria. Clinical sites in the US and Japan are anticipated to

recruit subjects.

3.6.1. Inclusion Criteria

Refer to protocol section 4.2 for inclusion criteria.

3.6.2. Exclusion Criteria

Refer to protocol section 4.3 for exclusion criteria.

3.7. DETERMINATION OF SAMPLE SIZE

In the blinded segment of the study approximately 300 subjects will be randomized 1:1 to bexagliflozin tablets, 20 mg or bexagliflozin tablets, placebo. With this sample size, the study has approximately 90% power to detect a 0.4% difference in HbA_{1c} between the active arm and the placebo arm, assuming a standard deviation of 1% HbA_{1c} for the study population and that 12% of the study participants will have dropped out by week 24.

With similar assumptions, a single group of 50 subjects has approximately 90% power to detect a difference from baseline of 0.5% HbA_{1c}, assuming a 12% dropout rate and a population standard deviation of 1%.

3.8. TREATMENT ASSIGNMENT & BLINDING

3.8.1. Treatment Assignment

The study will be conducted at multiple investigative sites and will likely involve a variable numbers of subjects at each site. Enrollment will be on a competitive basis but each site will be capped at 30 randomized subjects. However, after a site has recruited 30 randomized subjects, if a potential subject at that site is in the run-in phase and wishes to continue with the study, the subject will be allowed to continue and, if eligible, to be randomized.

Eligible subjects who complete the run-in period, meet all study inclusion/exclusion requirements, and have $HbA_{1c} \geq 7.5\%$ and $\leq 10.5\%$ at screening will be randomized in a 1:1 ratio to receive bexagliflozin tablets, 20 mg or placebo during the 24-week treatment period. Subjects will be assigned to treatment groups in sequential order as they qualify for the study, using a centrally located and managed Interactive Web Response System (IWRS). Randomization will be stratified by high ($8.5\% \leq HbA_{1c} \leq 10.5\%$) and low ($7.5\% \leq HbA_{1c} < 8.5\%$) baseline measurement at screening (Visit V1) and by country (US or Japan). Eligible subjects who have $HbA_{1c} > 10.5\%$ and $\leq 12.0\%$ at screening will be assigned to the open-label high glycemic group.

3.8.2. Blinding

Approximately 300 subjects will participate in the double-blind placebo-controlled portion of the trial. The sponsor, investigators, study coordinators, pharmacists, study subjects, the cardiovascular endpoint committee (CEC) members or Diabetic Ketoacidosis (DKA) adjudicator will be blinded to the study medication of these participants. Upon randomization, each subject will receive a subject randomization number and a drug kit. To maintain blinding of the individual treatment assignments, central laboratory glucose urinalysis data will not be made available to any study personnel or subjects.

If knowledge of the test substance is needed to manage a subject's condition, the investigator will contact the IWRS to obtain the treatment assignment. If unblinding occurs for any reason, the time and reason for breaking the blind will be recorded in the case report form (CRF) and the sponsor must be notified within 24 hours.

A designated statistician who is not involved with the study operation will hold the treatment codes. The unblinded treatment information can be provided to the DSMB to facilitate the evaluation of any clinically important increase in the rate of a serious suspected adverse reaction or to the designated safety contact when the treatment information is required to determine if an expedited safety report must be submitted to regulatory agencies.

The treatment assignment will continue to be withheld from the CEC members or the DKA event adjudicator at the conclusion of the study until all global investigational studies are completed and a meta-analysis to assess cardiovascular risk is conducted.

Subjects in the high glycemic group (HbA_{1c} > 10.5% and ≤ 12.0% at screening) will receive open labeled study drug. There will be no placebo cohort for the high glycemic group.

3.9. ADMINISTRATION OF STUDY MEDICATION

The following investigational products will be used for oral administration:

- Bexagliflozin tablets, 20 mg: tablets containing 20 mg of bexagliflozin
- Bexagliflozin tablets, placebo: tablets containing no bexagliflozin

Bexagliflozin tablets, 20 mg or placebo, should be taken at approximately the same time each day, before or after breakfast, with one cup (250 mL) of water.

On the day of scheduled clinic visits at which blood is to be drawn, administration of bexagliflozin should be delayed until after blood is drawn and bexagliflozin tablet should be taken in the clinic or at home with one cup (250 mL) of water.

3.10. STUDY PROCEDURES AND FLOWCHART

The activities that must be performed at each clinic visit are presented in Table 1. The required laboratory tests scheduled at each visit are listed in Table 2.

A visit window of ± 3 days is allowed for all visits except Visit 3. Visit 3 is the day of randomization and the basis for the visit window.

The study staff will contact each subject prior to a scheduled clinic visit to confirm the time of the visit and to remind the subject of proper fasting practice. A subject must be queried to assess compliance with an approximately 8 hours fast prior to blood draw to ensure the FPG and triglycerides values can be accurately determined. If a subject has not fasted for approximately 8 hours, the subject must return as soon as can be arranged (within 1 week) to provide a specimen after proper fasting.

Table 1 Schedule of Events

Procedure	Screening	Run-in	Treatment					Follow-up
Visit Number	V1/ Screening	V2/ Run In	V3/ Random-ization	V4	V5	V6	V7	V8/ Follow Up
Time to Randomization (week)	-3	-1	0	6	12	18	24	26
Informed Consent	X							
Screening for I/E Criteria	X		X					
Medical History	X							
Diet and Exercise Counseling		X						
Physical Examination			X					X
Abbreviated Physical Examination	X						X	
Vital Signs	X		X	X	X	X	X	X
Electrocardiography			X				X	
Clinical Laboratory Tests	X			X	X	X	X	X
Urine Pregnancy Test (local, WOCBP only)	X (all women)		X	X	X	X	X	X
Diary and Glucometer Dispensation and Review		X						
Dispensing Run-in Drug		X						
Diary and Glucometer Record Review			X	X	X	X	X	X
Dispensing Investigational Product			X		X			
Adverse Events Assessments		X	X	X	X	X	X	X
Concomitant Medication Assessments		X	X	X	X	X	X	X

Table 2 Schedule of Clinical Laboratory Tests

	Screening	Run-in	Treatment					Follow-up
Visit number	V1/ Screening	V2/ Run In	V3/ Randomization	V4	V5	V6	V7	V8/ Follow Up/ ET
Time to Randomization Visit (weeks)	-3	-1	0	6	12	18	24	26
Hematology	X			X	X		X	X
Serum Chemistry and Electrolytes	X			X	X	X	X	X
Glycemic Control	X			X	X	X	X	X
Serum Lipids	X				X		X	X
Urinalysis	X			X	X	X	X	X
Urine Pregnancy Test (WOCBP)	X (all women)		X	X	X	X	X	X
Population PK sampling				X	X			

4. ENDPOINTS

4.1. PRIMARY EFFICACY ENDPOINT

The primary efficacy endpoint is the change of HbA_{1c} from baseline to week 24 in the 20 mg dose group compared to placebo group.

4.2. SECONDARY EFFICACY ENDPOINTS

The key secondary efficacy endpoints include:

- Change in FPG from baseline over time
- Change in SBP from baseline over time
- Proportion of subjects achieving $HbA_{1c} \leq 7\%$ over time
- Change in body weight in subjects with baseline BMI ≥ 25 kg/m^2 over time
- Change in HbA_{1c} over time
- Change in HbA_{1c} among subjects who have baseline $HbA_{1c} > 10.5\%$ and $\leq 12.0\%$ over time

4.3. SAFETY ENDPOINTS

- Adverse events (AEs)
- AE of special interest
- Medication use
- Clinical laboratory results
- ECG results
- Vital signs
- Physical exam results

5. ANALYSIS SETS

5.1. SCREENED ANALYSIS SET

The Screened Analysis Set will include all subjects screened for eligibility. The screened population will include screen failures. Unless specified otherwise, this population will be used for summaries of subject disposition.

5.2. SAFETY ANALYSIS SET

All subjects who took at least one dose of study medications will be included in the safety analysis set. Safety analyses will be based on the medications that were taken first. The safety analysis set is the primary analysis set for safety evaluation.

5.3. INTENTION TO TREAT ANALYSIS SET

All subjects in double-blind treatment group who are randomized regardless of treatment adherence or availability of follow up data, or all subjects in the high glycemic group who enrolled into the study will be included in the Intention to Treat (ITT) Analysis Set. All analyses of the ITT analysis set will be based on each subject's randomized assigned treatment. The ITT analysis set will serve as the primary set for the efficacy analyses.

5.4. PER PROTOCOL ANALYSIS SET

The Per Protocol (PP) Analysis Set will include all subjects in double-blind treatment group and in the ITT who meet the study eligibility requirements and have no major protocol deviations that affect the validity of the efficacy measurements. Protocol deviations that may result in subject or visit exclusion from the PP Analysis Set are described in Section 5.5. The PP analysis set will serve as the secondary set for efficacy assessment.

5.5. PROTOCOL DEVIATIONS

Protocol deviations will be captured during monitoring visits and recorded in CRF. Deviation will be classified into each of the 9 categories: enrollment criteria, non-compliance, laboratory, dosing, visit schedule, visit/procedure requirement, concomitant medication, informed consent, and other. Date and details of the deviation will be recorded. The list of deviation will be reviewed before unblinding and major protocol deviations that, in the opinion of the medical monitor, could affect the primary and secondary variables will be determined. Some possible types of major protocol deviations are showing in Table 3.

Table 3 Some Possible Types of Major Protocol Deviations

Category	Criteria	Exclusion
Inclusion/Exclusion Criteria		
Ineligible subject is enrolled	- Subjects not satisfying HbA_{1c} inclusion criteria - Treated with SGLT2 within 3 months of screening	Subject exclusion
Prior or Concomitant Medication Restrictions		
Use of another SGLT2 inhibitor	Use of an SGLT2 inhibitor as the rescue medication for hyperglycemia	Visit exclusion [exclude data post SGLT2 starts]
Use of new Diuretic medication	Initiate of new diuretic medication within 12 weeks of randomization	Visit exclusion [exclude data post initiation of a new diuretic]
Randomization/Blinding		
Unblinding	Blind was broken (requested in IWRS)	Visit exclusion [exclude data post blind broken]
Dosing Non-Compliance	-Subject missed more than 50% of the investigational product doses between week 12 and week 24	Subject exclusion

6. GENERAL ASPECTS FOR STATISTICAL ANALYSIS

6.1. GENERAL METHODS

Statistical methodology and analyses are in accordance with the principles outlined by the International Conference on Harmonization (ICH) E9 guidelines. All statistical analyses will be conducted using SAS statistical software version 9.4 or higher.

Tables, listings and figures (TLFs) will be produced in accordance with the principles outlined by the ICH E3 guidelines. Unless otherwise specified, for all TLFs where applicable, double-blind group and high glycemic group will be presented in the same TLFs as different group. Double-blind group will be presented by treatment (Bexagliflozin 20 mg; Placebo) and overall where applicable.

Data summaries will use descriptive statistics (number of subjects [n], mean, standard deviation [SD], 1st quartile [Q1], median, 3rd quartile [Q3], minimum and maximum) for continuous variables, and frequency and percentage of subjects for categorical and ordinal variables, unless otherwise specified.

Unless otherwise specified, for double-blind group all statistical tests will be one-sided using a 0.025 level of significance. All confidence intervals (CIs) will be two-sided 95% CIs. No formal hypothesis testing will be performed for the high glycemic group.

The analysis visit window will be assigned to data collection. One selected data point per visit will appear in summary tables and figures. Refer to section 6.4 for details. All visit assessment data will be included in shift tables and will appear in the subject listings.

No data imputation will be applied for missing values, unless otherwise specified.

6.2. KEY DEFINITIONS

6.2.1. Treatment Period

For double-blind treatment group, treatment period is defined as the period after subject is randomized to receive the double-blind study medication. For high glycemic group, treatment period is defined as the period when subject is dosed with active study medication.

6.2.2. Baseline Values

For subject who is dosed, baseline is defined as the last non-missing value on or prior to the first dose of study medication in treatment period. For subject who is not dosed in

double-blind treatment group, baseline is defined as the last non-missing value on or prior to the randomization date.

6.2.3. First Dose Date

Two "first dose dates" will be required – one for the Run-In period and one for the treatment period. The first dose date for the Run-In period will be the date of administration of the first dose of placebo tablets during the Run-In period. The first dose date for the treatment period will be the date that the first dose of the study medication in the treatment period. Both first dose dates will be obtained from the electronic case report form (eCRF). Study analyses will use the treatment period first dose date.

6.2.4. Study Day

Study Day is the number of days starting from the first administration of study medication in treatment period, which is counted as Study Day 1. If the assessment date is after the date of the first study medication in treatment period, the study day is calculated as date of assessment - date of the first dose administration (in treatment period) + 1. If the assessment date is prior to the date of the first study medication in treatment period, the study day is calculated as date of assessment - date of the first dose administration (in treatment period).

6.2.5. Duration

Duration of treatment period will be determined as Duration = last dose date (in treatment period) - first dose date (in treatment period) + 1. Duration of Run-In period will be determined as Duration = last dose date (in run-in period) - first dose date (in run-in period) + 1.

6.2.6. End of Study

The end of study is defined as the date of final contact as entered on the End of Study page of the eCRF.

6.2.7. Patient Years

Patient years are calculated as sum of the duration from the first dose of double-blind treatment or active treatment in high glycemic group to the end of study/ 365.25 of all subjects in the specified analysis set and treatment arm. This is used as the denominator of the computation of incidence rate.

6.3. MISSING DATA

The handling of missing or incomplete data is described for each endpoint and data type (as needed) in Section 7 to 9.

6.4. ANALYSIS VISIT WINDOWS

Table 4 shows how data will be mapped to analysis visits prior to selection of records for analysis. All post-baseline visits, including unscheduled and early termination visits, will be mapped. After mapping the data to the analysis visits, the following rules will apply unless other handling is specified for a particular analysis.

- If multiple records are available within a single analysis visit window, the record closest to the planned assessment day will be selected for analysis.
- If 2 records are equidistant from the target day, then the later record will be selected.
- If a subject has no record in an analysis window, the subject will be considered missing at that visit.

Table 4 Analysis Visit Windows

Study Day Window	Scheduled day	Scheduled Visit/Week
Day 2 - 63	Day 42	Visit 4/Week 6
Day 64 – 105	Day 84	Visit 5/Week 12
Day 106 - 147	Day 126	Visit 6/Week 18
>= Day 148	Day 168 or (Last dose date + 1), select the less one	Visit 7/Week 24

* Week 24 is the end of treatment visit for those subjects completing the study per protocol. For endpoints of lab, vital sign and ECG, the first collection after assigned Week 24 and > 7 days from Week 24 visit will be considered as Week 26 visit. All data summary will base on the mapped visit.

6.5. POOLING OF CENTERS

Subjects will not be pooled based on site size, but rather by country (US or Japan), to ensure a sufficient number of subjects per treatment arm in both ITT and PP populations for analysis that contain country as a model effect.

7. DEMOGRAPHIC, OTHER BASELINE CHARACTERISTICS AND MEDICATION

7.1. SUBJECT DISPOSITION AND WITHDRAWALS

Subject disposition data will be listed. A disposition table will be presented by group (double-blind group and high glycemic group), treatment and overall where applicable, the number and/or percentage of subjects who have signed the informed consent, who have been screened, screen failed prior to Run-in, screen failed during the Day -7 to Day -1 Run-in, and who have entered treatment period, completed the study drug, discontinued study drug, completed the study, and discontinued from the study after entering treatment period. The reasons for early withdrawal after entering treatment period will be summarized. A separate table will display the number of subjects who have met the eligibility criteria at screening and Run-in period but have withdrawn prior to randomization (non-randomized subjects) with a summary of the reasons for withdrawal prior to randomization.

Assignment to the analysis sets (safety, ITT, and PP where applicable) will be summarized.

7.2. SUBJECT ELIGIBILITY AND PROTOCOL DEVIATIONS

All subjects, including screen failure subjects who violate the Inclusion/exclusion criteria will be listed. Reason for screen failure will be summarized.

Deviations that could affect the primary and secondary variables will be considered when determining a subject's eligibility for the PP population. The number and percent of subjects who had any major deviation and each type of major protocol deviation will be tabulated for the ITT Analysis Set.

7.3. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Demographic characteristics include age, gender, race, ethnicity and country. Baseline characteristics will include HbA_{1c}, baseline HbA_{1c} group ($\geq 8.5\%$ and $\leq 10.5\%$ vs. $\geq 7.5\%$ and $< 8.5\%$) at Visit 1, FPG, blood pressures, body weight, BMI. Summary descriptive statistics by treatment will include counts and percentages for discrete variables and estimation of means, standard deviations, medians, inter-quartile range (Q1, Q3), minimum, and maximum for continuous variables. Subjects' baseline demographic and personal baseline characteristics will be summarized by group (double-blind group and high glycemic group) and treatment for subjects in the safety, ITT, and PP analysis set where applicable. Subject age will be calculated from the date of informed consented.

7.4. MEDICAL HISTORY

Significant medical and surgical history, including dates of diagnoses and procedures and whether the condition is ongoing, if applicable, will be collected at the screening visit. Each condition will be recorded as a verbatim term, date of onset, date resolved, and a checkbox that indicates ongoing conditions. Significant surgical and medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 20.1.

Medical and surgical history will be summarized for the Safety Analysis Set by group, treatment, system organ class (SOC), and MedDRA preferred term (PT), overall.

Subject diabetes and cardiovascular diseases history will be summarized for all categorical variables by frequency and percentage. Listing will be provided.

7.5. MEDICATION

All prescription and over-the-counter medications, including vitamins and herbal supplements, that subjects receive during the trial must be documented on the CRF. The medication name, dose, frequency, route of administration, date(s) of administration and reason for administration must be recorded. This documentation should continue through the treatment period and the follow up period.

All medication will be coded using the World Health Organization Drug Dictionary (WHO-DD) version B2 September 2017 WHO DDE. Preferred drug name, Anatomical Therapeutic Chemical (ATC) class will be reported for inclusion in the database.

Medication summaries based on ATC level 2, and the preferred drug names will be produced for the Safety Analysis Set. The summaries will present, by group, treatment, the frequency and percentage of subjects who used any medication in an ATC class, or any medication based on a single preferred drug name. Medication summaries will be sorted by descending total frequency of ATC class and by PT within ATC class. Subjects will be counted only once for each medication class and each preferred drug name.

For subject listings, medications will be reported based on ATC class and PT; multiple medications for an individual subject will be listed by start date and then by stop date, from earliest to latest medications.

7.5.1. Prior Medication

Any medication with a stop date prior to or on first dose date for both treatment groups during the treatment period will be considered a prior medication.

No summary for prior medication will be presented. Prior and concomitant medications will be presented together on a single listing. Prior medication will be indicated in the listing.

7.5.2. Concomitant Medication

A concomitant medication is any medication that the subject has been taking prior to the first dose of study medication during the treatment period and that the subject is expected to continue to take for some portion of the study treatment period, as well as any medication other than the investigational product that the subject takes during the course of the study treatment period. In the case of completely missing stop date, medication will be assumed to be concomitant. If a medication can be considered as both prior and concomitant due to partial stop date, it will be considered as concomitant.

Any medication a subject takes to treat hyperglycemia after the first dose of double-blind study drug starts and continues for more than 2 weeks before study treatment stops is considered a rescue therapy and should be recorded in the concomitant medication log. Increase total daily dose of metformin, the background hypoglycemic therapy, is also considered rescue medication if the dose is intensified during the double-blind treatment period for more than 2 weeks before study treatment stops. All medication given to treat hyperglycemia will be recorded in CRF.

Concomitant medications will be presented in a summary table by group and treatment as well as in a subject listing. Rescue medications will be summarized in a separate table and will be indicated in the same listing for concomitant medications.

8. EFFICACY

Efficacy data include HbA_{1c}, FPG, SBP and body weight. All continuous efficacy data will be summarized for observed and change values by treatment and visit for double-blind treatment group and high glycemic group separately. Proportion of subjects with HbA_{1c} of ≤ 7% will be summarized by frequency and proportion by treatment and visit for double-blind treatment group and high glycemic group separately. No formal inference analyses will be performed for the open-label study.

8.1. PRIMARY EFFICACY ENDPOINT AND ANALYSIS

The primary efficacy hypothesis is that bexagliflozin reduces HbA1c after 24 weeks of treatment when compared to placebo for the double-blind groups.

8.1.1. Primary Efficacy Analysis

Let $\mu_{Bexagliflozon}$ and $\mu_{placebo}$ represent the mean changes from baseline in HbA_{1c} at Week 24 for bexagliflozin and placebo arms, respectively. The following hypotheses will be tested:

$$H0: \mu \text{ bexagliflozin} - \mu \text{ placebo} \geq 0$$

$$Ha: \mu \text{ bexagliflozin} - \mu \text{ placebo} < 0$$

This hypothesis will be tested based on ITT analysis set. Missing data will be imputed via multiple imputations, following which the MMRM will be repeated on the complete datasets with results combined across complete datasets using standard multiple imputations techniques; HbA_{1c} values collected after the start of rescue medication will not be excluded. A pattern-mixture model will be used to explore the impact of the missing data. Imputation will be conducted within each treatment arm and subgroup (treatment completer vs. treatment terminated early) under the assumption that non-adherent subjects with missing data will follow the same trajectory of non-adherent subjects with values observed. All efforts will be made to retain subjects in the study. If treatment is discontinued, subjects are encouraged to remain in the study and complete all scheduled visits, including final follow-up visit. However, with all efforts, insufficient values may be available for reliable imputation. This analysis will be performed only if ≥3% (approximately 5) of subjects in each treatment arm completed scheduled week 24 visit after treatment is stopped. This number is chosen to allow sufficient data to establish a regression model for imputation. For this analysis, the following three-step approach outlined below will be used:

a. Non-monotone (intermediate visits) missing data will be imputed first using the Monte Carlo Markov Chain (MCMC) method under the MAR assumption in all treatment arms (using the MCMC statement in PROC MI). Multiple chains option

(CHAIN=MULTIPLE option in the MCMC statement of PROC MI) will be used. For the non-monotone imputation of the HbA_{1c} missing data, a multivariate normal model will be used including variables for the HbA_{1c} at baseline and all post-baseline visits within each treatment group. Five hundred imputed datasets will be generated.

b. After the non-monotone missing data have been imputed, the remaining monotone missing data will be imputed within each treatment arm and subgroup, defined by whether subject completed 24 week of study treatment, using regression approach. The predictors for the regression imputation model at any time point will be country, and HbA_{1c} at all previous time points, including baseline. Imputations will be performed using a sequence of regression-based imputations (using PROC MI statement MONOTONE REG) at each post-baseline time point.

c. Imputed data in each of the multiple imputed datasets will be analyzed using a mixed model repeated measures (MMRM) approach. The MMRM model will include treatment, visit, treatment-by-visit interaction, the baseline HbA_{1c} value and country (US or Japan) as a fixed effect covariate. The analysis will evaluate the mean change from baseline in HbA_{1c} over the 24 week double blind treatment period. An unstructured covariance will be used to model the within subject correlation. The Kenward Roger approximation will be used to estimate denominator degrees of freedom. If the model with the unstructured covariance structure does not converge, an autoregressive(1) covariance structure will be used. HbA_{1c} values obtained after the start of rescue medication will not be excluded from the analysis. The treatment and treatment by visit interaction terms allow for comparisons of the treatment groups at each visit, and over week 6 to week 24. Least squares (LS) mean treatment differences between the bexagliflozin group and the placebo group at week 24 will be estimated from the model The LS mean results from all imputed datasets will be combined using the Rubin's combination rule (PROC MIANALYZE).

If no sufficient week 24 values from treatment discontinued subjects are available for imputation, MMRM analysis will be performed using all available data. Method described in above step c will be utilized.

Descriptive statistics (n, mean, Q1, median, Q3, SD, minimum, and maximum) will be reported by treatment group and visit, along with the LS means, differences between LS means, a 2-sided 95% confidence interval for each difference, p-values from the model effects. In addition, the LS means with standard errors (SEs) of the change from baseline over time and difference between treatment groups with 95% CIs will be presented graphically for the ITT population.

For supportive analyses, regardless of having sufficient week 24 values from treatment discontinued subjects are available for imputation or not, MMRM analysis will be performed using all available data. The primary efficacy endpoint will also be analyzed with all available data using the PP analysis set using MMRM method.

8.1.2. Sensitivity Analyses

Randomized subjects who withdraw consent to participate in the study will not be replaced. The drop-out rate is estimated to be 12%. To the extent possible, attempts will be made to minimize the amount of missing data through measures planned in the study conduct.

To investigate the possible implications of missing values in efficacy assessments, the number, timing, pattern, and reason for the missing value will be summarized. The reason for withdrawn will be reviewed. If there are missing values for the primary analysis, only available data will be analyzed and data obtained after rescue will not be excluded. The drop-out patterns will be assessed by a Kaplan-Meier plot of time to discontinuation by treatment group to assess whether they differ between treatment groups.

To evaluate the impact from missing data and rescue medication, the following sensitivity analyses will be conducted:

1) Tipping point analysis will be conducted. Data after rescue medication will be considered as missing. Following steps will be followed:

    a. Create monotone missing data as step a in Section of 8.1.1;

    b. Regression approach under MAR will be used to generate complete datasets. The predictors for the regression imputation model at any time point will be country, treatment, and HbA_{1c} at all previous time points, including baseline. The first imputed values will be penalized or rewarded based on the treatment subject received:

        - Subjects in the bexagliflozin group: missing value will be analyzed assuming the treatment effect is worsened by $\delta 1$ (where δ = 0.1 to 0.5, in step of 0.1) compared to the subjects who have no missing value in the reduction of HbA_{1c} value;

        - Subjects in the placebo group: missing value will be analyzed assuming the treatment effect is better by $\delta 2$ (where δ = 0 to 0.5, in step of 0.1) compared to the subjects who have no missing value in the reduction of HbA_{1c} value;

The penalty or reward will not be applied to the values from later time points as these values are penalized or rewarded through regression on previous time points.

c. Imputed datasets will be analyzed and results combined as step c in Section of 8.1.1.

For each combination of (δ1, δ2), 100 imputed datasets will be obtained. These 25 combinations will be separately analyzed to explore under which condition where the null hypothesis can no longer having evidence to be rejected.

2) Additionally, the impact of the data after the start of rescue medication will be evaluated. HbA_{1c} values collected after the start of rescue medication will be considered missing, and the MMRM analyses will be performed. Model as specified in the primary analysis will be used.

3) Furthermore, data that are missing or after rescue medication will be imputed using LOCF. All observed data and the imputed values will be analyzed using MMRM model. Model as specified in the primary analysis will be used.

8.1.3. Subgroups

The primary efficacy endpoint on available data will be summarized and analyzed by the following subgroups using MMRM model:

- Age (<65 years or ≥65 years), adjusted by country and baseline HbA_{1c} value;
- Gender (male or female) , adjusted by country and baseline HbA_{1c} value;
- Race (Asian; Black of African American; White or Caucasian; Other) , adjusted by country and baseline HbA_{1c} value;
- Baseline HbA_{1c} (≥ 8.5% and ≤ 10.5% vs. ≥ 7.5% and < 8.5%), adjusted by country;
- Country (US or Japan), adjusted by baseline HbA_{1c}.

8.2. SECONDARY EFFICACY ENDPOINT(S) AND ANALYSES

Analysis for secondary endpoints will be based on the ITT analysis set and repeated for PP analysis set. No formal hypothesis testing will be performed for secondary efficacy endpoints. All nominal p-values will be presented for exploratory purpose.

8.2.1. Change from baseline in FPG over time

The comparison of change in FPG between randomized treatments of the double-blind groups at each visit will be carried out using the similar MMRM model as section 8.1.1. The model will include terms for treatment, visit, treatment-by-visit interaction,

country as fixed effects and the corresponding baseline FPG value as covariate. LS means for each treatment and the difference between treatments, treatment comparison p-values and difference for each visit will be estimated from the model, with the two-sided 95 % CIs of the treatment difference also presented. Summary table will be presented and figures will be displayed for the ITT analysis set in both double-blind and high glycemic groups. The by-visit summary statistics will be provided for safety analysis set for high glycemic group.

8.2.2. Change in SBP over time

Change from baseline in SBP at each visit for the double-blind groups will be analyzed in a similar manner as change in FPG, Section 8.2.1, based on ITT analysis set for all subjects. Baseline SBP will be used in place of baseline FPG. Country and baseline HbA_{1c} value will also be considered as fixed effects. Summary table will be presented and figures will be displayed for the ITT analysis set in both double-blind and high glycemic groups.

8.2.3. Proportions of subjects with HbA_{1c} of ≤ 7% over time

Summary on the proportion of subjects achieving HbA_{1c} < 7% at each visit will be presented over time for both double-blind and high glycemic groups. For the double-blind groups, data will be analyzed using Generalized Estimating Equation (GEE) logistic regression. The fixed effect will include country, treatment, visit, treatment-by-visit interaction and baseline HbA_{1c}. An unstructured correlation structure will be used (or autoregressive(1) if the model with the unstructured structure does not converge). The odds ratios of bexagliflozin group over the placebo group at each visit will be estimated from LS means based on the model with the corresponding p-values and their two-sided 95 % CIs presented.

8.2.4. Change in body weight in subjects with baseline BMI ≥ 25 kg/m^2 at week 24

Change from baseline in body weight for the double-blind groups will be analyzed using an analysis of covariance model (ANCOVA), based on ITT analysis set for all subjects with baseline BMI ≥ 25 kg/m^2. Baseline body weight will be used as covariate, and country and baseline HbA_{1c} value will also be used as fixed effects.

Summary table will be presented and figures will be displayed for both double-blind groups and high glycemic group based on the ITT analysis set.

8.2.5. Change from baseline in HbA_{1c} in subjects who have baseline HbA_{1c} ≥ 7.5% and ≤ 10.5% over time

Change from baseline in HbA1c will be analyzed as part of primary efficacy analysis

(Section 8.1.1).

8.2.6. Change from baseline in HbA_{1c} in high glycemic group (subjects who have baseline HbA_{1c} > 10.5% and ≤ 12.0%) over time

The change from baseline in HbA_{1c} over time for high glycemic group will be presented at each time point descriptively in a separate table. Point estimate of the change and its corresponding 95% CI (normal assumption) will be constructed for each time point. All observed values will be summarized at each visit that HbA_{1c} is assessed. In addition, if week 24 value is missing or subject used rescue medication before week 24 assessments, the last post-baseline value prior to the start of rescue medication will be carried forward to week 24. The LOCF week 24 will also be summarized.

9. SAFETY

The analysis population used for safety analyses will be the safety analysis set. Safety data include AEs; physical examination results; vital signs, including blood pressures; ECG results; and clinical laboratory results, including serum chemistry, hematology, serum lipids, and urinalysis.

Observed data will be summarized by group and treatment as counts and percentages for discrete variables and means, standard deviations, medians, inter-quartile range, minimum, and maximum for continuous variables. All subjects who receive at least one dose of study medication will be included in the safety analysis. High glycemic subjects safety observations will be summarized as a stand-alone group in the summary tables.

9.1. EXTENT OF EXPOSURE

For each group, study drug exposure will include:

- Treatment duration by treatment
- Total dose received by treatment

Treatment duration of tablets (in weeks) is calculated as (the date of the last dose of study drug – the date of the first dose of study drug + 1) / 7 and rounded to 1 decimal place.

The specific definitions of the first dose and last dose dates of study drug are given below:

- First dose date: The date of the first dose of study treatment in treatment period obtained from the study drug exposure CRF.

- Last dose date: The date of the last dose of study medication in the 24 week treatment period for subjects who have completed the study or discontinued early. Total dose received will be calculated as number of tablets dispensed – number of tablets returned.

Summary statistics for treatment duration (in weeks) and total dose received, as well as a frequency summary of treatment duration categories (e.g., < 1, 1 - < 6 weeks), will be provided.

9.2. TREATMENT COMPLIANCE

Subjects will be provided with dosing instructions each time study medication is dispensed. Subjects will also be instructed to bring their medication with them to every visit. During the run-in period, subjects will be considered compliant in investigational product administration by missing no more than one day of run-in medication. Subjects who are not compliant during the run-in period will be excluded from treatment period. If, in the judgment of the investigator, it was appropriate for the subject to omit these doses, this requirement may be waived (e.g., if the subject was hospitalized overnight during run-in).

At each visit after the start of the run-in period, the study staff will review the self-monitored blood glucose (SMBG) control diary and medication use with the subject and record the drug consumption in the CRF. Reasons for non-adherence will also be recorded in the protocol deviation log if applicable.

Compliance in the run-in and treatment period is calculated for as follows:

- Percent compliance = (number of tablets taken / number of tablets should have taken) x 100.
- Number of tablets taken = number of tablets dispensed – number of tablets returned.
- Number of tablets should have taken = (number of tablets supposed to take in a day) x (number of exposure days).
- Number of exposure days = last dose date – first dose date + 1.

If any of the bottles dispensed is not returned, it will not be possible to compute the compliance. In this case, the number of tablets taken and compliance will be considered as missing. Summary statistics for tablet compliance (%) will be provided by group and treatment for the treatment period. A frequency summary of compliance will also be presented with the following categories: < 75%, 75-<100%, 100-120%, and > 120%.

9.3. ADVERSE EVENTS

Adverse events will be collected and recorded from the time a subject signs the Informed Consent Form (ICF) to the last scheduled contact. And new SAEs reported by the subject to the investigator that occur after the last scheduled contact, and are determined by the investigator to be reasonably associated with the use of the investigational product, should be reported to the sponsor or designated personnel. This may include SAEs that are captured on follow-up telephone contact or at any other time point after the defined study period (i.e., up to last scheduled contact). The investigator should follow SAEs identified after the last scheduled contact until the events are resolved, or the subject is lost to follow-up. The investigator should continue to report any significant follow-up information to the sponsor until the event has been resolved. This study requires that subjects be actively monitored for SAEs for at least 4 weeks after the last treatment.

For all AEs, preferred AE terms and system organ class SOC will be coded using terminology from the MedDRA, version 20.1.

A Treatment-emergent Adverse Event (TEAE) is defined as an AE that begin after the first administration of the study medication during the treatment period or existing AEs that worsen after the first dose of the study medication during the treatment period. All reported AEs will be listed, but only TEAEs will be summarized in tables.

Drug-related AEs will be considered those to be possibly, probably and definitely related to study medication administration based on the investigators assessment.

Unless otherwise specified, AEs will be summarized by SOC and PT, with SOCs and PTs within SOCs presented in descending order of subject incidence.

9.3.1. Derived Data

AE onset day is calculated as (date of AE start – date of first dose in treatment period + 1) if it is after the first dose of study drug. AE onset day is calculated as (date of AE start – date of first dose in treatment period) if it is before the first dose of study drug. The onset day will be missing if the start date is missing.

9.3.2. Data Summarization

AE summary tables are listed below:

- An overall summary of the number and percentage of subjects reporting any TEAEs, serious TEAEs, treatment-related TEAEs, serious treatment-related TEAE, any TEAEs leading to treatment discontinuation, any TEAEs leading to subject discontinuation and TEAEs leading to death.

- TEAEs overall and by SOC and PT
- TEAEs by severity, overall and by SOC and PT
- Serious TEAEs, overall and by system organ class and preferred term
- TEAEs by relationship to study treatment, overall and by system organ class and preferred term
- TEAEs leading to treatment discontinuation, overall and by SOC and PT
- TEAEs leading to study discontinuation, overall and by SOC and PT
- Most common TEAEs. Most common TEAEs are defined as TEAEs that occur in > 5% of the subjects in either of the treatment groups.

For summary tables, subjects having more than 1 event with the same PT will be counted once for that term. Subjects having more than 1 event with the same SOC will be counted once for each event and once for that SOC. For tabulations by severity, only a subject's most severe event within the category (e.g. overall, SOC, PT) will be counted; similarly, for tabulations by relationship, only a subject's most related event within a category will be counted. The denominator for percentages will be the number of subjects in the Safety Analysis Set for the given treatment group (i.e., the N's for the columns).

Listings will be provided for all AEs and the following subsets:

- All TEAEs
- All TEAEs at least possibly related to study medication
- Serious AEs
- AEs leading to treatment discontinuation
- AEs leading to death

9.3.3. AE of Special Interest

AE of special interest include Urinary Tract Infections (UTI), Genital Mycotic Infections (GMIs), diuretic effects, hypotension episodes, hypoglycemia, hepatotoxicity, Major Adverse Cardiovascular Event (MACE), falls and fractures, malignancy, hypersensitivity reactions, DKA, pancreatitis, amputations and renal failure events. These AEs of interest except MACE and amputations will be prospectively identified based on the MedDRA preferred terms in the AE log by a medical expert prior to the data base lock and unblinding of the individual subject treatment assignment. The list of AE of interest will be confirmed in a peer review process. MACE will be identified by the investigator and documented in the CRF, and subsequently adjudicated by an independent committee. Adjudicated results will be used for summary. Adjudicated results will be used for summary. Amputation events will be recorded in the procedures and

amputation CRF. Cardiovascular events, hypoglycemia events by severity, and amputations will be summarized separately.

9.3.3.1. AE of special interest identified by PTs

Cardiovascular events considered as MACE by investigator will be submitted to an independent CEC for adjudication. The events of interest include cardiovascular mortality, MI, stroke, hospitalization for acute coronary syndromes, urgent revascularization procedures, and other possible serious cardiovascular events. Other AE of special interest will be identified from the PTs. The number and percentage of subjects experiencing these TEAEs of special interest will be summarized for each treatment group by type of event. The incidence rate of AE of special interest per 100 patient years will also be summarized. Each category of events will be displayed in a separate listing.

9.3.3.2. Hypoglycemic Events

Hypoglycemic event categories include:

Category	Description
Severe	Assistance required and blood glucose ≤70 mg/dL or no value available but responded to glucose treatment
Documented Symptomatic	Blood glucose ≤70 mg/dL and typical symptoms of hypoglycemia
Asymptomatic	Blood glucose ≤70 mg/dL and no typical symptoms of hypoglycemia
Probable Symptomatic	Typical symptoms of hypoglycemia and no value available but responded to glucose treatment
Relative	Typical symptoms of hypoglycemia and blood glucose >70 mg/dL

While each event meeting the criteria above will be entered into the hypoglycemia log, only critical (severe) hypoglycemia, documented symptomatic hypoglycemia, asymptomatic hypoglycemia, and probable hypoglycemia will be entered as AEs.

The number and percentage will be summarized by treatment for:

- Each category of hypoglycemic events;
- Any severe or documented hypoglycemic events.

9.3.3.3. Revascularization and Amputations

Revascularization and amputations information are collected in a separate form. Frequency and percentage will be summarized for:

- Type of procedures –cardiovascular related or amputation

- Subjects with any amputation
- Conditions that resulted in amputation
- Location of amputation

Only procedures performed after the first dose of study drug during the treatment period will be summarized.

9.4. LABORATORY EVALUATIONS

Laboratory tests will include hematology panel, chemistry panel, serum lipids, and urinalysis testing. Chemistry, and urinalysis will be performed at the following time points: at the screening visi and week 6, 12, 18, 24, and 26. Hematology will be performed at the screening visit and week 6, 12, 24 and 26. Serum lipids will be performed at the screening visit, and on week 12 and 24, and 26. The study staff will contact each subject prior to a scheduled clinic visit to confirm the time of the visit and to remind the subject of proper fasting practice. A subject must be queried to assess compliance with an approximately 10-hour fast prior to blood draw to ensure the FPG and triglycerides values can be accurately determined. If a subject has not fasted for approximately 8 hours, the subject must return as soon as can be arranged (within 1 week) to provide a specimen after proper fasting. A list of laboratory tests is included in Table 5.

Low density lipoprotein cholesterol (LDL-C) will be calculated by the Friedewald equation. If triglycerides are > 400 mg/dL, the calculated LDL-C value is invalid by this equation and will be set as missing. Direct LDL-C will be determined in subjects whose triglycerides are > 350 mg/dL at screening visit. All subsequent LDL-C of these subjects will be determined by the same direct LDL-C measurements only. If triglycerides are >350 mg/dL, a reflex test will be performed based on direct LDL measurement. The following algorithm will be used to obtain LDL-C values for the analyses:

1. Select subjects (based on the SI unit) who had screening triglycerides >3.4 or >350 based on the conventional unit
2. Take the LDL – direct measurement values only, throughout the study visits for those subjects
3. If screening triglycerides > 350 and no direct LDL-C values have been determined, take the calculated.

Among those subjects who have screening triglycerides >350 and have both calculated and direct LDL values, only take the direct LDL.

Urinalysis microscopy will be conducted if the subject has a positive result on the leukocyte esterase or nitrite dipstick tests to clarify the significance of the finding. Results of glucose measurement in the urinalysis must be suppressed from the

laboratory reports so the sponsor, investigators, study coordinators, pharmacists, study subjects, and the CEC members will remain blinded to the dosing assignment.

The baseline value will be the latest value obtained prior to Day 1. Change from baseline for all continuous parameters will be calculated as the post-baseline value minus the baseline value. All continues variables will be summarized by number of subjects [n], mean, SD, Q1, median, Q3, minimum and maximum and categorical variables will be summarized by frequency and percentage.

All scheduled and unscheduled results will be considered in tables that assess maximum grade or toxicity.

Observed values (in SI units) and change from baseline over time will be summarized by group and treatment. Laboratory data will be classified as low, normal, or high relative to the parameter's reference range. Laboratory abnormalities for each treatment will also be summarized with shift tables for selected parameters.

All laboratory data will be listed. For hematology, chemistry, and serum lipids, columns will be included for normal ranges and individual abnormal laboratory values will be flagged and clinical significance will be indicated. A listing for the microscopic examination will be provided for subjects who have a positive result from the urinalysis dipstick evaluation.

Table 5 List of Laboratory Tests

TEST NAME		SHIPMENT
Hematology		Ambient
• Hematocrit (Hct) • Hemoglobin (Hgb) • Mean corpuscular hemoglobin (MCH) • Mean corpuscular hemoglobin concentration (MCHC) • Platelet count	• Mean corpuscular volume (MCV) • Red cell distribution width (RDW) • Red blood cell (RBC) count • White blood cell (WBC) count with differential	
Serum Chemistry and Electrolytes		Ambient
• Albumin (ALB) • Alanine aminotransferase (ALT) • Aspartate aminotransferase (AST) • Alkaline phosphatase (ALK-P) • Blood urea nitrogen (BUN) • Glucose • Bicarbonate (HCO3) • Creatinine • Chloride (Cl)	• Total protein • Calcium (Ca) • Magnesium • Phosphorus • Potassium (K) • Sodium (Na) • Total bilirubin • Direct bilirubin • Uric acid	
Glycemic Control		Ambient
• Fasting plasma glucose (FPG) • Hemoglobin A_{1c} (HbA_{1c})		
Serum Lipids		Ambient
• Total cholesterol (TC) • High-density lipoprotein cholesterol (HDL-C) • Triglycerides (TG)	• Low-density lipoprotein cholesterol (LDL-C), calculated • LDL-C, direct	
Urinalysis		Ambient
• Appearance • Bilirubin • Color • Glucose • Ketones • Microscopic examination of sediment • UACR	• Nitrite • Occult blood • pH • Protein • Specific gravity • Urobilinogen • Leukocyte esterase	
Urine Pregnancy Test (WOCBP)		Local
Population PK Sampling Bexagliflozin plasma level		Frozen

9.5. VITAL SIGNS

Vital signs will be measured at the screening visit, on randomization visit, and at week 6, 12, 18, 24, and 26. Measurements of vital signs will include measurement of supine, sitting and standing BP measurements, and heart rate. Only the BP measured in the sitting position will be used to determine eligibility. Orthostatic systolic and diastolic BP will be calculated as supine measurement – standing measurement.

For BP, pulse rate, and respiration rate, observed values and change from baseline will be summarized by group, treatment and mapped visit using descriptive statistics (n, mean and median, standard deviation, Q1, and Q3, minimum and maximum). For BP, supine, sitting, standing, and orthostatic BP will be summarized. Subjects vital sign measurements, including scheduled and unscheduled visits, will be listed.

9.6. ELECTROCARDIOGRAM

A 12-lead ECG will be conducted at the screening visit, on randomization visit, and at week 24. ECG parameters measured will be the RR interval, PR interval, QRS duration, and QT interval. Each ECG should also be assessed by the investigator for signs of ischemia, clinically significant hypertrophy, and clinically significant T-wave abnormalities or arrhythmia.

If a subject's ECG parameters cannot be determined due to pacemaker placement or atrial fibrillation, the ECG parameters will be considered missing. Any machine generated values such as 0 or 9999 will be excluded from the analyses.

For each abnormal ECG result, the investigator shall ascertain if the observation represents a clinically significant change from the screening ECG for that individual subject (this determination, however, does not necessarily need to be made the first time an abnormal result is observed. The investigator may repeat the ECG to verify the results of the original result). If the ECG result is determined to be a clinically significant and abnormal change from baseline for that subject, it is considered an AE.

For the ECG parameters, observed values and change from baseline from scheduled visits will be summarized with descriptive statistics by group, treatment and overall at each visit. The maximum change from baseline from scheduled visits will also be provided for ECG parameters.

For the ECG overall assessment, the number and percentage of subjects in each overall assessment category (normal, abnormal but not clinically significant, abnormal and clinically significant, missing) will be presented by group and treatment at each visit.

9.7. PHYSICAL EXAMINATION

A complete physical examination will be conducted at the randomization visit and at week 26. The examination will include measurement of body weight, and a general assessment of all body systems including the skin, head, eyes, ears, nose, throat, neck, heart, lungs, abdomen, lymph nodes, and extremities. An abbreviated physical examination will be conducted at screening visit and at week 24, or clinically indicated. The examination will include body weight and height (height will be measured only at screening), and general assessment of the skin, heart, lungs and abdomen. Physical examination findings will be presented in a by-subject listing.

10. INTERIM ANALYSES

No interim analyses are planned.

11. DATA AND SAFETY MONITORING BOARD

An independent DSMB will review descriptive summaries of accumulating safety, subject disposition and limited efficacy data every 6 months, or a frequency recommended by the DSMB.

A designated statistician who is not involved with the study operation will hold the treatment codes. The unblinded treatment information can be provided to the DSMB to facilitate the evaluation of any clinically important increase in the rate of a serious suspected adverse reaction or to the designated safety contact when the treatment information is required to determine if an expedited safety report must be submitted to regulatory agencies. The data outputs for this review will be created by an unblinded team. Personnel involved in the conduct of the study will not participate in the preparation of these outputs, receive the data, or participate in the unblinded portions of the DSMB meetings. More details will be provided for DSMB charter and DSMB SAP.

12. CHANGE FROM ANALYSIS PLANNED IN PROTOCOL

Protocol section 9.3.2 says “Safety analyses will be based on the medication that was actually taken by each subject.” SAP changes to use ‘Safety analyses will be based on the medication that was taken first.’ in section 5.2.

Protocol section 9.5.1 includes sensitivity analysis by imputing missing HbA_{1c} data via multiple imputation (MI). SAP changes to impute missing HbA_{1c} data via MI in primary efficacy analysis per FDA comments.

Protocol section 9.5.1 includes a sensitive analysis using jump to reference multiple imputation method. This has been removed after tipping point analysis is added.

13. REFERENCE LIST

Little, R., Yau, L. Intent-to-Treat Analysis for Longitudinal Studies with Drop-Outs. Biometrics, 1996; 52: 1324—1333.

Ratitch, B., O'Kelly, M. Implementation of Pattern-Mixture Models using Standard SAS/STAT Procedures. Pharmasug 2011 – Paper SP04, PHARMASUG.

Bodner TE. What improves with increased missing data imputations? Structural Equation Modeling: A Multidisciplinary Journal 2008; 15: 651—675.

White IR., Royston P., Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Statistics in Medicine 2011; 30: 377—399.

14. PROGRAMMING CONSIDERATIONS

The following conventions will hold for programming of outputs:

- SAS® Version 9.4 will be used for programming and production
- The format of the table shells will be followed as closely as possible; however, in the course of programming and familiarization with the database, some changes may become necessary. All changes will be documented. Major changes will be documented through a formal amendment to this document.
- Patients in this study will be identified as "Subjects."
- Descriptive statistics will be displayed in the following order:

n
Mean
Standard deviation (SD)
Q1, Q3
Median
Minimum, Maximum

- Decimal places: For summary statistics, the minimum and maximum will be reported with the same number of decimal places as the collected measure, the mean, LS mean (if applicable) and median will have 1 more decimal place than the measure collected, and the SD and confidence interval (CI) will have 2 more decimal places than the collected measure. For frequency distributions, percentages will be reported to 1 decimal place. For p-values, 4 decimal places will be reported or the SAS® p-value format of "< 0.0001" or "> 0.9999" will be reported.
- Unless otherwise noted, the denominator for percentages is the number of subjects in the applicable analysis population and treatment group.
- If the frequency for a particular table cell is zero, then "0", properly aligned, will be displayed (i.e. "0 (0.0%)" will not be displayed.)
- Non-numeric values: Where variables are recorded using < (e.g., "< 10" or "≤ 10") the numeric portion of the result will be used (e.g., < 10 and ≤ 10 becomes 10) for summary; where variables are recorded using > (e.g., "> 10" or " ≥ 10") the numeric portion of the result will be used (e.g. > 10 and ≥ 10 become 10) for summary; the actual recorded results, (e.g. "< 10" or "> 10") will appear in listings.

14.1. GENERAL CONSIDERATIONS

- A separate SAS program will be created for each output.

- Output files will be delivered in Word format.
- Numbering of TFLs will follow ICH E3 guidance

14.2. TABLE, LISTING, AND FIGURE FORMAT

14.2.1. General

- All TLFs will be produced in landscape format, unless otherwise specified.
- All TLFs will be produced using the Courier New font, size 9
- The data displays for all TLFs will have a minimum 1-inch margin on all 4 sides.
- Headers and footers for figures will be in Courier New font, size 9.
- Legends will be used for all figures with more than 1 variable, group, or item displayed.
- TLFs will be in black and white (no color), unless otherwise specified
- Specialized text styles, such as bolding, italics, borders, shading, and superscripted and subscripted text, will not be used in the TLFs, unless otherwise specified. On some occasions, superscripts 1, 2, or 3 may be used (see below).
- Only standard keyboard characters will be used in the TLFs. Special characters, such as non-printable control characters, printer-specific, or font-specific characters, will not be used. Hexadecimal-derived characters will be used, where possible, if they are appropriate to help display math symbols (e.g., μ). Certain subscripts and superscripts (e.g., cm2, Cmax) will be employed on a case-by-case basis.
- Mixed case will be used for all titles, footnotes, column headers, and programmer-supplied formats, as appropriate.

14.2.2. Headers

- All output should have the following header at the top left of each page:
  Theracos Sub, LLC
  Protocol Number: THR-1442-C-419
- Draft or Final in top right corner.
- All output should have Page n of N at the top of each page. TLFs should be internally paginated in relation to the total length (i.e., the page number should appear sequentially as page n of N, where N is the total number of pages in the table).
- The date output was generated should appear along with the program name as a footer on each page.

14.2.3. Display Titles

- Each TLF should be identified by the designation and a numeral. (i.e., Table 14.1.1). ICH E3 numbering is strongly recommended but sponsor preferences should be obtained prior to final determination. A decimal system (x.y and x.y.z) should be used to identify TLFs with related contents. The title is centered. The analysis set should be identified on the line immediately following the title. The title and table designation are single spaced. A solid line spanning the margins will separate the display titles from the column headers. There will be 1 blank line between the last title and the solid line.

Table x.y.z
First Line of Title
Second Line of Title if Needed
ITT Analysis Set

14.2.4. Column Headers

- Column headings should be displayed immediately below the solid line described above in initial upper-case characters.
- In the case of efficacy tables, the variable (or characteristic) column will be on the far left followed by the treatment group columns and total column (if applicable). P-values will be presented in a separate comparison column (if applicable).
- For numeric variables, include “unit” in column or row heading when appropriate.
- Analysis set sizes will be presented for each treatment group in the column heading as (N=xx) (or in the row headings if applicable). This is distinct from the ‘n’ used for the descriptive statistics representing the number of subjects in the analysis set.
- The order of treatments in the tables and listings will be active treatment first, then placebo, followed by a total column (if applicable).

14.2.5. Body of the Data Display

14.2.5.1. General Conventions

Data in columns of a table or listing should be formatted as follows:

- alphanumeric values are left-justified;
- whole numbers (e.g., counts) are right-justified; and
- numbers containing fractional portions are decimal aligned.

14.2.5.2. Table Conventions

- Units will be included where available
- If the categories of a parameter are ordered, then all categories between the maximum and minimum category should be presented in the table, even if n=0 for all treatment groups in a given category that is between the minimum and maximum level for that parameter. For example, the frequency distribution for symptom severity would appear as:

Severity Rating	N
severe	0
moderate	8
mild	3

  Where percentages are presented in these tables, zero percentages will not be presented and so any counts of 0 will be presented as 0 and not as 0 (0%).

- If the categories are not ordered (e.g., Medical History, Reasons for Discontinuation from the Study, etc.), then only those categories for which there is at least 1 subject represented in 1 or more groups should be included.
- An Unknown or Missing category should be added to any parameter for which information is not available for 1 or more subjects.
- Unless otherwise specified, the estimated mean and median for a set of values should be printed out to 1 more significant digit than the original values, and standard deviations should be printed out to 2 more significant digits than the original values. The minimum and maximum should report the same significant digits as the original values. For example, for systolic blood pressure:

```
n                 XX
Mean             XXX.X
SD                 X.XX
Q1               XXX.X
Median           XXX.X
Q3               XXX.X
Minimum          XXX
Maximum          XXX
```

- P-values should be output in the format: “0.xxxx”, where xxxx is the value rounded to 4 decimal places. Any p-value less than 0.0001 will be presented as <0.0001. If the p-value is returned as >0.9999 then present as >0.9999
- Percentage values should be printed to one decimal place, in parentheses with no spaces, one space after the count (e.g., 7 (12.8%), 13 (5.4%)). If the rounded percentage is 0.0, display as '<0.1'. Unless otherwise noted, for all percentages, the number of subjects in the analysis set for the treatment group who have an observation will be the denominator. Percentages after zero counts should not be

displayed and percentages equating to 100% should be presented as 100%, without any decimal places.

- Tabular display of data for concomitant and rescue medications should be presented by treatment class with the highest occurrence in the total column in decreasing order. Tabular display of data for medical history and adverse event data should be presented by the SOC using descending order. Within the drug class and SOC, medical history (by preferred term), drugs (by ATC2 code), and adverse events (by preferred term) should be displayed in decreasing order in the total column. If incidence for more than 1 term is identical, they should then be sorted alphabetically. Missing descriptive statistics which cannot be estimated should be reported as "-".
- The percentage of subjects is normally calculated as a proportion of the number of subjects assessed in the relevant treatment group (or overall) for the analysis set presented. However, careful consideration is required in many instances due to the complicated nature of selecting the denominator, usually the appropriate number of subjects exposed. Describe details of this in footnotes or programming notes.

- For categorical summaries (number and percentage of subjects) where a subject can be included in more than one category, describe in a footnote or programming note if the subject should be included in the summary statistics for all relevant categories or just 1 category and the criteria for selecting the criteria.

- Where a category with a subheading (such as system organ class) has to be split over more than one page, output the subheading followed by "(cont)" at the top of each subsequent page. The overall summary statistics for the subheading should only be output on the first relevant page.

14.2.5.3. Listing Conventions

- Listings will be sorted for presentation in order of treatment groups as above, subject number, visit/collection day, and visit/collection time.
- Missing data should be represented on subject listings as either a hyphen ("-") with a corresponding footnote ("- = unknown or not evaluated"), or as "N/A", with the footnote "N/A = not applicable", whichever is appropriate.
- Dates should be printed in SAS® DATE9.format ("ddMMMyyyy": 01JUL2000). Missing portions of dates should be represented on subject listings as dashes (--JUL2000). Dates that are missing because they are not applicable for the subject are output as "N/A", unless otherwise specified.
- All observed time values must be presented using a 24-hour clock HH:MM or HH:MM:SS format (e.g., 11:26:45, or 11:26). Time will only be reported if it was measured as part of the study.
- Units will be included where available

14.2.5.4. Figure Conventions

- Unless otherwise specified, for all figures, study visits will be displayed on the X-axis and endpoint (e.g., treatment mean change from Baseline) values will be displayed on the Y-axis.

14.2.6. Footnotes

- A solid line spanning the margins will separate the body of the data display from the footnotes.
- All footnotes will be left justified with single-line spacing immediately below the solid line underneath the data display.
- Footnotes should always begin with "Note:" if an informational footnote, or 1, 2, 3, etc. if a reference footnote. Each new footnote should start on a new line where possible.
- Subject specific footnotes should be avoided, where possible.
- Footnotes will be used sparingly and must add value to the table, figure, or data listing. If more than six lines of footnotes are planned, then a cover page may be used to display footnotes, and only those essential to comprehension of the data will be repeated on each page.
- The last line of the footnote section will be a standard source line that indicates the name of the program used to produce the data display, date the program was run, and the listing source (i.e., 'Program : myprogram.sas Listing source: 16.x.y.z').

15. QUALITY CONTROL

SAS programs are developed to produce clinical trial output such as analysis data sets, summary tables, data listings, figures or statistical analyses. INC Research SOP 03.010 and 03.013 provide an overview of the development of such SAS programs.

INC Research SOP 03.009 describes the quality control procedures that are performed for all SAS programs and output. Quality control is defined here as the operational techniques and activities undertaken to verify that the SAS programs produce the proper clinical trial output by checking for their logic, efficiency and commenting and by review of the produced output.

16. INDEX OF TABLES

17. INDEX OF FIGURES

18. INDEX OF LISTINGS

19. MOCK-UPS

Attachment 1: Planned Table Shells

Attachment 2: Planned Listing Shells

Attachment 3: Planned Figure Shells

20. APPENDICES