Statistical Analysis Plan

Study Code	D168BC00001
Edition Number	3.0
Date	21Sep2017

A 24-WEEK INTERNATIONAL, MULTICENTER, RANDOMIZED, OPEN-LABEL, ACTIVE-CONTROLLED, PARALLEL GROUP, PHASE 3B TRIAL WITH A 28-WEEK EXTENSION TO EVALUATE THE EFFICACY AND SAFETY OF SAXAGLIPTIN CO-ADMINISTERED WITH DAPAGLIFLOZIN COMPARED TO INSULIN GLARGINE IN SUBJECTS WITH TYPE 2 DIABETES WHO HAVE INADEQUATE GLYCEMIC CONTROL ON METFORMIN WITH OR WITHOUT SULFONYLUREA THERAPY

A 24-WEEK INTERNATIONAL, MULTICENTER, RANDOMIZED, OPEN-LABEL, ACTIVE-CONTROLLED, PARALLEL GROUP, PHASE 3B TRIAL WITH A 28-WEEK EXTENSION TO EVALUATE THE EFFICACY AND SAFETY OF SAXAGLIPTIN CO-ADMINISTERED WITH DAPAGLIFLOZIN COMPARED TO INSULIN GLARGINE IN SUBJECTS WITH TYPE 2 DIABETES WHO HAVE INADEQUATE GLYCEMIC CONTROL ON METFORMIN WITH OR WITHOUT SULFONYLUREA THERAPY

Study Statistician PPD

22 Sep, 2017

Date

A 24-WEEK INTERNATIONAL, MULTICENTER, RANDOMIZED, OPEN-LABEL, ACTIVE-CONTROLLED, PARALLEL GROUP, PHASE 3B TRIAL WITH A 28-WEEK EXTENSION TO EVALUATE THE EFFICACY AND SAFETY OF SAXAGLIPTIN CO-ADMINISTERED WITH DAPAGLIFLOZIN COMPARED TO INSULIN GLARGINE IN SUBJECTS WITH TYPE 2 DIABETES WHO HAVE INADEQUATE GLYCEMIC CONTROL ON METFORMIN WITH OR WITHOUT SULFONYLUREA THERAPY

Global Product Statistician PPD

22 Sept 2017
Date

TABLE OF CONTENTS

AMENDMENT HISTORY

Date	Brief description of change
April 6, 2017	Initial Approved SAP
September 8, 2017	• Definition of confirmed hypoglycemia was updated. • In section 6.3.3 the definition of Randomized Subjects Data Set definition was updated by including all randomized subjects who receive at least one dose of study medication. • Section 6.3.4 is added to include the definition of Full Analysis Set. Following subsections were renumbered. • In section 6.3.5 the definition of Evaluable Subjects Data Set was updated. • Section 7.1.9 for analysis of 6-point self-monitored blood glucose data was updated. • Table 7.2-1 for list of relevant protocol deviations and their exclusion level was updated. • Section 7.3.2 for demographics and baseline characteristics was updated. • In section 7.4.1 additional categories were added for duration of exposure summary in the short term open-label treatment period. • In section 7.4.3 calculation for overall insulin compliance was added. • Section 7.5.2 was updated by adding new sensitivity analyses for primary endpoint. • Section 7.5.3 was added for handling missing efficacy data which includes additional MI based sensitivity analysis. Following sections were renumbered. Section 8.2 for selecting observations when there are multiple observations in a single visit window was updated. • Section 8.3 for laboratory evaluations was updated to include only data from central laboratory in safety tables • Section 8.4 was updated by adding visit windows for HbA1c and weight. • Section 10 was added to list the changes in analysis from protocol. Following section was renumbered. • Start of long-term treatment period was clarified in section 6.1. • Inserted texts from the core SAP and removed the references to core SAP.

1 BACKGROUND AND RATIONALE

This Phase 3 study is part of a clinical program to support the development of a saxagliptin 5 mg - dapagliflozin 10 mg fixed-dose combination (FDC) therapy for the treatment of T2DM. In subjects with T2DM who are inadequately controlled on metformin therapy with or without a sulfonylurea, this study will compare the efficacy and safety of the addition of saxagliptin-dapagliflozin versus the addition of insulin. The complementary mechanisms of action of saxagliptin and dapagliflozin, in combination with metformin, have been shown in previous studies to provide superior HbA1c lowering compared to either of the individual agents alone in subjects with inadequately controlled T2DM.

As a FDC therapy, saxagliptin-dapagliflozin provides a new treatment option for T2DM patients. Dapagliflozin inhibits renal glucose reabsorption and acts independently of insulin, while saxagliptin enhances glucose-mediated insulin secretion by a glucose-dependent mechanism (via incretin effect). Saxagliptin and dapagliflozin have demonstrated, both individually, in combination and in combination with metformin with or without SU, a favorable safety and tolerability profile. They have shown as single agents, as well as in combination with metformin, a low propensity for hypoglycemia consistent with their respective glucose dependent mechanism of action, therefore addressing a potential key concern when adding a glucose lowering agent. Both drugs have either demonstrated moderate weight reduction (dapagliflozin) or weight neutrality (saxagliptin) and do not require dose titration or injection, simplifying therapy compared to insulin.

While insulin is an effective glycemic treatment for T2DM, insulin is associated with several undesirable side effects that can negatively affect patient compliance and limit its effectiveness. These include increased risk of hypoglycemia and weight gain. Hypoglycemia is a clinically important barrier to optimizing treatment and there is emerging evidence that hypoglycemia is associated with negative cardiovascular outcomes. Over 85% of patients with T2DM are overweight or obese, and additional weight gain is undesirable and often results in reduced treatment compliance by the patients. In addition, not all patients are willing or able to inject insulin and keep up with the regular blood glucose monitoring required by insulin regimens.

Research Hypothesis:

In subjects with type 2 diabetes with inadequate glycemic control treated with metformin with or without SU, co-administration of saxagliptin and dapagliflozin over an open-label treatment period of 24 weeks will result in glycemic control, measured by HbA1c that is non-inferior compared to the addition of insulin glargine.

Schedule of Analyses:

This analysis plan presents the objectives, endpoints, and analyses that will be conducted once all randomized subjects have either completed or been discontinued from the short-term open-label period. All relevant queries must be resolved and the database must be locked for this period before the analyses. Short-term plus long-term objectives, endpoints, and analyses performed at the time of the final database lock are also specified within this SAP.

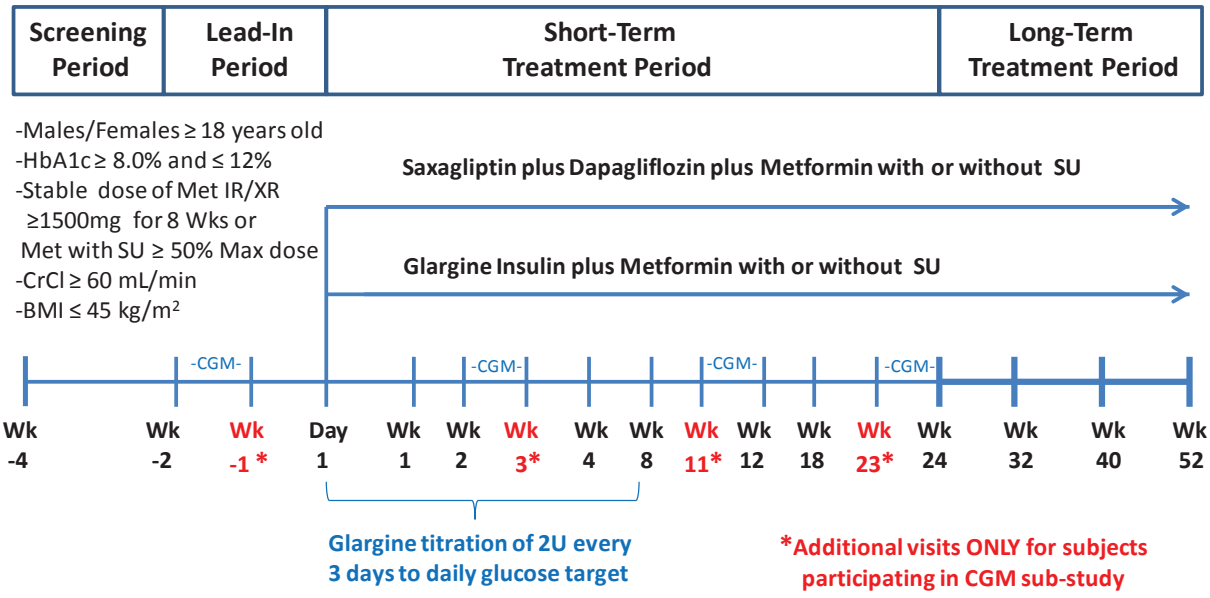
2 STUDY DESCRIPTION

2.1 Study Design

The CV181369 study is a randomized, open-label, two arm, parallel group, active-controlled, multicenter, Phase 3b study to evaluate the efficacy and safety of saxagliptin co-administered with dapagliflozin compared to insulin glargine. Subjects with documented T2DM and inadequate glycemic control (central laboratory HbA1c value ≥ 8% and ≤ 12% at screening) on a stable dose of metformin ≥ 1500 mg/day for at least 8 weeks prior to screening visit with or without a stable dose of SU of at least 50% the maximal dose per local label for at least 8 weeks prior to screening visit will be enrolled. The primary endpoint of HbA1c will be assessed at 24 weeks; however, the trial will continue to 52 weeks as a long-term extension.

The study design schematic is presented in Figure 2.1-1.

Figure 2.1-1: Study Design Schematic



Approximately 171 sites will randomize a combined total of approximately 598 subjects (299 subjects per treatment arm). Allowing approximately 12 months for patient recruitment, this study will be conducted over 24 months.

Continuous Glucose Monitoring (CGM) is a useful technology (in addition to HbA1c) to qualitatively, as well as quantitatively, monitor glycemic control in the form of time spent in the euglycemic/hyperglycemic/hypoglycemic (blood glucose ≤ 70 mg/dL) ranges and the mean amplitude of glucose excursions (MAGE). In this protocol, a subpopulation of approximately 250 randomized subjects (~125 subject/treatment arm) who agree to participate (separate informed consent) will have CGM performed for periods of 7 days at 4 different timepoints throughout the study (Week -2 to -1 (referred to as Baseline), Week 2 to 3 (referred to as Week 2), Week 11 to 12 (referred to as Week 12), and Week 23 to 24 (referred to as Week 24)).

2.2 Treatment Assignment

CCI

Subjects entering the 24-week open-label short-term treatment period

Following completion of the lead-in period, subjects who meet the criteria will be randomly assigned by the IVRS at the Day 1 Randomization visit, to one of the following two (2) open-label treatment arms in a 1:1 ratio:

- saxagliptin 5 mg and dapagliflozin 10 mg
- titrated insulin glargine

Randomization will be stratified by current use of background medication (metformin alone vs. metformin plus SU) at baseline to ensure equal representation across all treatment groups and also within the sub-study. Subjects will continue to receive their stable dose of metformin with or without SU throughout the study.

During the first 12 weeks following randomization, there will be no protocol specified rescue criteria. During this time, subjects receiving open-label insulin glargine will be titrated to target, which is driven over an (8) week period, then at Investigator's discretion with minimal adjustment recommended. Titration of open-label saxagliptin and dapagliflozin will not be allowed at any time during the study.

Subjects entering the 28-week open-label long-term treatment period

Subjects who complete the Short-term Treatment Period will continue into the Long-term Period taking the same open-label study medication (saxagliptin 5 mg QD and dapagliflozin 10 mg QD or insulin glargine) that they were randomized to on study Day 1. Subjects will continue to receive their stable dose of metformin with or without SU throughout the study.

2.3 Blinding and Unblinding

This is an open-label study. The investigator, BMS and AstraZeneca personnel, and subjects will be unblinded to treatment allocation throughout the short-term and long-term treatment periods. The database used for the analysis of the short-term treatment period of the study will be locked after all subjects have terminated the short-term treatment period of the study. Analyses will not be performed by treatment arm until the Short-term database lock. There will be one study report produced at the end of Short-term Period and a second study report produced at the end of Short-term plus Long-term period.

For the short-term treatment period for each subject, the HbA1c values will be masked to the investigator. After the short-term treatment period for each subject, the HbA1c values will no longer be masked. All other laboratory values will be unmasked to both the investigator and the sponsor.

2.4 Protocol Amendments

There has currently been one protocol amendment and the information has been incorporated into this document.

2.5 Cardiovascular Adjudication Committee

A Clinical Event Committee (CEC) blinded to the treatment of the subjects, will independently adjudicate all events of heart failure with hospitalization.

2.6 Hepatic Adjudication Committee

An Independent Hepatic Adjudication Committee, blinded to the treatment of the subjects, will determine the probability that drug-induced liver injury (DILI) is the cause of liver-related abnormalities, included but not limited to:

- AST and/or ALT > 3X ULN and TB >2 X ULN (within 14 days of the AST and/or ALT elevation)
- Hepatic events timely related to death (within 30 days before death)
- AST and/or ALT > 10X ULN

A separate Adjudication Charter will define and describe the procedure for the handling, reporting, and classification of these events.

3 OBJECTIVES

3.1 Primary Objective

To examine whether the mean change from baseline in HbA1c with co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU is non-inferior to titrated insulin glargine plus metformin with or without SU after 24 weeks of open-label treatment.

3.2 Secondary Objectives

- To compare the mean change from baseline in total body weight with co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU versus titrated insulin glargine plus metformin with or without SU after 24 weeks of open-label treatment.
- To compare the proportion of confirmed hypoglycemia [defined as: blood glucose □ 70mg/dL (3.9 mmol/L) recorded in the hypoglycemia module of the CRF] with co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU versus titrated insulin glargine plus metformin with or without SU during 24 weeks of open-label treatment.
- To compare the proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0%, without any hypoglycemia, with co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU versus titrated insulin glargine plus metformin with or without SU after 24 weeks of open-label treatment.
- To examine whether the change from baseline in the mean value of 24-hour glucose readings measured by Continuous Glucose Monitoring (CGM) with co-administered saxagliptin 5 mg and

dapagliflozin 10 mg plus metformin with or without SU is non-inferior to titrated insulin glargine plus metformin with or without SU after 2 weeks of open-label treatment.

- To examine whether the proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0%, with co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU is non-inferior to titrated insulin glargine plus metformin with or without SU after 24 weeks of open-label treatment.

3.3 Exploratory Objectives

- To assess the change from baseline in mean amplitude of glucose excursions (MAGE) of 24-hour glucose readings measured by CGM with co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU versus titrated insulin glargine plus metformin with or without SU after 2 weeks, 12 weeks and 24 weeks of open-label treatment.
- To assess the change from baseline in the mean value of 24-hour glucose readings measured by CGM with co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU versus titrated insulin glargine plus metformin with or without SU after 12 weeks and 24 weeks of open-label treatment.
- To assess the change from baseline in within-subject, within-day standard deviation of 24 hour glucose readings measured by CGM with co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU versus titrated insulin glargine plus metformin with or without SU after 2 weeks, 12 weeks and 24 weeks of open-label treatment.
- To assess the change from baseline in the percentage of time spent within the euglycemic target range of ≥ 71 mg/dL (3.9 mmol/L) and ≤ 180 mg/dL (10.0 mmol/L) as measured by CGM with co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU versus titrated insulin glargine plus metformin with or without SU after 2 weeks, 12 weeks and 24 weeks of open-label treatment.
- To assess the change from baseline in the percentage of time spent with glucose ≤ 70 mg/dL (3.9 mmol/L) as measured by CGM with co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU versus titrated insulin glargine plus metformin with or without SU after 2 weeks, 12 weeks and 24 weeks of open-label treatment.
- To assess the change from baseline in the percentage of time spent with glucose ≤ 70 mg/dL (3.9 mmol/L) as measured by CGM between midnight (12am) and 6am with co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU versus titrated insulin glargine plus metformin with or without SU after 2 weeks, 12 weeks and 24 weeks of open-label treatment.
- To assess the mean change from baseline in HbA1c with co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU versus titrated insulin glargine plus metformin with or without SU after 52 weeks of open-label treatment.
- To assess the mean change from baseline in total body weight with co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU versus titrated insulin glargine plus metformin with or without SU after 52 weeks of open-label treatment.
- To assess the proportion of confirmed hypoglycemia [defined as: blood glucose ≤ 70 mg/dL (3.9 mmol/L) recorded in the hypoglycemia module of the CRF] with co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU versus titrated insulin glargine plus metformin with or without SU during 52 weeks of open-label treatment.

- To assess the proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0%, without any hypoglycemia, with co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU versus titrated insulin glargine plus metformin with or without SU after 52 weeks of open-label treatment.
- To assess the proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0%, with co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU versus titrated insulin glargine plus metformin with or without SU after 52 weeks of open-label treatment.
- To assess the proportion of subjects requiring rescue or discontinuation for lack of glycemic control with co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU versus titrated insulin glargine plus metformin with or without SU during 24 weeks and 52 weeks of open-label treatment.
- To assess the time to treatment intensification (addition of non study insulin or other anti-diabetic therapies for rescue therapy or discontinuation for lack of glycemic control) with co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU versus titrated insulin glargine plus metformin with or without SU during 24 weeks and 52 weeks of open-label treatment.
- To assess the proportion of subjects achieving HbA1c ≤ 6.5% with co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU versus titrated insulin glargine plus metformin with or without SU after 24 weeks and 52 weeks of open-label treatment.
- To assess the proportion of subjects achieving HbA1c ≤ 6.5% without any hypoglycemia with co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU versus titrated insulin glargine plus metformin with or without SU after 24 weeks and 52 weeks of open-label treatment.
- To assess change from baseline in average glucose values and postprandial glucose values measured by 6-point self-monitored blood glucose (SMBG) profiles with co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU versus titrated insulin glargine plus metformin with or without SU after 12 weeks, 24 weeks and 52 weeks of open-label treatment.
- To assess changes from baseline in subject-reported treatment satisfaction, quality of life, and barriers to medication adherence achieved with co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU versus titrated insulin glargine plus metformin with or without SU after 12 weeks, 24 weeks and 52 weeks of open-label treatment.

3.4 Safety

To assess the safety and tolerability of co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU versus titrated insulin glargine plus metformin with or without SU after 24 weeks and 52 weeks of open-label treatment.

4 ENDPOINTS

4.1 Primary Endpoint

Change from baseline in HbA1c at Week 24.

4.2 Secondary Endpoints

The secondary efficacy endpoints for the Short-term treatment Period include:

1) Change from baseline in total body weight at Week 24
2) Proportion of subjects with confirmed hypoglycemia [defined as: blood glucose ≤ 70 mg/dL (3.9 mmol/L) recorded in the hypoglycemia module of the CRF] at Week 24
3) Proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0%, without any reported hypoglycemia (for the duration of the Short-term Period) at Week 24
4) In a sub-study, change from baseline in the mean value of 24-hour glucose readings measured by Continuous Glucose Monitoring (CGM) at Week 2
5) Proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0% at Week 24

4.3 Exploratory Endpoints

1) In a sub-study, change from baseline in mean amplitude of glucose excursions (MAGE) of 24 hour glucose readings measured by Continuous Glucose Monitoring (CGM) at Week 2, Week 12, and Week 24 (Phase V® CGM Analytics variable name MAGE_SC see appendix 1).
2) In a sub-study, change from baseline in the mean value of 24 hour glucose readings measured by Continuous Glucose Monitoring (CGM) at Week 12 and Week 24 (Phase V® CGM Analytics variable name BGMEANALL see appendix 1).
3) In a sub-study, change from baseline within-subject, within-day in the standard deviation of 24 hour glucose readings measured by Continuous Glucose Monitoring (CGM) at Week 2, Week 12 and Week 24 (Phase V® CGM Analytics variable name WPWDSDDY see appendix 1).
4) In a sub-study, change from baseline in the percentage of time spent in the euglycemic target range of ≥ 71 mg/dL (3.9 mmol/L) and ≤ 180 mg/dL (10.0 mmol/L) as measured by CGM at Week 2, Week 12, and Week 24 (Phase V® CGM Analytics variable name CGM_GE71_LE180_PCTTIME see appendix 1).
5) In a sub-study, change from baseline in the percentage of time spent with glucose ≤ 70 mg/dL (3.9 mmol/L) as measured by CGM at Week 2, Week 12, and Week 24 (Phase V® CGM Analytics variable name CGM_HYPO_PCTTIME see appendix 1).
6) In a sub-study, change from baseline in the percentage of time spent with glucose ≤ 70 mg/dL (3.9 mmol/L) as measured by CGM between midnight (12am) and 6am at Week 2, Week 12, and Week 24 (Phase V® CGM Analytics variable name CGM NOC_HYPO_PCTTIME see appendix 1).
7) Change from baseline in HbA1c at Week 52.
8) Change from baseline in total body weight at Week 52
9) Proportion of subjects with confirmed hypoglycemia [defined as: blood glucose ≤ 70 mg/dL (3.9 mmol/L) recorded in the hypoglycemia module of the CRF] at Week 52

10) Proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0%, without any hypoglycemia, at Week 52
11) Proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0%, at Week 52
12) Proportion of subjects requiring rescue or discontinuation for lack of glycemic control) at Week 24 and Week 52
13) Time to treatment intensification (addition of non study insulin or other anti-diabetic therapies for rescue therapy or discontinuation for lack of glycemic control) at Week 24 and Week 52
14) Proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c ≤ 6.5%, at Week 24 and Week 52
15) Proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c ≤ 6.5%, without any hypoglycemia at Week 24 and Week 52
16) Change from baseline in average glucose values and postprandial glucose values measured by 6-point SMBG profiles at Week 12, Week 24, and Week 52
17) Change from baseline in patient-reported treatment satisfaction, quality of life, and barriers to medication adherence at Week 12, Week 24 and Week 52

4.4 Safety Endpoints

The safety endpoints include:

- The proportion of subjects with hypoglycemia events and the frequency and ADA classification of the hypoglycemia events during the Week 24 and Week 52 treatment periods
- The proportion of subjects experiencing adverse events (AE) and marked abnormalities in clinical laboratory tests during the Week 24 and Week 52 treatment periods
- The change from baseline at each post-baseline time point of assessment of selected safety clinical laboratory parameters, physical measurements, vital signs, and electrocardiogram data during the Week 24 and Week 52 treatment periods

5 SAMPLE SIZE AND POWER

The change from baseline in HbA1c at Week 24 will be assessed comparing co-administered saxagliptin 5 mg and dapagliflozin 10 mg plus metformin with or without SU versus titrated insulin glargine plus metformin with or without SU.

Power calculations for longitudinal repeated measures analyses depend on many factors, including the pattern of drop-outs over time and correlations among the various time points included in the model. The choice of these parameters will affect any estimates of power, and their true values may not be known. Based on comparisons of results of longitudinal repeated measures analyses and analysis of covariance (ANCOVA) using last observation carried forward (LOCF) from previous diabetes trials, the estimated standard errors of the treatment differences were similar between analyses. Therefore, power calculations are based on ANCOVA with LOCF, with the expectation that this will provide a good estimate of the power for the primary analysis using a longitudinal repeated measures model. To demonstrate non-inferiority of saxagliptin plus dapagliflozin to insulin for changes from baseline to Week 24 in HbA1c within a non-inferiority margin of 0.30%, assuming a standard deviation 1.1%, and at a one-sided significance level of 0.025, 284 evaluable subjects will be needed in each treatment group to provide approximately 90% power (given a true difference of zero between the 2 treatment

groups). Assuming that 5% of subjects do not have a post-baseline assessment, a total of approximately 598 subjects (299 subjects per treatment arm) need to be randomized. Assuming that 50% of screened subjects will fail to meet screening criteria, a total of 1196 subjects need to be screened.

6 STUDY PERIODS, TREATMENT REGIMENS AND POPULATIONS FOR ANALYSES

6.1 Study Periods

Study CV181-369 consists of 4 study periods:

1) **Screening period (Period A)**. This period starts with enrollment (i.e., signature of the protocol-specific informed consent form constitutes the first procedure of the Screening period) and ends on the start of lead-in period. Subjects should be on a stable dose of metformin 1500 mg/day for at least 8 weeks prior to screening visit with or without a stable dose of a SU of at least 50% the maximal dose per local label.
2) **Lead-in period (Period B)**. During this period, subjects will continue their diabetes management with stable dose of metformin 1500 mg/day with or without a stable dose of SU, diet and exercise. No placebo or study medication will be provided during the lead-in period.
3) **24-week short-term open-label treatment period (Period C)**. On the Day 1 visit, subjects who meet all the protocol-specific enrollment and randomization criteria will be randomized.
4) **28-week long-term open-label treatment period (Period D)**. After completing the short-term period, subjects will enter the 28-week open-label long-term treatment period. The start date of long-term treatment period is defined as the end date of short-term treatment period. This is the date when the subject entered the long-term treatment period and was dispensed long-term study medication.

Follow-up non-treatment phase (Period X). Subjects who discontinue study treatment early will follow the same visit schedule as subjects who remain on treatment.

6.2 Treatment Regimens

On the Day 1 visit, subjects who meet all protocol enrollment and randomization criteria will be randomized into one of the two open-label treatment arms, in a 1:1 ratio:

- Saxagliptin 5 mg QD and Dapagliflozin 10 mg QD
- Titrated insulin glargine

Randomization will be stratified by current use of background medication (metformin alone vs. metformin plus SU) at baseline to ensure equal representation across all treatment groups and also within the sub-study. Subjects will continue to receive their stable dose of metformin with or without SU throughout the study. The “as randomized” treatment group is defined as the treatment group to which a subject was randomized at the start of the open-label treatment period (even if the treatment they received was different). The primary efficacy analyses will be performed using the Randomized Subjects Data Set.

The “as treated” treatment group is the same as the “as randomized” treatment group, except in cases where information was available which indicated that a subject received a different treatment for the

entire course of their participation in the study (or period). In this case, the "as treated" treatment group is set to the treatment the subject actually received. In case a subject never received the treatment as assigned by randomization, then the "as treated" treatment group is the first treatment received.

6.3 Populations for Analyses

6.3.1 Enrolled Subjects Data Set

This consists of all subjects who signed informed consent.

6.3.2 Lead-in Subjects Data Set

The lead-in subject data set includes data collected from all subjects who have at least one vital sign measurement during Lead-in Period.

6.3.3 Randomized Subjects Data Set

The randomized subject data set will consist of all randomized subjects who receive at least one dose of study medication. Whenever using the randomized subject data set, subjects will be presented in the treatment group to which they were randomized at the start of the Short-term treatment Period (even if the treatment they received was different). This is also known as the Intent-to-Treat (ITT) population. This will be the primary efficacy data set.

6.3.4 Full Analysis Set

The full analysis set (FAS) is defined as all randomized subjects who take at least one dose of the study medication and have a baseline value for HbA1c. Analysis of the FAS will be based on the randomized treatment.

6.3.5 Evaluable Subjects Data Set

The Evaluable Subjects Data Set will be a subset of the Randomized Subjects, with all data points collected after a relevant protocol deviation excluded from the data set. Relevant protocol deviations are defined as deviations that could potentially affect the interpretability of the study results. This is also known as the Per-protocol population. All decisions to exclude patients and/or data from the Randomized Subjects Data Set will be made prior to un-blinding the study and agreed by the study team.

6.3.6 Treated Subjects Data Set

The Treated Subjects Data Set for ST and ST+LT will consist of all subjects who receive at least one dose of study medication during the short-term treatment period. This will be the primary safety data set. Data in this data set will be analyzed based on randomized treatment, except in cases where a subject received a different treatment for the entire course of his/her participation in the treatment period. In this case, safety data for such a subject will be analyzed based on the first treatment the subject actually received.

6.3.7 Short-term Completers Data Set

The short-term Completers Data Set will consist of all subjects in the Randomized Subjects Data Set who were not rescued, and completed ST treatment and entered the LT treatment period. It is a subset of the Randomized Subjects Data Set.

Whenever using the Short-term Completers Data Set, subjects will be presented in the treatment group to which they were randomized at the start of the ST treatment period (even if the treatment they received in the ST or LT was different.

7 STATISTICAL ANALYSES

7.1 General Methods

7.1.1 Definitions

7.1.1.1 *Baseline Value*

Unless otherwise stated, for each subject, baseline value of a parameter (e.g., efficacy laboratory parameter, safety laboratory test, ECG or physical measurement endpoint) is defined as the last assessment on or prior to the date of the first dose of the study medication.

7.1.1.2 Change and Percent Change from Baseline

Change from baseline to any Week *t* in short-term treatment period or ST +LT is defined as follows:

$$C_{\text{Week t}} = M_{\text{Week t}} - M_{\text{baseline}},$$

where:

- $C_{\text{Week t}}$ is the change from baseline at Week *t*,
- $M_{\text{Week t}}$ is the measurement at Week *t*,
- M_{baseline} is the measurement at baseline.

Percent change from baseline to any Week *t* in short-term treatment period is defined as follows:

$$P_{\text{Week t}} = 100 \times (M_{\text{Week t}} - M_{\text{baseline}}) / M_{\text{baseline}}.$$

Where $P_{\text{Week t}}$ is the percent change from baseline at Week *t*, and $M_{\text{Week t}}$ and M_{baseline} are defined as above.

The "Week *t*" to which a measurement belongs is determined using the conventions described in Section 8.2 and Section 8.4.

7.1.1.3 Handling of Missing Data

The main analysis for change from baseline specified for primary and secondary endpoints in following sections will use the repeated measures model. For subjects who started rescue medication, measurements taken prior to the date of the first dose of rescue medication will be used. This model assumes that the time course of the endpoint values for subjects who discontinue treatment or are rescued at a specified time point is consistent with the time course for subjects who are ongoing at that time point.

While utilizing models such as ANOVA, for example, in analysis of change (or percent change) from baseline as well as glycemic response endpoint at Week 24 LOCF, the measurement assigned as the Week 24 measurement will be used. If no Week 4 measurement is available (subject has

discontinued before Week 24, or measurement not taken at Week 24, though subject was not discontinued), then the last available earlier post- baseline measurement will be used (LOCF). For subjects who started rescue medication prior to Week 24, their last post-baseline measurement taken prior to the date of the first dose of rescue medication will be used.

Section 8 provides additional information regarding handling of missing or partial dates, inclusion of values, windowing, and values obtained post treatment.

7.1.2 Longitudinal Repeated Measures Analysis

A longitudinal repeated measures analysis using 'direct likelihood' will be performed. The SAS procedure PROC MIXED will be used. The preferred model will include the fixed categorical effects of treatment, week and treatment-by-week interaction as well as the continuous fixed covariates of baseline measurement and baseline measurement-by-week interaction.

The following model will be used:

$$C_{ijk} = \text{intercept} + \beta_1 [M_{baseline,ij}] + \tau_i + \omega_m + \alpha_k + (\alpha\ \tau)_{ik} + (\alpha\ M_{baseline})_{ijk} + error_{ijk}$$

(Model 7.1.2.1)

where

- C_{ijk} is the change from baseline for subject j in treatment group i at time k,
- β_1 is the slope coefficient for the baseline measurement,
- $M_{baseline,ij}$ is the baseline measurement of subject j in treatment group i,
- τ_i is the mean effect of treatment group i,
- ω_m is the mean effect of stratum m
- α_k is the mean effect at time k
- $(\alpha\ \tau)_{ik}$ is the interaction term between treatment group i and time k.
- $(\alpha\ M_{baseline})$ is baseline measurement-by-week interaction term for subject j in treatment group i at time k, and
- $error_{ijk}$ is the error term for subject j in treatment group i at time k.

An unstructured matrix for the within-subject error variance-covariance will be used. The denominator degrees of freedom will be calculated according to the Kenward-Roger method.

In case of non-convergence of the preferred model or memory space issues, the following back-up models are defined:

- The first backup model is the same as the preferred model but the Kenward-Roger method will be replaced by Satterthwaite approximation.
- The second backup model is the same as the preferred model but without the term for baseline measurement-by-week interaction.

The second back-up model will only be provided if the first back-up model does not converge or has memory issues.

The model will provide least-squares mean estimates, standard errors and 2-sided 95% confidence intervals for mean change at all time points within and between treatments.
Pre-specified stratification factors other than site in the randomization (e.g. pre-enrollment factor) will be added as an additional fixed effect in the mixed model.

Assessment of Treatment-by-Baseline Interaction:

Treatment-by-baseline interaction will be assessed for the analyses of the primary efficacy endpoint. In the repeated measures model (7.1.2.1), the interaction will be tested by including the additional terms for the treatment-by-baseline interaction.

The test for interaction will be performed at the 0.10 level of significance. If the treatment- by-baseline interaction is not significant, the original model (7.1.2.1) will be used. Otherwise, the interaction will be assessed as qualitative or quantitative. Assessment of the interaction type will be based on regression lines plotted for each treatment group. The intercepts and slopes for these regression plots will be obtained from the analysis model including the interaction term. The intercepts will be estimated by the least squares treatment means. The abscissa of the plots will range from the minimum baseline value from all subjects included in the analysis to the maximum baseline value.

If the regression lines do not cross, or the crossing is judged not severe (i.e., the crossing occurs near the boundary or beyond the range of baseline values), then the interaction will be considered quantitative and this does not compromise the validity of the treatment comparisons. In this case, the treatment comparisons will be made using the model without the interaction term.

Otherwise, the interaction will be considered qualitative and treatment comparisons will not be presented as a result of the complete model. In this case, the impact of the baseline value on treatment effect will be investigated by summarizing the data in subsets defined by baseline categories.

Subgroup Analyses and Assessment of Treatment-by-Subgroup Interaction:

Treatment effects will be assessed for subgroups based on age group, gender, race, region, and baseline HbA1c, as defined in Table 7.3.2A and Table 7.3.2B. The analyses will be based on model 7.1.2.1 with additional covariates of subgroups, treatment-by-subgroup, time-by-subgroup, and treatment-by-time-by-subgroup. Additionally, the treatment-by-stratification will be assessed. Tests of the treatment by subgroup interaction will be assessed using contrasts of the treatment effect by subgroups at Week 24. The model to assess the treatment-by-baseline HbA1c interaction will include baseline HbA1c as a continuous variable. Adjusted mean change from baseline and its difference from the control group will be given for each subgroup at week 24.

7.1.3 Analysis of Covariance (ANCOVA)

In summaries of efficacy endpoints examining ***changes from baseline at Week 24***, ANCOVA of the differences between post-baseline and baseline measurements will be performed, with treatment group as an effect and the baseline measurement as a covariate.
The following ANCOVA model will be used:

$$D_{t,ij} = intercept + \beta\,[Y_{0,ij}] + \tau_i + \omega_m + error_{ij}$$ Where (Model 7.1.3.1)

- $D_{t,ij} = Y_{24,ij} - Y_{0,ij}$ = the Week *24* change from baseline of subject *j* in treatment group *i* (as defined in Section 7.1.1.2 and Section 8 on conventions),
- $Y_{0,ij}$ is the baseline measurement of subject *j* in treatment group *i*,
- $Y_{24,ij}$ is the Week *24* measurement of subject *j* in treatment group *i*,
- β is the slope of $D_{t,ij}$ regressing on the baseline measurement and,
- τ_i is the mean effect of treatment group *i*.
- ω_m is the mean effect of stratum m, and
- *Intercept*, ® and τ_i are unknown parameters to be estimated from the data.

The model will provide least squares mean estimates and 2-sided 95% confidence intervals for mean changes from baseline within and (when warranted) differences in mean change from baseline between treatments. Where applicable, t-statistics corresponding to the Type III sums of squares for the differences in the least squares means will be used to obtain p-values for treatment group comparisons.

7.1.4 Descriptive Summaries of Continuous Variables

Descriptive summaries of continuous variables in terms of change or percent change from baseline values will be provided, including ns, means, medians, and standard error (SE). In addition, 95% confidence interval for the mean (percent) change from baseline will be calculated for continuous efficacy variables. They will be presented by treatment group and time point where applicable.

7.1.5 Descriptive Summaries of Categorical Variables

Descriptive summaries of categorical variables will consist of frequencies and percentages for each treatment group and overall, where applicable.

7.1.6 Descriptive Summaries of Change from Baseline in Categorical Variables

Descriptive summaries of change from baseline in categorical variables will be provided using shift tables. Frequencies and percentages of subjects within each treatment group will be generated for levels of cross-classifications of baseline and the on-treatment value of the parameter. The on-treatment value can either be the value at a certain time point, or e.g. for laboratory tests, the minimum/maximum value in the direction of toxicity, which has been observed during the short-term treatment period. Treatment group differences will not be assessed in summaries of shifts.

7.1.7 Kaplan-Meier Curve and Estimates for Time-to-Event Analyses

Kaplan-Meier plots[2] of time to event variables will be displayed by treatment group. Unless otherwise specified, the plot will be presented only when there are at least 5 events in one treatment group. Additionally, a table will accompany the plot and will display the Kaplan-

Meier estimates of the cumulative proportion (with 95% CI calculated based on Greenwood's method[3] when applicable) of subjects with event at specific time points by treatment group.

If the estimated lower bound of 95% CI is below 0 or the estimated upper bound of 95% CI is over 1, then it will be restricted to 0 or 1 respectively.

7.1.8 Proportion of Subjects Achieving Pre-defined Characteristic

Unless otherwise specified, the proportion of subjects *with a pre-defined characteristic* at Week *t* will be analyzed using logistic regression with adjustment for baseline HbA1c value and/or the stratification factor. Subjects in the randomized treatment group who do not demonstrate achieving therapeutic glycemic response at a specified timepoint or are rescued will count as treatment failures from that timepoint. When proportion of responders (e.g., meeting HbA1c criteria; composite endpoints) is needed, the estimates, confidence intervals, and tests will be obtained using this methodology with adjustment for baseline variable (e.g., adjustment for baseline HbA1c) and stratification factor. For each treatment group, the probability of response is first modeled using a logistic regression model with baseline (e.g., baseline HbA1c) and stratification factor as the covariates. Treatment group estimates of response rate are then obtained by integrating each group's modeled probability of response over the observed distribution of baseline covariate (combined across groups). For analysis of proportion of subjects achieving therapeutic glycemic response (HbA1c < 7%), the difference in response rate between the saxa-dapa arm and the insulin arm will be displayed along with standard error and 95% confidence intervals using the methodology of Zhang, Tsiatis and Davidian and Tsiatis, Davidian, Zhang and Lu. Non-Inferiority will be demonstrated if the upper bound of the 95% confidence interval is less than 10%.

When there are less than 5 responders in any treatment group, the unadjusted (and difference in) proportions, exact 95% confidence interval, and p-values from the Fisher's exact test (when applicable) will be provided.

7.1.9 Mean Daily Glucose and Mean Post-prandial Glucose (Self-monitored)

All subjects will perform a 6 Point SMBG profile for any 3 days (to be included in the calculation of mean daily glucose, the measurements do not need to be recorded over 3 consecutive days, nor do they need to have 3 complete days of measurements.). These are scheduled between Week (-2) to Week (-1), Week 11-12, Week 23-24 and Week 51-52. Blood glucose readings will consist of 3 pre-prandial measurements and 3 post-prandial measurements. Meals are considered to be breakfast, lunch and dinner. A minimum of 2 days of all 6-point SMBGs are required to complete the 6-point SMBG profile for each period. Similarly a minimum of 2 days of all post prandial SMBGs will be required for summary of postprandial averages. The pre-prandial and postprandial measurements collected at the nominal time points will be considered for analysis.

Glucose concentrations at each of the 6 time points will be averaged over the 3 days to derive the mean glucose concentrations at each of the 6 time points. The mean daily glucose (MDG) will then be calculated as the average over the average glucose concentrations at the 6 time points. Only complete pairs of the average preprandial and postprandial blood glucose values will be used for the calculation. For example, if the available average time point data for a subject at a visit are

pre-breakfast, post-breakfast, and pre-dinner, then the daily average will be calculated as (pre-breakfast + post-breakfast)/2. The average value for pre-dinner is excluded for the calculation since the average value for the post-dinner is missing. The mean post-prandial glucose will be calculated as the average over the average glucose concentrations at the 3 post-prandial time points.

7.1.10 Mean Amplitude of Glucose Excursions (MAGE) of continuous glucose monitoring system readings

Mean amplitude of glucose excursions (MAGE) is calculated using data from continuous glucose monitoring (CGM). MAGE for a 24-hour period equals "*the arithmetic mean of the blood glucose increases or decreases (from blood glucose nadirs to peaks or vice versa) when both ascending and descending segments exceeded the value of one standard deviation of the blood glucose for the same 24-hour period*"[1]. Therefore, it corresponds to the mean of absolute differences between consecutive maxima and minima, as long as these differences are greater than one standard deviation for the same 24-hour period.

Here an example of calculation as described in the original publication of MAGE in [3].

Figure 7.1.10-1: Blood glucose curve of a subject studied on intermediate-acting insulin regimen.

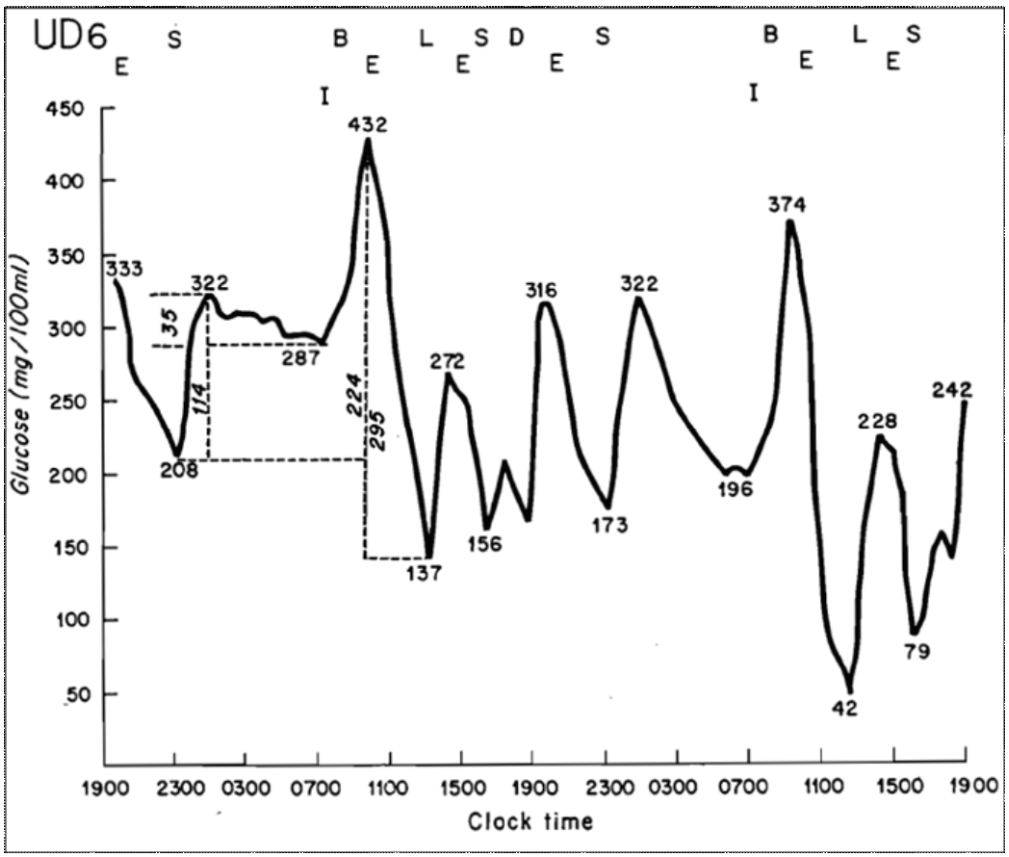


Figure from Service, et al., 1970 [3]

As stated in [3], the calculation of MAGE for subject in Figure 7.1.3-1"*is as follows: the first excursion, 333 to 208 mg./100 mL ., is 125 mg./100 mL., exceeding the value of one standard deviation; 62. The first blood glucose increase, from 208 to 322 mg./100 mL, is 114 mg/100 ml [...] and also exceeds one standard deviation; however, the subsequent blood glucose decrease of 35 mg./100 mL (from 322 to 287 mg./100 mL) is less than one standard deviation and so this excursion is not counted. The next blood glucose increase, from 208 to 432 (224 mg./100 mL), and decrease, from 432 to 137 mg./100 mL (295 mg./100 mL), both exceed the value of one standard deviation; therefore this glycemic excursion is 295 because the first excursion is in a peak-to-nadir direction and the direction of calculation (peak-to-nadir or nadir-to-peak) is established, for that CBGA, by the direction of the first excursion.*" Continuing, the next peak is 272 mg./100 mL which corresponds to an excursion of 135 mg./100 mL. The next excursion is 116 mg./100 mL to a nadir at 156 mg./100 mL, followed by an excursion of 160 mg./100 mL to a peak at 316 mg./100 mL. The next nadir is at 173 mg./100 mL corresponding to an excursion of 143 mg./100 mL and it is followed by a peak at 322 mg./100 mL and a respective excursion of 149 mg./100 mL and by a nadir at 196 mg./100 mL and a respective excursion of 126 mg./100 mL. The next excursion is 178 mg./100 mL to a peak at 374 mg./100 mL, followed by one of 332 mg./100 mL to a nadir at 42 mg./100 mL and one of 186 mg./100 mL to the subsequent peak of 228 mg./100 mL. Finally the last couple of excursions are 149 mg./100 mL to the next nadir at 79 mg./100 mL and 163 mg./100 mL to the final peak point at 242 mg./100 mL. Therefore, MAGE for this subject profile corresponds to the arithmetic mean of all excursions and equals approximately 177 mg./100 mL.

Note although the above example refers to glucose expressed in mg/dL, calculation of MAGE with glucose readings expressed in standard international units (i.e. mmol/L) will follow the same principle.

For calculating MAGE the following definitions apply:

- **Segment:** the ascending or descending part between a nadir and a peak.
- **Qualifying excursion:** A change from nadir to peak (or from peak to nadir) where both the excursion and the following segment in the opposite direction exceed 1 standard deviation for the whole 24 hour period for the subject.

MAGE is calculated as the arithmetic mean of all qualifying excursions during the 24 hours.

Phase V® is collecting the data and performing the analyses for CGM in this study. More details on the collection and analysis is provided in Appendix 1.

7.2 Study Conduct

Subjects who deviate from protocol conditions (e.g., important inclusion/exclusion criteria) will be reported as having significant protocol deviations. Significant protocol deviations that are determined to affect the primary efficacy results are deemed Relevant Protocol Deviations (RPDs). A list of relevant protocol deviation criteria, along with the consequent handling of data should those deviations occur, is given in Table 1. There will be no data exclusion for significant protocol deviations only.

Table 7.2-1: List of Relevant Protocol Deviations

RPD #	RPD Criteria	Exclusion Type	Exclusion Level
1	Randomized subjects without type 2 diabetes or with central laboratory HbA1c more than 0.2% outside of specified limits of ≥ 8.0% and ≤ 12%	Complete	This will be assessed using the screening value or at an unscheduled visit prior to randomization.
2	Randomized subjects who did not receive stable dose of metformin with or without a stable dose of SU	Complete	Subjects not taking ≥ 1500 mg metformin for at least 8 weeks prior to enrollment
3	Randomized subjects with abnormal TSH and free T4 missing or abnormal at enrollment or at an unscheduled visit prior to randomization	Complete	Because the T4 test is collected reactively to TSH, this deviation would only occur if there were an abnormal TSH and free T4 value
4	Randomized subjects with history of hemoglobinopathy, (with the exception of sickle cell trait (SA) or thalassemia minor), and/or chronic or recurrent hemolysis.	Complete	
5	Randomized subjects who used antihyperglycemic medication (other than protocol required medications) for 14 or more consecutive days during the short-term open-label treatment period.	Partial /Complete	Exclusion will start from the 14th consecutive day that the medication was taken/ Subjects who were on treatment < 14 days and used prohibited antihyperglycemic medication will be excluded
6	Randomized subjects that were treated with any systemic corticosteroid therapy for ≥ 5 consecutive days initiated during the short-term open-label treatment period.	Partial /Complete	Exclusion will start from the 5th day that the medication was taken. Subjects who were on treatment < 5 days and used prohibited corticosteroid medication will be excluded
7	a. Randomized subjects who took no dose of metformin or dose of metformin outside dose range for ≥ 2 consecutive weeks in the short-term open-label treatment period b. Randomized subjects who were on stable dose of SU at baseline but took no dose of SU or different SU dose for ≥ 2 consecutive weeks in the short-term open-label treatment period	Partial /Complete	Exclusion will start from the 14th consecutive day that metformin was either not taken or outside range of ≥ 1500 or dose of SU was changed for subjects who were taking SU. Subjects who were on treatment < 14 days and did not take metformin/SU according to protocol will be excluded
8	Randomized subjects who receive no study medication (saxagliptin, dapagliflozin, or insulin glargine) for ≥ 2 consecutive weeks in the short-term open-label treatment period	Partial /Complete	Exclusion will start from the 14th consecutive day that the medication was not taken in an interruption. Subjects who were on treatment < 14 days and did not receive study medication (saxagliptin,

Table 7.2-1: List of Relevant Protocol Deviations

RPD #	RPD Criteria	Exclusion Type	Exclusion Level
			dapagliflozin, or insulin glargine) will be excluded
9	Randomized subjects who received incorrect study medication (saxagliptin, dapagliflozin, or insulin glargine) or dosing for ≥ 2 consecutive weeks in the short-term open-label treatment period	Partial/Complete	Exclusion will start from the 14th consecutive day that the incorrect medication was taken. Subjects who were on treatment < 14 days and did not receive correct study medication (saxagliptin, dapagliflozin, or insulin glargine) or dosing will be excluded
10	Randomized subjects who are judged to be noncompliant in terms of overall compliance, i.e., who took less than 80% or more than 120% of their prescribed dose of study medication during the short-term open-label treatment period.	Complete	

* If a patient has an unacceptable/missing value at screening visit but then has an unscheduled visit prior to randomization with an acceptable value, the patient will not be considered an RPD.

7.3 Study Population

7.3.1 Subject Disposition

The disposition of subjects for the screening period, lead-in period, and the 24-week short-term treatment period will be summarized. The summary of status in the screening period will include all subjects in Enrolled Subjects Data Set. Reasons for discontinuation for subjects who discontinued from the lead-in period and from the short-term treatment period will be tabulated and listed. The summary will be presented by randomized treatment group and overall.

The summary of status in the combined short-term plus long-term treatment periods will include all subjects in Randomized Subjects Data Set and be presented by randomized treatment group. This summary will include subjects entering, completing and discontinuing the long-term treatment period with reasons for discontinuation.

7.3.2 Demography and Other Baseline Characteristics

Demographic and other baseline characteristics, including diabetes-related characteristics and renal function characteristics will be summarized overall and by treatment group (where applicable) using the Randomized Subjects Data Set and/or the Evaluable Subjects Data Set
Demographic and common baseline characteristics are listed in Table 7.3.2A. Common diabetes related baseline characteristics are listed in Table 7.3.2B. Common renal function baseline characteristics are listed in Table 7.3.2C.

Table 7.3.2A: Demographic and Common Baseline Characteristics

Characteristic	Summarized as	Categories
Gender	Categorical	Male, Female
Age	Categorical and Continuous	< 65 yrs ≥ 65yrs ≥65yryrs
Female Age	Categorical	≤ 50 yrs > 50 yrs
Race	Categorical	White, Black or African American, Asian, Other
Ethnicity	Categorical	Hispanic/Latino, Non Hispanic/Latino
Body Weight	Continuous	-
Height	Continuous	
Body Mass Index	Categorical and Continuous	< 25 kg/m2 ≥ 25 kg/m2 ≥ 27 kg/m2 ≥ 30 kg/m2
Geographic Region	Categorical	As defined in Appendix 4

Table 7.3.2B: Common Diabetes-Related Baseline Characteristics

Characteristic	Summarized as	Categories
Duration of Type 2 Diabetes	Categorical and Continuous	< 3 yrs ≥ 3 and ≤ 10 yrs > 10 yrs
HbA1c	Categorical and Continuous	<8% ≥ 8 and < 9% ≥ 9%
FPG	Continuous	-

Table 7.3.2C: Common Renal Function Baseline Characteristics

Characteristic	Summarized as	Categories
eGFR (MDRD)	Categorical and Continuous	<60 mL/min/1.73m^2 >=60-<90 mL/min/1.73m^2 >= 90 mL/min/1.73m^2

All summaries of continuous characteristics will be based on non-missing observations. For categorical characteristics, percentages will be calculated out of the total number of subjects in the data set, overall and by treatment group, where applicable (i.e., each denominator includes the number of subjects with missing/unknown values for the characteristic).

7.3.3 Specific and General Disease History

The number (percent) of subjects with diabetes, diabetes-related disease histories will be summarized by treatment group and overall using the Randomized Subjects Data Set for the short-term treatment period.

The number (percent) of subjects with general medical history findings will also be summarized by treatment group and overall using the Randomized Subjects Data Set for the short-term treatment period.

7.4 Extent of Exposure

7.4.1 Study and Rescue Medication

The extent of exposure to study medication during the 24-week short-term open-label treatment period is defined as the difference between the last and the first dose of study medication of the short-term open-label treatment period plus 1 day. The extent of exposure to study medication will be summarized using the Treated Subjects data set for the short-term open-label treatment period and the short-term open-label treatment period prior to rescue, where the number and percent of subjects with an extent of exposure within specified day ranges (1-6, 7-14, 15-28, 29-42, 43-56, 57-70, 71-84, 85-168, 169-182, >= 183 days) will be presented by treatment group. The mean (SD), median and range of the number of days of exposure will also be presented. Summaries will be presented excluding and including periods of interruptions (defined by record of 0 tablets of study medications on the case report form (CRF). Total daily insulin dose at Week 12 and Week 24 during the 24-week open-label treatment period will be summarized for the subjects taking Insulin. All rescue medication use during the 24-week short-term open-label treatment period will be summarized and listed by treatment group.

The extent of exposure to study medication during the combined ST+LT treatment periods is defined as the difference between the last dose of ST or LT treatment and the first dose of study medication plus 1 day. The extent of exposure to study medication will be summarized using the Treated Subjects data set for the combined ST+LT treatment periods, regardless of and prior to rescue. The number and percent of subjects with an extent of exposure within pre-specified day ranges (1-90, 91-180, 181-270, and > 270 days) will be presented by treatment group. The mean (SD), median and range of the number of days of exposure will also be presented. Summaries will be presented including periods of interruptions (defined by record of 0 tablets of study medications on the case report form (CRF)). All rescue medication use during the combined ST+LT treatment periods will be summarized and listed by treatment group.

A listing of subjects by batch number of study medication will also be generated for both ST and ST+LT treatment periods.

7.4.2 Concomitant Medications

Concomitant medications for the short-term treatment period will be any medication taken from start of the short-term treatment period up to the end of the short-term treatment period.

Concomitant medications for combined short-term plus long-term treatment periods will be any medication taken from start of the ST treatment period up to the end of the combined ST+LT treatment periods.

Concomitant medications for both the ST and ST+LT will be summarized using the Treated Subjects dataset by drug class and (generic) drug name. A summary table by drug class and generic drug name will be generated for each of the following:

- all concomitant medication
- all concomitant diuretic medication

- all concomitant ARB and/or ACE-I medication
- all concomitant anti-hypertensive medication

In addition, a listing of all non-study medications taken during the study (including prior and concomitant) will be produced.

Missing and partial date handling of start and stop dates of concomitant medications is described in Section 8.7. The WHO dictionary will be used to code the non-study medication.

7.4.3 Measurements of Treatment Compliance

Percent treatment compliance will be calculated for ST and ST+LT treatment period for each treatment group. For each subject, percent compliance for saxagliptin plus dapagliflozin treatment group is defined as the total number of tablets taken divided by the total number of tablets that should have been taken, multiplied by 100. Percent compliance for Insulin treatment group is calculated as number of days insulin dose was taken divided by the number of days insulin dose should have been taken (duration of exposure), multiplied by 100. A subject is considered compliant if percent compliance is $\geq$ 80% and $\leq$ 120%. The number and percent of subjects considered compliant will be summarized for each treatment group using the Treated Subjects data set.

For the Short-term:
The number of tablets that should have been taken is calculated as 1 + the number of days from the first short-term treatment period dose recorded in the "Record of Study Medication" to the last short-term treatment dose, times the prescribed daily dose (i.e. 2 tablets per day, one from each bottle). The number of tablets taken is the total number of tablets recorded (sum of bottle 1 and bottle 2) as taken based on the CRF, summed over the days counted when calculating the number of tablets that should have been taken, including the day of the last short-term treatment dose.

For the combined ST and LT:
The number of tablets that should have been taken is calculated as 1 + the number of days from the first treatment period dose recorded in the "Record of Study Medication" to the last treatment dose, times the prescribed daily dose (ie, 2 tablets per day, one from each bottle). The number of tablets taken is the total number of tablets recorded (sum of bottle 1 and bottle 2) as taken based on the CRF, summed over the days counted when calculating the number of tablets that should have been taken, including the day of the last treatment dose.

7.5 Efficacy

7.5.1 Overall Efficacy Summary

Efficacy analyses will be performed using the randomized subjects data set. For the non-inferiority endpoints, the analysis will also be performed in the evaluable subject data set. All analyses included in the hierarchical testing below, including the primary analysis, will use the randomized

subject data set. The analyses using the evaluable subject data set will not be included in the hierarchical testing.

All the main analyses will be conducted using values prior to rescue/intensification of treatment or treatment discontinuation. Values collected after this time will be excluded from analyses. Sensitivity analysis using data regardless of rescue or treatment discontinuation will also be performed for the primary efficacy endpoint to assess the robustness of the primary efficacy results.

The primary endpoint is the change in HbA1 from baseline at Week 24 visit. The primary endpoint will be tested for non-inferiority for saxagliptin plus dapagliflozin versus insulin at the alpha = 0.025 level (one sided), within a non-inferiority margin of 0.30%. If non-inferiority is demonstrated for the primary endpoint, the statistical tests for the secondary efficacy endpoints will be performed. The secondary endpoints then will be tested sequentially in the order that they appear below and in the objectives section of the protocol. Each comparison will be tested at the alpha = 0.05 (two-sided) level. Statistical tests between the dapagliflozin plus saxagliptin group and insulin group will be only performed for a given secondary endpoint if all previous sequential tests are statistically significant. Otherwise, the testing procedure will stop at the secondary endpoint that does not reach statistical significance. These secondary endpoints are all testing superiority, with the exception of the fourth and fifth bullets below which are testing non-inferiority.

- Change from baseline in total body weight at Week 24 (tested for superiority)
- Proportion of subjects with confirmed hypoglycemia [defined as: blood glucose ≤ 70 mg/dL (3.9 mmol/L) recorded in the hypoglycemia module of the CRF] at Week 24 (testing superiority)
- Proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0%, without any reported hypoglycemia (for the duration of the Short-term Period) at Week 24 (testing superiority)
- In a sub-study, change from baseline in the mean value of 24-hour glucose readings measured by Continuous Glucose Monitoring (CGM) at Week 2 (testing non-inferiority, using a non-inferiority margin of 12 mg/dL)
- Proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0% at Week 24 (testing non-inferiority, using a non-inferiority margin of 10%)

Analyses of secondary and exploratory endpoints specifically designed for sub-studies (CGM) will be performed using the subsets of the Randomized Subjects data set for the sub-studies.

7.5.2 Primary Efficacy Analysis

The primary endpoint is the change in HbA1 from baseline at Week 24 visit. The primary endpoint will be tested for non-inferiority for saxagliptin plus dapagliflozin versus insulin at the alpha = 0.025 level (one sided). The primary analysis of the change in HbA1c from baseline at Week 24 visit will be based on a longitudinal repeated measures analysis using 'direct likelihood'. The model will use subjects in the primary efficacy data set (i.e., randomized subjects data set) who have a baseline assessment and at least one post-baseline open-label

treatment period assessment. The SAS procedure PROC MIXED will be used. The preferred model will include the fixed categorical effects of treatment, week, randomization stratification factor (background medication of metformin alone vs. metformin + SU) and treatment-by-week interaction as well as the continuous fixed covariates of baseline measurement and baseline measurement-by-week interaction (see Model 7.1.2.1 in Section 7.1.2). An unstructured matrix for the within-subject error variance-covariance will be used. The denominator degrees of freedom will be calculated according to the Kenward-Roger method. A number of back-up models are defined in the Core statistical analysis plan (Section 7.1.2) in case of non-convergence of the preferred model or other issues. Data collected outside the analysis window (defined in Section 8 of this SAP) after subjects discontinued study medication will also be excluded from the primary analysis. Point estimates and 95% confidence intervals for the mean change within each treatment group as well as the difference in mean change between the saxagliptin plus dapagliflozin versus insulin will be calculated. If the upper limit of the 95% confidence interval of the difference is < 0.3%, then saxagliptin plus dapagliflozin as add-on therapy to metformin with or without SU will be considered non-inferior to insulin as add-on therapy to metformin with or without SU.

Sensitivity Analyses

The following sensitivity analyses will be carried out for the primary end point:

- Primary analysis will be repeated using Evaluable Subjects data set.
- Primary analysis will be repeated using all available values regardless of rescue or treatment discontinuation
- An ANCOVA analysis using values prior to rescue or treatment discontinuation (LOCF will be used if the week 24 value is not available)
- Analyses to address missing values (described below in section 7.5.3)
  - Analysis using multiple imputation return-to-baseline
  - Tipping point analysis

7.5.3 Handling of missing efficacy data

Missing data in this study may result from patients discontinuing from the study prematurely or missing intermediate visits or selected assessments while remaining on study. Every reasonable effort will be made to obtain the protocol-required data for all study assessments that are scheduled for all patients who have been enrolled. For efficacy analyses, missing observations will not be imputed, except those inherited from the mixed model repeated measures (MMRM) which assumes that data are missing at random (MAR).

MAR refers to missingness that is independent of missing responses, conditionally on observed responses and covariates. As the imputation strategy should always consider the dropout patterns and the time-course of the efficacy measurements by treatment, the pre- and post-withdrawal values will be assessed to understand the impact of dropouts on the efficacy results.

The primary efficacy endpoint of HbA1c data will be visually examined to explore the missingness patterns by (1) plotting each individual patient's HbA1c change trajectory in completers, side-by-side with those from the early study drug withdrawal, and by (2) plotting the early study drug withdrawal's last HbA1c change data, overlaid with the box plot by visit from completer population. Two additional methods of sensitivity analyses will be performed to compare the results from an MAR-based analysis versus an MNAR–based analysis under several MNAR scenarios. The first method is the return-to-baseline using Multiple Imputation in which patients with missing data known or believed to have discontinued protocol therapy were assumed to have a washout ("return to baseline") of any potential treatment effect. More specifically for each imputation of the multiple imputation, for those subject with missing Week 24 measurements for HbA1c. The second method is the Tipping Point Analysis, which assumes that patients from the experimental treatment arm who discontinue treatment or initiate a rescue therapy would have values worse by some amount "delta" compared to efficacy values of similar patients who continue with study treatment and do not require rescue. The detailed analysis plan and implementation of these methods are described in section 7.5.3.1, and 7.5.3.2.

The analysis for the treatment difference will be based on an ANCOVA model with change from baseline to week 24 as the dependent variable and covariate/factors of treatment group, randomization strata, and baseline HbA1c.

7.5.3.1 MMRM Model for Change from Baseline Based on Return-to-Baseline Multiple Imputation

The return-to-baseline MI (Multiple Imputation) imputes the change from baseline to Week 24, and samples will be drawn from a Normal distribution with mean 0 and variance of pooled data. Let $X = (X_{obs}, X_{miss})$ be the complete data at Week 24. X is consisted of observed measurements X_{obs} and the missing observations X_{miss}. In return-to-baseline imputation, when X is change from baseline (CHG), each missing observation X_{miss} is imputed by a random draw from a Normal distribution with mean 0 and variance v_{imp}:

$$X_{mis} \sim N(0, v_{imp}),$$

The variance v_{imp} is calculated from the observed changes:

$$v_{imp} = (1 + 1/N_c)\, v_c,$$

where v_c is the variance of the change among completers across all treatment arms, and N_c is the number of completers.

An ANCOVA model will be used to analyze the imputed datasets for change from baseline with treatment group and randomization strata as fixed factors and Baseline HbA1c as a covariate. The least square mean (LS mean) estimates will be combined using Rubin's combination rules for statistical inference. All available measurements will be used including observations post-rescue and post-discontinuation of treatment.

In this study, the primary analysis is comparing Saxagliptin + Dapagliflozin + Metformin group to control group Insulin Glargine + Metformin. The above approach will be conducted using the Insulin Glargine + Metformin as the control group.

7.5.3.2 Tipping Point Analysis based on ANCOVA Model for Change from Baseline Using Multiple Imputation

In order to address the impact of missing data due to initiation of rescue therapy, or premature treatment discontinuation on the primary efficacy analysis, a Tipping Point Analysis using multiple imputation will be performed to compare the results from analysis assuming a missing at random (MAR) mechanism versus analysis assuming a missing not at random (MNAR) mechanism. The specific MNAR assumption that will be considered in this framework is that patients from the experimental treatment arm who discontinue study treatment prematurely or who initiate a rescue therapy would have, on average, their efficacy values post rescue/post treatment discontinuation worse by some amount delta compared to efficacy values of similar patients who continue with the study treatment and do not require rescue therapy. Delta is considered a sensitivity parameter representing a degree of departure from the MAR assumption. The aim of the Tipping Point Analysis is to find a "tipping point" corresponding to a value of delta where the study conclusion of a non-inferior treatment effect would no longer hold. An interpretation of clinical plausibility of the assumption underlying the tipping point will be provided.

The tipping point approach based on multiple imputation of values at time points after treatment discontinuation or initiation of rescue can be performed with a specified adjustment (referred to as delta adjustment or shift) applied to values imputed under an MAR-based imputation model for the appropriate subset of patients. In order to find a tipping point, a series of imputations will be performed with increasing values of delta.

For the primary endpoint which is a continuous variable, we will use an additive delta adjustment (a shift) as follows:

$$Y_{j(adj)}^{(m)} = Y_{j(imp)}^{(m)} + \delta$$

Where -

- $Y_{j(imp)}^{\{m)}$ are values imputed using a MAR-based imputation model in the mth imputed dataset, m=1,…,M (number of imputations)
- δ is a mean shift (delta adjustment) parameter for adjusting imputed values

The main steps in the implementation of the Tipping Point Analysis are described below.

Step 1: Investigate missing data patterns in the source data (containing observed values prior to initiation of rescue therapy or premature study treatment discontinuation). If missing data have both monotone and non-monotone patterns, use a multivariate regression imputation model and the Markov chain Monte Carlo (MCMC) method to partially impute non-monotone data under the MAR assumption. The variables used as explanatory variables for imputation include Randomization strata, HbA1c baseline, and post-baseline HbA1c at each time point. The imputations will be done separately within each treatment arm. If there are not enough observations to properly estimate covariance parameters, then stratification variables could be removed from the MCMC step, and/or the treatment arm can be included as an explanatory variable in the model.

As a result, each imputed dataset will only have a monotone missing data pattern.

Step 2: Impute remaining monotone missing data using an MAR-based regression imputation model for all patients who discontinued study treatment prematurely or initiated a rescue therapy. Apply a shift (an additive delta adjustment) to imputed values of patients in the experimental treatment arm (Saxagliptin + Dapagliflozin + Metformin). The variables used as explanatory variables for imputation include Treatment, Randomization strata, HbA1c baseline, and post-baseline HbA1c at each time point.

Imputations with delta adjustment described above will be performed with varying values of delta in order to perform a series of analyses with progressively larger values of delta until a tipping point is reached. A tipping point will correspond to the smallest value of delta for which the primary hypothesis (non-inferiority) is no longer held (the upper bound of 95% confidence interval of the difference is greater than 0.3).

Step 3: At each level of delta, analyze each of multiple imputed datasets using the same ANCOVA model as used for the primary analysis. Combine estimates obtained from multiple imputed datasets based on Rubin's combination rules.

Step 4: Using 95% upper CI, find the tipping point, ie, the value of delta parameter for which the primary non-inferiority hypothesis is no longer held. In this study, we will compare the non-inferiority of Saxagliptin + Dapagliflozin + Metformin vs. Insulin Glargine + Metformin for each Delta value. We will find the delta value at which the test of non-inferiority hypothesis is no longer held. A clinical interpretation about the plausibility of the assumptions underlying the tipping point will be provided.

7.5.4 Secondary Efficacy Analyses

If non-inferiority is demonstrated for the primary endpoint, the statistical tests for the secondary efficacy endpoints will be performed. The secondary endpoints then will be tested sequentially in the order that they appear in the objectives section of the protocol. Each comparison will be tested at the alpha = 0.05 (two-sided) level for superiority. Statistical tests between the dapagliflozin

plus saxagliptin group and insulin group will be only performed for a given secondary endpoint if all previous sequential tests are statistically significant. Otherwise, the testing procedure will stop at the secondary endpoint that does not reach statistical significance.

The continuous secondary endpoints (i.e., the change from baseline in total body weight, the change from baseline at Week 24 visit in the mean value of 24-hour glucose readings obtained from CGM) will be analyzed using a longitudinal repeated measures analysis, similarly to the one used for primary efficacy analysis. The proportion of subjects achieving HbA1c < 7.0% will be analyzed using the methodology of Zhang, Tsiatis, and Davidian and Tsiatis, Davidian, Zhang, and Lu with adjustment for baseline HbA1c value and/or the stratification factor. The other binary endpoints (hypoglycemia endpoint and composite response endpoints of glycemic control with hypoglycemia) will be analyzed using logistic regression with adjustment for baseline HbA1c value and/or the stratification factor. In addition to point estimates and 95% confidence intervals, p-values will be calculated for all secondary endpoints. However, no claim will be based on endpoints for which the statistical testing is not performed for the endpoint as per the testing strategy as described above. A clear distinction will be made between p-values whereby claims can and cannot be made. All secondary efficacy analyses will use subjects in the primary efficacy data set (i.e., randomized subjects data set) who have a baseline assessment and any post-baseline open-label treatment period assessment.

For the two secondary endpoints which are tested for non-inferiority, the analysis will be repeated using evaluable subjects data set to examine the robustness of the results. The non-inferiority margins for the two secondary endpoints testing for non inferiority are specified in Table 7.5.4-1.

Table 7.5.4-1: Non-inferiority margins specified for primary/secondary endpoints

Secondary Endpoint	Non-inferiority Margin
Change from baseline in HbA1c at Week 24	0.30%
Change from baseline in mean 24-hour glucose readings measured by Continuous Glucose Monitoring (CGM) at Week 2	12 mg/dL
Proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0% at Week 24	10%

7.5.5 Exploratory Efficacy Analyses

None of the exploratory endpoints will be statistically tested.

7.5.5.1 Mean change from baseline at Week t

The analyses of change from baseline in exploratory endpoints will be performed using the longitudinal repeated measures model (Model 7.1.2.1, including the stratification factor) as described in Section 7.1.2.

7.5.5.2 Glycemic Rescue or Discontinuation from Study treatment due to Lack of Efficacy

The proportion of subjects requiring treatment intensification (glycemic rescue or discontinue study treatment for lack of efficacy) up to Week 24 / 52 will be assessed by summarizing the difference in the percentages between treatment groups and by a time-to-event analysis using the Kaplan-Meier methodology.

Time to treatment intensification will be analyzed using a Cox proportional hazards model. Estimates of hazard ratio and 95% confidence intervals will be provided. Kaplan-Meier plots[1] of time to intensification at Weeks 24 and 52 will be calculated and plotted by treatment group. Unless otherwise specified, the plot will be presented only when there are at least 10 events in one treatment group. Additionally, a table will accompany the plot and will display the Kaplan-Meier estimates of the cumulative proportion (with 95% CI calculated based on Greenwood's method[2] when applicable) of subjects with event at specific time points by treatment group. If the estimated lower bound of 95% CI is below 0 or the estimated upper bound of 95% CI is over 1, then it will be restricted to 0 or 1 respectively. For the analysis during 24 / 52 weeks, all subjects will be censored at Week 24 / 52 visit (end of ST treatment/end of LT treatment) if treatment intensification has not occurred by then. Subjects rescued at Week 24 / 52 will be counted as having an event for the analysis.

Composite response endpoints (glycemic control without hypoglycemia), will be analyzed using logistic regression with adjustment for baseline HbA1c value and/or the stratification factor. None of the exploratory endpoints will be statistically tested

In addition, an analysis of the proportion of subjects not rescued and achieving glycemic response (defined as HbA1c < 7%) in subjects from the Short-term Completers Data Set at time points during the LT period will be performed using the same methodology described above (with adjustment for end of ST HbA1c measurement). Subjects discontinued or missed measurements (regardless of the type of missing) at the specific time point will be considered as not achieving glycemic response.

7.5.5.3 Summary of change from baseline for lipids

The change from baseline in lipids data during the short-term and long-term treatment periods will be summarized for each treatment group.

7.5.6 Summary of Primary Efficacy Endpoint within Subgroups

Treatment effects will be assessed for subgroups based on age group, gender, race, region, and baseline HbA1c, as defined in Table 7.3.2A and Table 7.3.2B. The analyses will be based on model 7.1.2.1 with additional covariates of subgroups, treatment-by-subgroup, visit-by-subgroup, and treatment-by-visit-by-subgroup. Additionally the treatment-by-stratification interaction will be assessed. The nominal p-value for subgroup-by-treatment interaction will be presented. The p-value for the test of treatment by age, female age, baseline HbA1c, and baseline eGFR interaction will include corresponding continuous subgroup variables in the model.

Adjusted mean change from baseline and its difference from the control group will be given for each subgroup at week 24.

7.5.7 Pharmacokinetic Analyses

Not applicable

7.5.8 Biomarker Analyses

Not applicable.

7.5.9 Outcomes Research Analyses

The Phase V® Health Outcomes Information System Diabetes Module will be used for the patient-reported outcomes (PRO) assessments. The self-administered PRO questionnaires consist of validated generic and diabetes-specific modules of treatment satisfaction, quality of life, barriers to medication adherence and weight perception. Mean change from baseline in these parameters will be analyzed using longitudinal repeated measures model with terms for treatment group, baseline value, randomization stratification factor (back ground medication of metformin alone vs. metformin + SU), time, the interaction of treatment and time, and the interaction of baseline value and time. Point estimates and 95% confidence intervals will be calculated for the adjusted mean changes within each treatment group as well as for the differences in adjusted mean changes between treatment groups.

Phase V® is collecting the data and analyzing the data for patient reported outcomes in this study. More details on the collection and analysis is provided in Appendix 2.

7.5.10 Other Analysis

Not applicable.

7.5.11 Interim Analyses

There will be produced one study report at the end of Short-term Period and a second study report at the end of Short-term plus Long-term period.

7.6 Safety

The assessment of safety will be based on the analyses of AEs, vital signs, physical examinations, ECGs, hypoglycaemia, and clinical laboratory evaluations. All safety analyses will be performed using treated subjects data set.

The Treated Subjects Data Set will be used for all safety analyses, including data after rescue. Safety analyses will be performed using data from the 24-week open label period (at the time of the primary endpoint analysis) and then again at the final database lock (52 weeks) on the combined ST+LT treatment period. Sensitivity analyses on data collected prior to rescue will be performed for selected Adverse Events (AEs) as described in Section 7.6 of the ST Core SAP and Section 7.5 of the ST+LT Core SAP.

7.6.1 Adverse Events

.Adverse events analyses will be performed for the ST and ST+LT periods.

Adverse Events (AEs) will be classified by Primary System Organ Class (SOC) and Preferred Term (PT) according to the Medical Dictionary for Regulatory Activities (MedDRA). Summaries of AEs will use the version of MedDRA that is current at the time of database lock for each study. Counting rules for adverse events are described in Section 8.7.

In summaries by SOC and PT, AEs will be sorted by decreasing frequency within each SOC and PT. In summaries by PT, AEs will be sorted by decreasing frequency within PT according to the highest of either of the combinational dose group across the study.
Separate pages to capture events of hypoglycemia are contained within the CRF. Hypoglycemia or discontinuation due to hypoglycemia would not be reported on an AE CRF page unless the event fulfilled criteria for a Serious AE (SAE) in which case an SAE form would be completed. Hypoglycemia events that are reported as SAEs will be included in all summaries of AEs or SAEs (see Section 7.6.1.1). Separate summaries will be provided including hypoglycemia events reported on that special CRF pages.

7.6.1.1 All Adverse Events

An overall summary of adverse events at subject level, including AEs, SAEs, death, hypoglycemia, treatment-related events and events leading to the discontinuation of study medication will be performed for the short-term treatment periods. All adverse events (serious and non-serious, excluding hypoglycemic events that are not reported as SAEs) will be summarized by system organ class, preferred term for the short-term treatment periods. For the analyses of the short-term treatment period, the summary of AEs and analyses by SOC and PT will be performed for the primary and sensitivity safety analyses. In addition, a subject listing of all reported adverse events will be produced, displaying all events (including pre-treatment events) that occurred prior to the start date of long-term treatment period, if any. All adverse events (serious and non-serious) including all hypoglycemic events will also be summarized by treatment group, where applicable.

AEs and SAEs with an onset from Day 1 of short-term treatment up to and including 4 days and 30 days respectively, after the last dose date in the short-term treatment period (or up and including to the start date of the long-term treatment period whichever comes first) will be considered as occurring during the short-term treatment period.

In addition, the following summaries will be provided for the short-term treatment period (excluding hypoglycemic events that are not reported as SAEs):

- Most common adverse events by preferred term and treatment group (i.e., reported by ≥ 2 % of subjects in any treatment group),
- Adverse events by system organ class, preferred term, intensity and treatment group,

- Adverse events related to study medication by system organ class, preferred term and treatment group.
- Proportion of subjects with adverse events by SOC and PT in subgroups of subjects defined by age category (< 65 and ≥ 65 yrs), gender, race, and female age category (≤ 50 and > 50 yrs.

In addition to analyses of event incidence at the subject level, recurrence analyses will be performed at the event level. In order to prepare these summaries, the CRF data will be processed to categorize each line of patient data as a new occurrence or a continuation of an existing event. This determination will be based upon onset and resolution dates. Each line of patient data will represent the maximum severity observed at the time of occurrence as well as the last known assessed relationship to study medication by the investigator.

The following summary information will be provided:

Total number and rate (exposure adjusted) of occurrences for all AEs occurring in at least 5% of the subjects treated.

Additionally, a listing will be provided displaying the unique instances of all AEs, i.e., after duplicates have been eliminated and overlapping and contiguous occurrences of the same event have been collapsed. Because of how hypoglycemic events are assessed and captured in the database, each reported event is assumed to be unique and multiple hypoglycemic events will not be collapsed.

Similar analyses of adverse events will be performed for ST+LT treatment period. AEs and SAEs with an onset from Day 1 of short-term treatment up to and including 4 days and 30 days respectively, after the last dose date in the ST or LT treatment period will be included in summaries of the combined ST+LT treatment periods.
No formal comparisons will be made between treatments. No formal statistical testing will be performed, only summary statistics will be provided.

7.6.2 Deaths

All deaths recorded on the status page, the AE page, or SAE page (with a death date, cause of death, outcome or SAE categorization present) of the CRF will be considered a death in the analyses. Any deaths that occur during the study will be described in depth as narrative in the CSR. A listing of all deaths that occur prior to the start date of long-term treatment period will be produced for ST and listing of all deaths that occur during the study will be produced for ST+LT.

7.6.3 Serious Adverse Events

All SAEs (including hypoglycemic events) will be described in narratives, regardless of investigator assessment of causality.

SAEs with an onset from Day 1 of short-term treatment up to and including 30 days after the last dose date in the short-term treatment period (or up to the start date of the long-term treatment period whichever comes first) will be considered as occurring during the short-term treatment period.
SAEs with an onset from Day 1 of ST treatment up to and including 30 days after the last dose date in the ST or LT treatment period will be considered as occurring during the combined ST+LT treatment periods.

SAEs occurring during the ST/ST+LT treatment period will be summarized by system organ class, preferred term and treatment group for both the primary and sensitivity safety analyses. In addition, the proportion of subjects with related SAEs will be presented by system organ class, preferred term and treatment group.

A listing of all SAEs for ST will be produced, displaying all SAEs (including pre-treatment events) that occurred prior to the start date of long-term treatment period.
A listing of all SAEs for ST+LT will be produced, displaying all SAEs (including pre-treatment events) that occurred during the study.

7.6.4 Adverse Events Leading to Discontinuation of Study Medication

AEs with an onset during the short-term treatment period reported with an action taken of discontinuation of study medication will be summarized by system organ class, preferred term and treatment group for both the primary and sensitivity safety analyses. This summary will include hypoglycemia events that reported as SAEs. AEs leading to discontinuation with an onset date on or prior to the start date of the long-term treatment period will be summarized. In addition, a subject listing of discontinuations due to AEs will be provided, displaying all events that led to discontinuation with an onset date prior to the start of long-term treatment period.
AEs with an onset during the combined ST+LT treatment periods reported with an action taken of discontinuation of study medication will be summarized by SOC, PT, and treatment group for the primary safety analysis. This summary will include hypoglycemia events that reported as SAEs. When summarizing AEs leading to discontinuation for ST+LT, no upper cutoff day windows (ie, 4 days and 30 days from last dosing date for AEs and SAEs respectively) will be applied. In addition, a subject listing of discontinuations due to AEs will be provided, displaying all events that led to discontinuation.

7.6.5 Adverse Events of Special Interest (AEOSI)

Separate summaries will be provided for the following adverse events of special interest (AEOSI). Except as otherwise noted, to identify each type of adverse event of special interest in this section, a list of PTs will be selected before database lock and unblinding of the database.
AEs and SAEs of special interest with an onset from Day 1 of short-term treatment up to and including 4 days and 30 days respectively, after the last dose date in the short-term treatment period (or up to the start date of the long-term treatment period whichever comes first) will be considered as occurring during the short-term treatment period.

AEs and SAEs of special interest with an onset from Day 1 of ST treatment up to and including 4 days and 30 days respectively, after the last dose date in the ST or LT treatment period will be considered as occurring during the combined ST+LT treatment periods.

7.6.5.1 Hypoglycemia

Separate pages to capture events of hypoglycemia are contained within the CRF.

Hypoglycemic events with an onset from Day 1 of the short-term open-label treatment up to and including 4 days after the last dose date in the short-term open-label treatment period (or up to and including the start date of the long-term treatment period, whichever comes first) will be considered as occurring during the Short-term open-label treatment period. Hypoglycemic events with an onset from Day 1 of the short-term open-label treatment up to and including 4 days after the last dose date in the long-term treatment period will be considered as occurring during the Short-term plus long-term treatment period.
Hypoglycemic events will be categorized using two methods. One classification is based on the ADA recommendations[2]:

Severe hypoglycemia: "An event requiring assistance of another person to actively administer carbohydrate, glucagons, or other resuscitative actions. These episodes may be associated with sufficient neuroglycopenia to induce seizure or coma. Plasma glucose measurements may not be available during such an event, but neurological recovery attributable to the restoration of plasma glucose to normal is considered sufficient evidence that the event was induced by a low plasma glucose concentration." **Error! Bookmark not defined.**
Documented symptomatic hypoglycemia:"An event during which typical symptoms of hypoglycemia are accompanied by a measured plasma glucose concentration ≤ 70mg/dl (3.9mmol/l)." **Error! Bookmark not defined.**
Asymptomatic hypoglycemia: "An event not accompanied by typical symptoms of hypoglycemia but with a measured plasma glucose concentration ≤ 70mg/dl (3.9mmol/l). Since the glycemic threshold for activation of glucagon and epinephrine secretion as glucose levels decline is normally 65–70mg/dl (3.6–3.9mmol/l) (24–26) and since antecedent plasma glucose concentrations of ≤ 70mg/dl (3.9mmol/l) reduce sympathoadrenal responses to subsequent hypoglycemia (1,11,20), this criterion sets the lower limit for the variation in plasma glucose in nondiabetic, nonpregnant individuals as the conservative lower limit for individuals with diabetes." **Error! Bookmark not defined.**
Probable symptomatic hypoglycemia: "An event during which symptoms of hypoglycemia are not accompanied by a plasma glucose determination (but that was presumably caused by a plasma glucose concentration ≤ 70mg/dl [3.9mmol/l]). Since many people with diabetes choose to treat symptoms with oral carbohydrate without a test of plasma glucose, it is important to recognize these events as "probable" hypoglycemia. Such self-reported episodes that are not confirmed by a contemporaneous low plasma glucose determination may not be suitable outcome measures for clinical studies that are aimed at evaluating therapy, but they should be reported." **Error! Bookmark not defined.**

Relative hypoglycemia: "An event during which the person with diabetes reports any of the typical symptoms of hypoglycemia, and interprets those as indicative of hypoglycemia, but with a measured plasma glucose concentration > 70mg/dl (3.9mmol/l). This category reflects the fact that subjects with chronically poor glycemic control can experience symptoms of hypoglycemia at plasma glucose levels > 70mg/dl (3.9mmol/l) as plasma glucose concentrations decline toward that level (27,28). Though causing distress and interfering with the patient's sense of well-being, and potentially limiting the achievement of optimal glycemic control, such episodes probably pose no direct harm and therefore may not be a suitable outcome measure for clinical studies that are aimed at evaluating therapy, but they should be reported." **Error! Bookmark not defined.**

The second classification of hypoglycemic events follows:

Major episodes of hypoglycemia - defined as symptomatic episodes requiring external (3rd party) assistance due to severe impairment in consciousness or behavior with a capillary or plasma glucose value < 3 mmol/L (< 54 mg/dL) and prompt recovery after glucose or glucagon administration.

Minor episodes of hypoglycemia - defined as either a symptomatic episode with a capillary or plasma glucose measurement below 3.5 mmol/L (63 mg/dL) regardless of need for third-party assistance or an asymptomatic capillary or plasma glucose measurement below 3.5 mmol/L (63 mg/dL), that does not qualify as a major episode.

Other episodes of hypoglycemia - defined as episodes reported by the investigator that are suggestive of hypoglycemia but do not meet the above criteria.

All analyses of hypoglycemic events will be performed both overall and by each class of events both including and excluding rescue data, with the ADA classification system and excluding rescue data analysis serving as the primary analysis. Data collected in the hypoglycemia CRF modules of the CRF will be included in the analysis. All blood glucose measurements collected in the hypoglycemia module of CRF will be considered for classifying the hypoglycemia events irrespective of the method using which glucose data were collected.

The total number of events by treatment group will be summarized for the short-term open-label treatment period. Hypoglycemic events with an onset during the short-term open-label treatment period and leading to discontinuation of study medication will be summarized by treatment group. When summarizing hypoglycemic events leading to discontinuation no upper cutoff day windows are applied. For short-term open-label period analyses the only upper cutoff date is the start date of the long-term treatment period.

A summary of the incidence of confirmed hypoglycemia, defined as blood glucose value ≤ 70mg/dL as recorded in the hypoglycemia module of CRF will be provided.

A listing of subjects will be produced and it will display all hypoglycemic events with an onset date/time from the start date/time of short-term open-label treatment period up to the start date of the long-term treatment period

7.6.5.2 Confirmed Adjudicated Cardiovascular Adverse Events

The number and percentage of subjects with confirmed cardiovascular events (ie., heart failure as determined by the adjudication committee) will be summarized by preferred term and treatment group.

7.6.5.3 Confirmed Adjudicated Hepatic Adverse Events

The number and percentage of subjects with confirmed hepatic events (as determined by the adjudication committee) will be summarized by preferred term and treatment group. In addition, a listing of all adjudicated hepatic adverse events by subjects will be provided.

7.6.5.4 Other AEOSIs

Adverse events of special interest (AEOSI) will be defined based on lists of preferred terms. These lists will be reviewed and finalized prior to database lock and unblinding of the database.
The following summaries and listings of AEOSI will include all data regardless of use of rescue medication:

- AEs of genital infection
- AEs of genital infection by gender
- AEs of Urinary Tract Infection (UTI)
- AEs of UTI by gender
- AEs of renal impairment/failure
- AEs of volume depletion
- AEs of fracture
- AEs of hepatic disorder
- AEs of hypersensitivity reactions
- AEs of severe cutaneous adverse reactions
- AEs of gastroenteritis and upper respiratory tract infections
- AEs of decreased lymphocyte count
- AEs of pancreatitis
- AEs of all malignancies
- AEs of cardiac failure
- AEs of DKA

The number and percentage of subjects with each of these events will be summarized by preferred term and treatment group in the short-term treatment period and in the combined ST+LT treatment periods.

The number and percentage of subjects with events of urinary-tract infection will be summarized for the subgroups defined on the basis of categorized variables including incidence (1, 2, 3 or > 3). A similar summary will be produced for events of genital infection.

7.6.6 Laboratory Evaluation

Unless otherwise specified, laboratory data obtained after the start of study medication dosing up to and including 4 days (30 days for liver function laboratory tests) after the last short-term dosing date (or up to and including the start of the long- term treatment period, whichever comes first) will be considered as obtained during the short-term treatment period. Laboratory data obtained from the day after the last study medication + 4 days (30 days for liver function laboratory

tests) up to the last visit date of the follow-up period will be considered as obtained during the follow-up period. Laboratory data obtained after the start of ST treatment up to and including 4 days (30 days for liver function laboratory tests) after the last dose of ST or LT medication will be considered as obtained during the combined ST+LT treatment periods.

Listings for lab data will include everything in the database.

For liver safety, a summary of proportion of subjects with elevated liver test including elevated AT (ALT and/or AST) and total bilirubin (see Appendix 1 for definition) will be provided. In addition, a summary of proportion of subjects with elevated liver test and/or reported AE of hepatic disorder will also be provided.
All laboratory evaluations performed by central laboratories will be included in summary tables. All lab tables for the CSRs will be produced in both US and SI units whenever available.

7.6.6.1 Marked Laboratory Abnormalities

Laboratory abnormalities will be evaluated based on marked abnormality values (MA). The pre-defined criteria for marked abnormalities are detailed in Appendix 3. If both the baseline and on-treatment values of a parameter are beyond the same MA limit for the parameter, then the on-treatment value will be considered a MA only if it is more extreme (farther from the limit) than was the baseline value. If the baseline value is beyond the low MA limit, and the post-baseline value is beyond the high MA limit (or vice-versa), then the post-baseline value will be considered a MA.

Laboratory abnormalities occurring during the short-term treatment period and the combined ST+LT treatment periods will be summarized by treatment group. In the short-term treatment period the summaries will be presented for both the primary and sensitivity safety analyses. The direction of change (high or low) in MA will be indicated in the tables.

For each subject with an MA for a parameter, all the subject's values of that parameter will be listed.

7.6.6.2 Change from Baseline for Selected Laboratory Parameters Over Time

All analyses of laboratory data will use observed data regardless of rescue. Visit windows are provided in Section 8.3 in order to link each laboratory test to a scheduled visit. Change from baseline during the ST treatment period and the combined ST+LT treatment periods for selected laboratory parameters will be summarized descriptively by treatment group using n's, means, medians, SEs, and 95% Cis

- hematocrit
- hemoglobin
- platelet count
- white blood cell (WBC) count
- total bilirubin

- alanine aminotransferase (ALT)
- alkaline phosphatase
- aspartate aminotransferase (AST)
- Estimated GFR using MDRD equation:

  $eGFR\ (mL/min/1.73m^2) = 175 \times (Scr)^{-1.154} \times (Age)^{-0.203} \times (0.742 \text{ if female}) \times (1.210 \text{ if African-American})$
- creatinine kinase (creatine phosphokinase) (CK)
- creatinine, serum (S_{cr})
- electrolytes - sodium, potassium, chloride, magnesium and calcium
- total protein, serum
- albumin (CV181169 only)
- inorganic phosphorus
- urinary albumin to creatinine ratio
- creatinine clearance

7.6.7 Vital Signs

Vital signs data obtained after the start of study medication dosing up to and including 4 days after the last short-term dosing date (or up to and including the start of the long- term treatment period, whichever comes first) will be considered as obtained during the short-term treatment period.

Vital signs data obtained after the start of study medication dosing up to and including 4 days after the last ST or LT dosing date will be considered as obtained during the combined ST+LT treatment period.

Visit windows are provided in Section 8.3 in order to link each vital sign measurement to a scheduled visit. Values and changes from baseline for vital sign measurements will be summarized by treatment group at each scheduled visit using descriptive statistics (using available data regardless of rescue for subjects in Treated Subjects Data Set).

7.6.8 Electrocardiograms

The normality/abnormality of the ECG tracing, as determined by the investigator, will be summarized using frequency tables on number of subjects who have a normal/abnormal ECG tracing at Week 24 and at Week 52 by the ECG tracing at baseline.

7.6.9 Pregnancy Test Results

By-subject listings of pregnancy test results will be provided using Treated Subject Data Set for the ST and the combined ST+LT treatment periods.

8 CONVENTIONS

Data collected outside the analysis window (defined in Section 8.4) after subjects discontinued open-label study medication will be excluded from analyses.

8.1 Duration of Type 2 Diabetes

Duration of Type 2 diabetes is calculated as the number of years from Type 2 diabetes diagnosis date to informed consent date:

(1 + consent date - diagnosis date) / 365.25.

The duration of diabetes will be included in the baseline diabetes characteristics listing.

If the date Type 2 diabetes was diagnosed is partially missing, the following rules will take effect:

- Missing day, but month and year are present: the day will be imputed as the 15th day of the month.
- Missing day and month, but year is present: the day and month will be imputed as 30 June of the year (provided that the imputed date is less than the consent date).
- Missing year, but day and month are present: No imputations will occur, and the subject will be excluded from all summaries related to duration of Type 2 diabetes.
- Missing day, month and year: No imputations will occur, and the subject will be excluded from all summaries related to duration of Type 2 diabetes.
- If any such imputed date falls after the consent date, then the diagnosis date will be taken as equal to the consent date.

These durations, even if partially imputed, will be listed. However, only the portion of the date of diagnosis actually observed, rather than imputed dates, will be displayed in listings.

8.2 Missing and Multiple Measurements

For listings of efficacy and safety measures, missing values will be represented as not reported.

If the blood pressure measurements are taken at a wrong position, e.g., sitting instead of standing, then these measurements will be excluded from the summary/analysis.

Some laboratory samples may be inadvertently analyzed multiple times for the same test, producing multiple lab results on the same collection date and time for the same subject.

In case of multiple observations within a single visit window, the following rules apply:

- If there are two or more observations within the same visit window, the nonmissing observation closest to the target day will be used in the analysis

- If two observations are equidistant from the target day, the non-missing observation with the later collection date will be used in the analysis
- If two or more observations are collected on the same day and have a collection time associated with them, the non-missing observation with the later collection time will be used in the analysis
- If two or more observations are collected on the same day and time, all non-missing and the average of the observations will be used in the analysis.
- If two or more observations are collected on the same day, all non-missing but with no collection time associated with at least one of them, the average of the observations will be used in the analysis.

If a visit window does not contain any observations, the data will be missing for that visit.

8.3 Laboratory evaluations

All laboratory evaluations performed by central laboratories and local laboratories that are included in the database will be listed but only the data from central lab will be included in summary tables.

8.3.1 Post treatment efficacy and safety evaluations

While efficacy and safety observations will be listed regardless of whether the subject was taking open-label study drug, observations may not contribute to summaries/analyses if they are measured after the last dose of the open-label study medication given during the treatment periods, as indicated below:

For efficacy parameters:

- Lipids and metabolic surrogate markers will be summarized only if measured on or prior to the 4th day after the last study drug treatment date.
- FPG and spot urinary glucose to creatinine ratio will be summarized/analyzed only if measured on or prior to the first day after the last study drug treatment date.
- HbA1c and body weight will be summarized/analyzed only if measured on or prior to the 8th day after the last study drug treatment date.

For safety parameters:

- SAEs and the lab measurements for the liver function tests will be included in the summaries only if occurred/measured on or prior to the 30^{th} day after the last study drug treatment date.
- All other safety events (non-severe AEs, hypoglycemia, etc) and measurements (safety lab, vital signs, etc) will be included in the summaries only if occurred/measured on or prior to the 4^{th} day after the last study drug treatment date.

8.4 Longitudinal Assessments

Day 1 for the short-term treatment period is the start date of short-term treatment medication.

Visit windows are specified below.

Table 8.4-1: Visit Windows for the Short-term Analyses

Visit	Treatment Period	Target Day	Day Range
Week 1	ST	8	2-10
Week 2	ST	15	11-22
Week 4	ST	29	23-43
Week 8	ST	57	44-71
Week 12	ST	85	72-106
Week 18	ST	127	107-148
Week 24	ST	169	149 to last day of short-term

Table 8.4-2: Visit Windows for the Short-term + Long-term analyses

Visit	Treatment Period	Target Day	Day Range
Week 1	ST	8	2-10
Week 2	ST	15	11-22
Week 4	ST	29	23-43
Week 8	ST	57	44-71
Week 12	ST	85	72-106
Week 18	ST	127	107-148
Week 24	ST	169	149 - 197
Week 32	LT	225	198 - 252
Week 40	LT	281	253-324
Week 52	LT	365	325 to last day of long-term treatment

Table 8.4-3:	**HbA1c, Weight - Visit Windows for the Short-term analyses**		
Visit	Treatment Period	Target Day	Day Range
Week 4	ST	29	2-43
Week 8	ST	57	44-71
Week 12	ST	85	72-106
Week 18	ST	127	107-148
Week 24	ST	169	149 to last day of short-term

Table 8.4-4:	**HbA1c, Weight - Visit Windows for the Short-term + long term analyses**		
Visit	Treatment Period	Target Day	Day Range
Week 4	ST	29	2-43
Week 8	ST	57	44-71
Week 12	ST	85	72-106
Week 18	ST	127	107-148
Week 24	ST	169	149 -197
Week 32	LT	225	198 - 253
Week 40	LT	281	254-324
Week 52	LT	365	325to last day of long-term treatment

8.5 Post-Treatment Efficacy Observations

While short-term treatment period (up to and including the start of long-term treatment period, where applicable) efficacy observations will be listed regardless of whether the subject was taking study drug, observations may not contribute to summaries/analyses if they are measured after the last dose of short-term study medication (up to and including the start of long-term treatment period, where applicable) as indicated below:

- Lipids and metabolic surrogate markers will be summarized only if measured on or prior to the 4th day after the last study drug treatment date.
- FPG, fasting insulin, fasting C-peptide, and spot urinary glucose to creatinine ratio will be summarized/analyzed only if measured on or prior to the first day after the last study drug treatment date.
- HbA1c, body weight, BMI and waist circumference, will be summarized/analyzed only if measured on or prior to the 8th day after the last study drug treatment date.

Similar rule applies for the combined ST+LT treatment periods.

8.6 Assignment of Doses to Adverse Events and Laboratory Assessments

In case of missing dates, prior to assigning the treatment that the subject received at the onset of an AE or at the time of a laboratory assessment, imputation rules will be applied as follows:

For AEs, a missing or incomplete onset date will be imputed according to the following conventions:

- If the onset date for an AE is missing or incomplete, an imputed date will be derived to slot the event to an appropriate analysis period. This derived date will not be reported in summary tables or listings. Every effort will be made to determine the actual onset date for the event or to obtain a reliable estimate for the onset date from the investigator.
- If an onset date is missing, the derived onset date will be calculated as the first non-missing valid date from the following list (in order of precedence):
  - First active study medication date
  - Consent date
  - Visit date corresponding to the visit at which the event was reported.
  - If a valid non-missing date is not available for any of these dates, the derived onset date will be set to missing.
- If an onset date is incomplete, the derived onset date will be calculated using the following algorithm:

- Calculate a surrogate date as the first non-missing valid date from the following list (in order of precedence):
  - First active study medication date
  - Consent date
  - Visit date corresponding to the visit at which the event was reported
  - If a valid non-missing date is not available for any of these dates, the surrogate date will be set to missing.
- Based on the information provided, set the derived date to the earliest possible date. If only a year is provided, set the derived date to January first of that year. If a year and month is provided, set the derived date to the first day of that month.
- If the surrogate date is non-missing then:

- If the derived date is on or after the surrogate date use the derived date as calculated
- If the derived date is prior to the surrogate date and the surrogate date is consistent with the partial data provided for the onset date, use the surrogate date as the derived date
- If the derived date is prior to the surrogate date and the surrogate date is not consistent with the partial data provided for the onset date then set the derived onset date to be the latest possible date based on the partial onset date information provided. If only a year is provided, set the derived date to December 31st of that year. If a year and month is provided, set the derived date to the last day of that month.

– If all three dates used to determine the surrogate date are missing, then based on the information provided, set the derived date to the earliest possible date. If only a year is provided, set the derived date to January first of that year. If a year and month is provided, set the derived date to the first day of that month.

A drug treatment file will be created, containing any starting and stopping dose as well as intermediate dose changes within each study period, with dates as recorded on the CRF. In this context,

- The date of the first dose of study medication is defined as the earliest start date with number of tablets > 0 reported on the study medication page.
- date of the last dose of study medication is defined as the latest start or stop date with number of tablets > 0 reported on the study medication page.

8.7 Concomitant Medications

Start and stop date of all concomitant medications are collected on the CRF. In order to classify medication as prior, current or concomitant, partial, missing or invalid start and stop dates will be imputed where possible as follows:

- If the reported start date is missing or invalid and the informed consent date is not missing or invalid, then the imputed start date is set equal to the informed consent date. If the consent date is missing or invalid and the birth date is not missing or invalid, then the imputed start date is set equal to the birth date. If the start date, the consent date and the birth date are all invalid or missing, then the imputed start date is set equal to missing.
- If the reported start date is partially entered, then the imputed start date is set equal to the earliest possible reported start date based on the partial entered reported start date.

- If the reported end date is missing, continuing, unknown or invalid, then the imputed end date is set equal to the most recent database extraction date.
- If the reported end date of the medication is partial, then the imputed end date is set equal to the last possible reported end date based on the partial entered reported end date.

Imputed dates will not appear on the listings of non-study medication.

8.8 Counting Rules for Adverse Events

Where a subject has the same adverse event, based on preferred terminology, reported multiple times in a single analysis period, the subject will only be counted once at the preferred terminology level in adverse event frequency tables.

Where a subject has multiple adverse events within the same system organ class in a single analysis period, the subject will only be counted once at the system organ class level in adverse event frequency tables.

When a subject has the same adverse event, based on preferred terminology, reported multiple times in a single analysis period, the following criteria, in order of precedence, will be used to

select the event to be included in summary tables:

- Relationship to study medication
- Intensity of event
- Onset date
- Related events will take precedence over unrelated events in determining the event to include in summary tables.
- More intense events will take precedence over less intense events in determining the event to include in summary tables.
- Earlier onset date events will take precedence over late onset date events in determining the onset to include in summary tables.

When reporting adverse events by intensity, in addition to providing a summary table based on the event selection criteria detailed above, summary tables will also be provided based on the most intense event during the analysis period - independent of relationship to study medication. For these tables, the following criteria, in order of precedence, will be used to select the event to be included in summary tables:

- Intensity of event
- Onset date

Fasting State

Lipids parameters listed in the protocol include Total -C, LDL-C, HDL-C and TG. For lipid parameters, only data collected in fasting state will be used for analysis. For FPG, fasting insulin, and fasting C-peptide, only assessments documented with the subject in fasting state will be summarized and listed.

8.9 Percent Compliance Calculation

See Section 7.4.3.

8.10 Strip Sign for Selected Laboratory Data

For selected laboratory test values that have been received with an operator sign as a part of the result (>, ε, <, or δ), a process to strip the operator sign will be applied and the resulting numeric values will be used for data analysis. The raw value with operator will remain as such on the CRF (or in the electronic record) and in the database.

9 CONTENT OF REPORTS

All analysis results for the short-term open-label period will be included in the ST CSR while all the analyses specified for the ST+LT will be included in the Final CSR.

10 CHANGES OF ANALYSIS FROM PROTOCOL

Changes from Protocol prior to ST data base lock are summarized in the following table:

	Changes	Reason for changes
1	Definition of Randomized Subjects Data Set was updated as follows: The randomized subject data set will consist of all randomized subjects who receive at least one dose of study medication. Whenever using the randomized subject data set, subjects will be presented in the treatment group to which they were randomized at the start of the Short-term treatment Period (even if the treatment they received was different). This is also known as the Intent-to-Treat (ITT) population. This will be the primary efficacy data set.	To be consistent with other saxagliptin plus dapagliflozin studies.
	Definition of Full Analysis Set was added.	To be consistent with other saxagliptin plus dapagliflozin studies.
	Summary and analysis of SMBG data was updated.	To be consistent with other saxagliptin plus dapagliflozin studies.
	Definition of Confirmed hypoglycemia was updated to consider all events with glucose values ≤ 70 mg/dL (3.9 mmol/L) as recorded in the hypoglycemia module of the CRF irrespective of the method using which glucose data were collected.	To avoid error in summarizing data based on glucose values which were not collected using consistent methods.
	MI sensitivity analysis for primary endpoint were added.	To be consistent with other saxagliptin plus dapagliflozin studies.

11 REFERENCES

1. Zhang M., Tsiatis A., Davidian M. Improving efficiency of inference in randomized clinical trials using auxiliary covariates. Biometrics. 2008; DOI: 10.1111/j.1541-0420.2007.00976.x.
2. Tsiatis A., Davidian M. Zhang M., Lu X. Covariate adjustment for two-sample treatment comparisons in randomized clinical trials: a principled yet flexible approach. Statistics in Medicine. 2007; DOI: 10.1002/sim.3113.
3. Service, F.J., Mean amplitude of glycemic excursions, a measure of diabetic instability. Diabetes. 1970, Sep ; 19, 644-55.
4. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. Journal of American Statistical Association. 1958; 53:457-481.
5. Greenwood M. The natural duration of cancer. Reports on public health and medical subjects. 1926; 33:1-26.
6. Workgroup on Hypoglycemia, American Diabetes Association. Defining and reporting hypoglycemia in diabetes: a report from the American Diabetes Association Workgroup on Hypoglycemia. Diabetes Care. 2005 May;28(5):1245-9.

APPENDIX 1 PHASE V INFORMATION FOR CONTINUOUS GLUCOSE MONITORING

All CGM monitoring assessments were undertaken using the Phase V® Health Outcomes Information System (PVOIS – CGM Module), Phase V Technologies, Inc. (Phase V), Wellesley Hills, MA. This PVOIS-CGM Module, consisting of validated measurement, instrumentation, scoring, algorithms and statistical programs has been continuously updated and expanded since its inception in 2004. Within the PVOIS-CGM system the two general applications are referred to as: 1) "PVOIS-CGM-Monitoring" which essentially includes the quality features that provide quality assurance and control over the CGM clinical trial tasks described above and 2) "PVOIS-CGM-Analytics" which includes the statistical transformations and conversions of the raw CGM data into the set of 170 or more "data analytics" variables.

A separate and distinct component which the PVOIS-CGM utilizes to gather physiological data directly from the patient, includes the devices, systems and software which govern the collection of the electronic source data in the study (PVOIS-CGM-Device). The selection and approval of this separate component is typically chosen by the Sponsor in consultation with Phase V.

The PVOIS is governed by a standardized set of Quality Operation Procedures ("QOP's"), Work Instructions ("WI's") and Task Specifications (TS's) which are part of the Phase V® ISO 13485/9001 Quality Management System ("PVT QMS").

Table 1. Phase V® CGM Analytics Cited in SAP

#	Data Panel	Phase V® CGM Analytics Variable	Description
1	**CGMVIS**	**BGMEANALL**	Mean of all sensor glucose values recorded in CGMRULNM units during the visit's active monitoring period (e.g. 130 milligram/deciliter)
2	**CGMVIS**	**MAGE_SC**	Mean amplitude of glucose excursions (MAGE) calculated using CGM sensor glucose data for a 24-hour period corresponding to the arithmetic mean of the blood glucose increases or decreases (from blood glucose nadirs to peaks or vice versa) when both ascending and descending segments exceed the value of one standard deviation of the blood glucose for the same 24-hour period"
3	**CGMVIS**	**WPWDSDDY**	Mean of the daily (24-hour) sensor glucose standard deviations (Within-patient and within-day standard deviation for glucose values) during the visit's active monitoring period* (e.g. 41.7 mg/dL)

#	Data Panel	Phase V® CGM Analytics Variable	Description
4	**CGMVIS/CGMDAY**	**CGM_GE71_LE180_PCTTIME**	Percent [0,100] of time spent within sensor glucose interval range [71mg/dL, 180mg/dL] during the visit's active monitoring period (e.g., 67.9%) Percent [0,100] of time spent within sensor glucose interval range [71mg/dL, 180mg/dL] during a single day of the active monitoring period (e.g., 67.9%)
5	**CGMVIS/CGMDAY**	**CGM_HYPO_PCTTIME**	Percent of time [0,100] spent with sensor glucose value less than or equal to 70 mg/dL during a single day of the active monitoring period (e.g., 4.2%) Percent of time [0,100] spent with sensor glucose value less than or equal to 70 mg/dL during a single day of the active monitoring period (e.g., 4.2%)
6	**GMVIS/CGMDAY**	**CGM NOC_HYPO_PCTTIME**	Percent of time [0,100] spent with sensor glucose value less than or equal to 70 mg/dL during military time interval [0000, 0600) during the visit's active monitoring period* (e.g., 4.2%) Percent of time [0,100] spent with sensor glucose value less than or equal to 70 mg/dL during military time interval [0000, 0600) during a single day of the active monitoring period (e.g., 4.2%)

APPENDIX 2 PHASE V INFORMATION ON PATIENT-REPORTED OUTCOMES MEASUREMENT

All patient-reported outcomes (PRO) assessments were undertaken using the Phase V® Health Outcomes Information System (PVOIS – PRO Module), Phase V Technologies, Inc., Wellesley Hills, MA. This PVOIS, consisting of validated measurement, instrumentation, scoring, algorithms and statistical programs has been continuously updated and expanded since its inception in 1987. It entails a modular approach which includes both core and disease-specific assessments for many chronic conditions such as diabetes, hypertension, cancer, HIV, obesity, lipodystrophy, migraine headache, allergic rhinitis, and dyslipidemia. Each disease or condition is assessed by include a Core assessment component and a Disease or Condition-Specific sub-module. In addition, there may be target population-specific sub-modules. The system consists of several components including patient questionnaires, data entry and management systems, database structures, scoring algorithms and statistical routines. The statistical database is maintained in SPSS for Windows Version 23 and all analytical output is generated using SPSS. A SAS statistical database and verification program may be generated using Phase V® conversion routines.

The Module described below is referred to as the PVOIS – PRO – Diabetes Module.

The quality-of-life (QOL) scales and subscales used in Protocol CV181369 include:

Quality-of-Life Scaling

- analogue perceived health (10-point rating -- overall, physical, emotional, personal, and job/work);
- functional health status (Duke activity index, diabetes-specific symptom interference index, general symptom interference index);
- general health perceptions (vitality, general health, sleep);
- mental and emotional health (psychological well-being and psychological distress);
- cognitive function (acuity, disorientation and detachment, and performance);
- symptom distress;
- sexual dysfunction.
- weight perception

A brief description of the quality-of-life scales and corresponding number of items available in the Phase V® Health Outcomes Evaluation System[7] is given in Sections 2.1.1 – 2.1.3 below. A double asterisk indicates those scales and subscales intentionally omitted from protocol CV181369.

Phase V® Generic Core Modules

A) Mental and Emotional Health (24 items) **

B) General Health Perceptions (11 items) **

C) Work/Daily Role Performance (11 items)

D) Symptom Distress (53 items)

E) Sexual Dysfunction (5 items)

F) Subjective Cognitive Functioning (25 items)**

G) Objective Cognitive Functioning** (5 psychomotor, recall, and memory tests);

H) Negative Life Events and Stress Indices;

I) Perceived Health (Analogue = 5 items and Likert General Symptom Interference Scale = 7 items)

J) Work/Role Disability (3 items) and

K) Health Care Utilization (5 items).

L) General Symptom Inference Scale

M) Duke Activity Scale

Phase V® Diabetes-Specific Modules

A) Diabetes-Specific Satisfaction with Treatment scales

B) Diabetes Symptom Interference scale (7 items)

C) Diabetes-Specific Symptom Distress module (24 items)

QOL Scale Summary Descriptions of Scales used in Protocol CV181369:

Health Care Utilization (Patient Reported): Five questions concerning frequency of hospitalizations, clinic or physician visits, nurse or other health care provider home visits, general assistance with chores and activities of daily living, telephone calls and consults to a physician, nurse or other health care provider.

Work/Disability Days (Patient Reported): 3 questions on bed days, missed days at work and reduction in the level of usual activities (non-paid, e.g. housework).

Perceived Health (Global Analogue Scale): 5 questions: Feeling past month 1) overall or in general, 2) physically, 3) emotionally, 4) personal life and 5) about job or work?

Functional Health Status: 12 functional levels of activities of daily living ranging from strenuous activity to basic activities such as dressing, bathing and eating. (Derived from the Duke Activity Scale).

Diabetes-Specific Symptom Interference: 7 questions concerning interference with 1) work, 2) social events, 3) recreational activities, 4) exercise and physical activities, 5) work effectiveness, 6) enjoying life and 7) feeling your best due to symptoms of diabetes (such as low or high blood sugar, dizziness, vision problems or problems with circulation).

General Symptom Interference: 7 questions concerning interference with 1) work, 2) social events, 3) recreational activities, 4) exercise and physical activities, 5) work effectiveness, 6) enjoying life and 7) feeling your best due to other more general symptoms or health problems such as fatigue, pain and depression.

Symptoms and Side-Effects Distress: 53 questions including diabetes-specific and general symptoms (prevalence, frequency and distress severity).

Mental and Emotional Health: 24 questions encompassing anxiety, depression, and loss of behavioral and emotional control **(Psychological Distress)**, life satisfaction, positive well being and emotional ties **(Psychological Well Being)**.

General Health Perceptions: 11 questions on sleep disturbance, vitality and general health status.

Cognitive Function and Performance: 15 questions assessing self-reported cognitive acuity, memory, reasoning, disorientation and detachment and 6 questions on self-rated cognitive performance.

Sexual Satisfaction and Dysfunction: 5 questions (separate questionnaires for males and females) concerning sexual interest and satisfaction and problems with sexual functioning.

Weight Perceptions: Weight Evaluation/Assessment: 1 item assessing the subject's perceptions of their weight ranging from very underweight to very overweight and Weight Concern: 1 item assessing the subject's distress associated with his or her weight.

Composite Psychosocial: Mean of the *subscales* of Mental Health and Health Perceptions.

Composite QOL: Mean of the *subscales* of Mental Health (6 scales), Health Perceptions and Sexual Satisfaction and Dysfunction.

Overall QOL (Item-wise): Mean of all *items* in the Mental Health and Health Perceptions scales.

Diabetes Treatment Satisfaction

The treatment satisfaction measures used in protocol CV181369 include those constructs that focus on the patient's expectations and experiences with the process and perceived outcomes of the therapeutic regimen.

The Diabetes Treatment Satisfaction module[8] was developed independently by Phase V Technologies using a compendium of existing diabetes satisfaction items and incorporating new items developed from a series of focus group studies with persons with type 1 and type 2 diabetes experienced with newer oral hypoglycemic agents, insulin formulations, and pump and inhaled insulin technology. Item pool selection, psychometric analysis, and field-testing was conducted by Phase V Technologies.

The following constructs were evaluated using the long-form diabetes treatment satisfaction module including:

- Life Interference
- Convenience

- Burden
- Acceptance of Negative Aspects (including side effects, stress, hassle)
- Acceptance of Positive Aspects (including effectiveness, ease of use, comfort)
- Overall Patient Preference (compared to other treatments).

The full diabetes treatment satisfaction battery comprises four sections dealing with diabetes treatment involving A) any form, B) insulin (regardless of delivery), C) insulin injections and D) insulin inhalers. Scales used in this protocol are described in Sections 2.2.1.

Overall Satisfaction with Treatment (Module A)

This 72-item section assesses satisfaction with diabetes treatment in general and is not targeted towards any one kind of treatment or delivery system. The subscales included:

Advocacy: 2 items on recommending and advocating the treatment to other persons with diabetes, including family and friends.

Burden: 14 items concerning multiple aspects of burden of the therapeutic regimen including adherence, diet, exercise, burden for performing daily activities, social activities and enjoying life.

Convenience: 6 items relating to ability to remember taking medication, overall convenience, being pleased with convenience, amount of time required to manage diabetes.

Efficacy: 3 items on the patient's perception of the treatment's ability to control blood sugar.

Flexibility: 4 items on how flexible the treatment is for scheduling and allowing variability in meals and overall flexibility.

General Satisfaction: 5 items on general satisfaction and being pleased with current medication.

Hassle: 8 items specific to the amount of bother and hassle of the regimen including dosing, treatment supplies, carrying supplies, supply disposal, pain and discomfort, and worries about hypoglycemia and hyperglycemia.

Interference: 11 items concerning how much the diabetes medication interferes with daily routine, meals, recreation, family life, sleep schedules, energy levels, making plans, traveling, having fun and overall quality of life.

Pain: 3 items concerning pain and discomfort.

Preference: 2 items rating how strong the desire to search out other regimens that might be better and to continue on current regimen.

Side effects: 5 items concerning gaining weight, unpleasant feelings, distress with hypoglycemia and hyperglycemia.

Social: 9 items rating the treatment's interference with social interactions with family and friends, travel, having fun, and problems in performing work and social roles.

Overall Satisfaction: mean of the 12 individual general satisfaction scales.

APPENDIX 3 MARKED ABNORMALITY CRITERIA

The criteria for marked abnormality for each variable are listed in the following table. Note that a post-baseline lab value will be considered a MA only if it satisfies the specified criteria and is more extreme (farther from the limit) than is the baseline value.

Clinical Laboratory variables	Units	Marked Abnormality Criteria	
		Low	High
Hematology			
HCT males/females	%	< 20.0%	> 55.0%
HCT males/females	%		> 60.0%
Hemoglobin males/females	g/dL	< 6 g/dL	> 18 g/dL
Hemoglobin males/females	g/dL		> 20 g/dL
Blood Chemistry			
Albumin	g/dL	≤ 2 g/dL	> 6 g/dL
Total protein	g/dL		> 10 g/dL
ALP	U/L		> 3X ULN
ALT	U/L		> 3X ULN
AST	U/L		> 3X ULN
ALT	U/L		> 5X ULN
AST	U/L		> 5X ULN
ALT	U/L		> 10X ULN
AST	U/L		> 10X ULN
ALT	U/L		> 20X ULN
AST	U/L		> 20X ULN
Total Bilirubin	mg/dL		> 2X ULN if PreRx ≤ ULN; > 3X ULN if PreRx > ULN
Glucose, Plasma Unspecified	mg/dL	< 54 mg/dL	> 350 mg/dL
Na (Sodium)	mEq/L	< 130 mEq/L	> 150 mEq/L
Na (Sodium)	mEq/L	< 120 mEq/L	
K (Potassium)	mEq/L	≤ 2.5 mEq/L	≥ 6.0 mEq/L
HCO3 (Bicarbonate)	mEq/L	≤ 13 mEq/L	
BUN	mg/dL		≥ 60 mg/dL
Creatinine	mg/dL		≥ 1.5X PreRx CREAT
Creatinine	mg/dL		≥ 2.5 mg/dL
CK (Creatine Kinase)	U/L		> 5X ULN
CK (Creatine Kinase)	U/L		> 10X ULN

APPENDIX 4 GEOGRAPHIC REGIONS

Geographic Region	Countries
North America	United States
Latin America	Mexico
Europe	South Africa, Germany, Poland, Spain, Sweden Czech Republic Hungary, Denmark, Romania