Statistical Analysis Plan

Study Code	D1689C00013
Edition Number	3.0
Date	8 October 2019

A 52-WEEK INTERNATIONAL, MULTICENTER, RANDOMIZED, DOUBLE-BLIND, ACTIVE-CONTROLLED, PARALLEL GROUP, PHASE 3B TRIAL WITH A BLINDED 104-WEEK LONG -TERM EXTENSION PERIOD TO EVALUATE THE EFFICACY AND SAFETY OF SAXAGLIPTIN CO-ADMINISTERED WITH DAPAGLIFLOZIN IN COMBINATION WITH METFORMIN COMPARED TO GLIMEPIRIDE IN COMBINATION WITH METFORMIN IN ADULT PATIENTS WITH TYPE 2 DIABETES WHO HAVE INADEQUATE GLYCEMIC CONTROL ON METFORMIN THERAPY ALONE

A 52-WEEK INTERNATIONAL, MULTICENTER, RANDOMIZED, DOUBLE-BLIND, ACTIVE-CONTROLLED, PARALLEL GROUP, PHASE 3BTRIAL WITH A BLINDED 104-WEEK LONG -TERM EXTENSION PERIOD TO EVALUATE THE EFFICACY AND SAFETY OF SAXAGLIPTIN CO-ADMINISTERED WITH DAPAGLIFLOZIN IN COMBINATION WITH METFORMIN COMPARED TO GLIMEPIRIDE IN COMBINATION WITH METFORMIN IN ADULT PATIENTS WITH TYPE 2 DIABETES WHO HAVE INADEQUATE GLYCEMIC CONTROL ON METFORMIN THERAPY ALONE

CRO Statistician

PPD

9 OCT 2019
Date

Global Product Statistician

John Monyak

Date

A 52-WEEK INTERNATIONAL, MULTICENTER, RANDOMIZED, DOUBLE-BLIND, ACTIVE-CONTROLLED, PARALLEL GROUP, PHASE 3BTRIAL WITH A BLINDED 104-WEEK LONG -TERM EXTENSION PERIOD TO EVALUATE THE EFFICACY AND SAFETY OF SAXAGLIPTIN CO-ADMINISTERED WITH DAPAGLIFLOZIN IN COMBINATION WITH METFORMIN COMPARED TO GLIMEPIRIDE IN COMBINATION WITH METFORMIN IN ADULT PATIENTS WITH TYPE 2 DIABETES WHO HAVE INADEQUATE GLYCEMIC CONTROL ON METFORMIN THERAPY ALONE

CRO Statistician

Emma May — Date

Global Product Statistician PPD — 9 Oct, 2019
Date

TABLE OF CONTENTS

AMENDMENT HISTORY

Date	Brief description of change
April 6, 2016	Initial Approved SAP
Version 2, September 27, 2017	• Inserted texts from the core SAP and removed the references to core SAP. • Definition of confirmed hypoglycemia was updated. • Start of long-term treatment period was clarified in Section 6.1. • Section 6.3.5 was added to include the definition of Full Analysis Set. Following subsections were renumbered. • In Section 6.3.6 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 • Section 7.5.2 was updated by adding new sensitivity analyses for primary endpoint. • Section 0 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.

Date	Brief description of change
Version 3, 1 October 2019	During ST CSR appendices development, following discrepancies were identified in SAP: • Section 7.6.1, 2nd paragraph: Reference to the section number of "Counting rules for AE", updated to be Section 8.8 • Section 7.1.10, Figure 7.1.3-1 Reference#7 mentioned in the foot note of this figure has been updated to #3, as listed in section 11 • APPENDIX 2, Quality-of-Life Scaling Protocol number corrected from CV181-369 to CV181-365 As per protocol version 5.0 • Sections 4.3 and 7.5.5.3 the time point for obtaining last MRI image is W156, the time point to be changed to W122 +/- 4 weeks As the ST+LT analysis was planned the following updates were made: • Sections 2.7, 7.5.8, 7.6.5.4, 12 were added • Clarification of, and details pertaining to, the ST analyses were added in order to reliably and robustly reproduce the results at the ST+LT timepoint, and for any future analysis. Confirmation was added throughout the document to identify which datasets and analyses were applicable to the short-term and/or short-term and long-term treatment periods. Section 7.1.2 – removed stratum from the model, and confirmed the interaction term Section 7.1.8 – removed integration of group's modelled probability of response over the observed distribution of baseline covariate Section 7.1.1.3, 7.5.3 and 7.6.6.2 – details of handling missing data were clarified Section 7.1.4 – updated the list of summary stats to include Q1, Q3, geometric mean and CV Table 7.2-1 - updated to make relevant protocol deviations (RPDs) applicable to the ST+LT period Tables 7.3.2.x – categories in table A updated in light of data, and table D added for Hematuria at the ST+LT treatment period

Date	Brief description of change
	Section 7.4.1 – clarification of how duplicate dosing data is handled at the ST+LT analysis
	Section 7.4.2 – sub-categories removed for allowed concomitant medication summary table
	Section 7.5.2 – the sensitivity analysis to be undertaken at the ST+LT treatment period were added
	Section 7.5.4 – confirmation that the proportion of subjects achieving a therapeutic glycaemic response is adjusted by baseline Hb1Ac
	Section 7.5.5 – clarification that no adjustments for multiplicity will be done - p-values are nominal only
	Section 7.5.9 – clarification of the author of the PRO report
	Section 7.5.5.2 – confirmation that model assumptions will be checked
	Section 7.5.5.3 – details of lipid summary added
	Section 7.5.5.4 and 8.2 – Updates in line with the errata to the ST CSR
	Section 7.5.8 – details of biomarker analysis added
	Section 7.6.1 – additional information about handling surgical or spontaneous amputations was added
	Section 7.6.3 – SAE listings added for ST+LT analysis
	Section 7.6.7, 8.3.1 and 8.5 – parameters were explicitly listed
	Section 7.6.5.1 – details were added to described how conflicting hypoglycaemia events were handled
	Section 7.6.5.5 – other AESI were defined more clearly
	Section 8.2 – handling of missing MRI data was added
	Table 8.4 – tables were updated and added to describe the various windows applied.

1 BACKGROUND AND RATIONALE

Type 2 diabetes mellitus (T2DM) is a chronic progressive disease, characterized by hyperglycemia and an increased risk of microvascular and macrovascular complications. Achieving and maintaining the glycemic treatment goal is challenging. Typically, the treatment paradigm consists of a step-wise addition of different classes of antihyperglycemic drugs, as most patients eventually require two or more agents to achieve or maintain glycemic targets. Among the medications approved for the treatment of T2DM, metformin is the recommended drug of choice for initiating oral therapy, while other classes of antidiabetic agents are typically added sequentially as second- and third-line agents. Despite multiple drugs and classes being available, many patients are still inadequately controlled. An ideal add-on to metformin would provide strong HbA1c reduction through complementary mechanisms of action (MoA), with weight loss, and no hypoglycemia.

Saxagliptin (BMS-477118) is a highly potent, selective, reversible, and competitive dipeptidyl peptidase 4 (DPP4) inhibitor. DPP4 is the enzyme responsible for the inactivation of glucagon-like peptide-1 (GLP-1) and glucose-dependent insulinotropic polypeptide (GIP). By inhibiting DPP4, saxagliptin potentiates active endogenous GLP-1 concentrations, augmenting the physiological mechanism of insulin secretion and decreasing glucagon release, thereby reducing post-prandial and fasting glucose levels in patients with T2DM.

Dapagliflozin is a stable, competitive, reversible, highly selective, and orally active inhibitor of human renal sodium glucose co-transporter 2 (SGLT2), the major transporter responsible for renal glucose reabsorption in the kidney. Dapagliflozin's mechanism of action results in the direct and insulin-independent elimination of glucose by the kidney.

This is a Phase 3b study performed as part of the development program for saxagliptin / dapagliflozin fixed-dose combination (FDC, BMS-986098) to improve glycemic control as an adjunct to diet, exercise, and metformin treatment when treatment with both saxagliptin and dapagliflozin is appropriate. The study is intended to compare effects of saxagliptin co-administered with dapagliflozin in combination with metformin to glimepiride in combination with metformin. Saxagliptin and dapagliflozin have demonstrated, both individually and in combination with metformin, a favorable safety and tolerability profile. They have demonstrated a low propensity for hypoglycemia. Both drugs have either demonstrated weight neutrality (saxagliptin) or moderate weight reduction (dapagliflozin). Dapagliflozin and saxagliptin have also demonstrated persistent beneficial effects on HbA1c. Thus, a second-line oral dual add-on therapy with saxagliptin co-administered with dapagliflozin could be a new option, as part of a triple combination that includes drugs with complementary mechanisms of action, low risk of hypoglycemia, and the potential for moderate weight loss, providing a more effective and patient-friendly approach to the treatment of T2DM.

The purpose of this Study-specific Statistical Analysis Plan (SAP) is to present a detailed plan of the analyses of the Phase 3b study CV181-365.

Research Hypothesis:

Co-administration of saxagliptin and dapagliflozin, in addition to metformin, results in better glycemic control, as measured by HbA1c, over a treatment period of 52 weeks, compared to the addition of glimepiride to metformin in subjects with inadequately controlled T2DM.

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 double-blind period. All relevant queries must be resolved, and the database must be locked for this period before the blind is broken (for the sponsor) for 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 CV181-365 study is an international multicenter, randomized, double-blind, active-controlled, parallel-group Phase 3b study. The primary endpoint is the change from baseline in HbA1c with saxagliptin in combination with dapagliflozin on top of background therapy with metformin, compared to glimepiride. The primary endpoint of HbA1c will be assessed at 52 weeks, however, the trial will continue with a 104-week blinded extension. Subjects with documented T2DM treated with a stable dose of metformin will be enrolled. All subjects should be treated according to regional standards of care for diabetes. Subjects must be receiving metformin in accordance with the product label for their country. Approximately 440 subjects meeting all eligibility criteria will be randomized (1:1) to receive either saxagliptin co-administered with dapagliflozin or glimepiride. Enrollment of subjects based on disease state, and geographic region will be monitored and may be capped to ensure adequate representation. All potentially eligible subjects will undergo a screening visit. Each subject will sign an Informed Consent Form (ICF) prior to having any screening evaluations performed. Subjects who fulfill all eligibility requirements will enter a 2-week screening period and a 2-week lead-in period, during which they will continue their background medication with at least 1500 mg metformin. Approximately 2 weeks after entering the lead-in period, subjects will be expected to undergo a randomization visit. Subjects may withdraw prior to the randomization visit, or be withdrawn by study staff for any reason. At this visit, subjects will be re-evaluated by study staff to determine if after testing performed at the screening visit, after assessment of compliance, and any clinical changes that may have occurred during the lead-in period, the subject remains eligible and committed to participation in the study. If for any reason, prior to or during the randomization visit the subject is no longer eligible or interested in participating in the trial, he or she will be considered a lead-in failure, will not be randomized, and will not have additional follow up. If a subject is committed to participation, completes the lead-in period, and continues to meet eligibility criteria at the randomization visit, he or she will be randomized and will receive either saxagliptin 5 mg co-administered with dapagliflozin 10 mg plus metformin and matching placebo or glipimeride plus metformin and matching placebo in a double-blind fashion. Subjects will return to the site every 3 weeks to enable

titration of glimepiride from Week 3 up to Week 12, and then approximately every 12 to 13 weeks until the last visit (up to Week 156) for assessment of events related to the objectives of the study, tolerability and safety. Assessment of treatment compliance and provision of investigational product (IP) will be done at these visits.

Saxagliptin 5 mg tablets or matching placebo tablets will be provided, administered orally once daily for the 52-week double-blind treatment period and the 104-week subject and site blinded extension period. Saxagliptin 5 mg is the maximum recommend daily dose. No up- or down-titration of saxagliptin will be allowed. Dapagliflozin 10 mg tablets (also the maximum recommended daily dose) or matching placebo tablets will be provided, administered orally once daily for the 52-week double-blind treatment period and the 104-week subject and site blinded extension period. No up- or down-titration of dapagliflozin will be allowed.

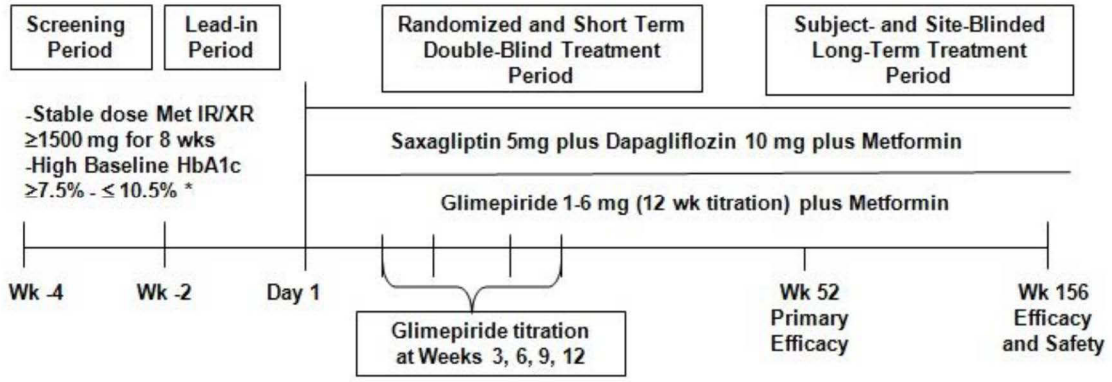
Glimepiride 1-6 mg or matching placebo capsules, administered orally once daily, will be provided for the 156-week blinded treatment period. During the first 12 weeks, the glimepiride dose will be slowly titrated in a stepwise blinded fashion depending on glycemic control. For the duration of the study, FPG will be measured at the study center using a glucose analyzer provided by BMS. The investigator's decision on dose titration (either upwards if no issues with hypoglycemia or downwards if recurrent hypoglycemia occurs) will take both the plasma glucose measurements (measured prior to the visit), and the investigator's measurements at the titration visits, into account. The glimepiride dose will be titrated to optimal effect (FPG $\leq$ 110 mg/dL [$\leq$ 6.1 mmol/L]) or the highest tolerable dose during the first 12 weeks. The starting dose for glimepiride is 1 mg per day (once daily), which can be further increased by increments of 1-2 mg at 3-week intervals to a maximum of 6 mg per day (maximum recommended dose in the Summary of Product Characteristics and thus the maximum dose to be used in this study). The titration steps will be 2 mg, 3 mg, 4 mg, and 6 mg once daily if needed. In subjects for whom titration is not medically indicated at Week 3, re-assessment for titration will occur at Week 6, Week 9, and Week 12.

During the randomized treatment period, diet and life-style modification will continue to be reinforced. All randomized subjects who are rescued with insulin (or other oral anti-diabetes therapy) and taking randomized IP should continue scheduled study visits as planned in the Clinical Study Protocol (CSP). Subjects who decide to prematurely discontinue the investigational products will always be asked about the reason(s) and the presence of any AEs. Subject should return and complete the procedures described for End of Treatment Visit (EOT) as soon as possible but at the latest 7 days after discontinuation of investigational product. The subjects with unresolved AEs should also be followed-up if medically indicated as judged by investigator. All subjects who discontinue from treatment will follow the same visit schedule as subjects who remain on treatment. The only exception to this requirement is when a subject withdraws consent for all study procedures or loses the ability to consent freely (i.e., is imprisoned or involuntarily incarcerated for the treatment of either a psychiatric or physical illness). Subjects lost to follow-up (as defined by; unable to reach the subject after 3 documented phone calls, fax, email, or attempts to contact him/her through subject locator agencies [if allowed per national regulation]

and having sent one letter by registered/certified mail; all should be documented in the subject's medical records).

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

Figure 2.1-1: Study Design Schematic



In addition, two exploratory sub-studies will be conducted along with the above main study.

The reduction in body weight consistently observed during treatment with dapagliflozin has been shown to result in total body fat reduction. A sub-study will investigate whether this reduction in body fat is equally distributed among subcutaneous and visceral adipose tissue or if there is a preferential loss of adipose tissue in the visceral compartment. Body composition measurements by magnetic resonance imaging (MRI) can separately quantify the subcutaneous and the visceral adipose tissue. Non-invasive quantification of liver lipid content by magnetic resonance imaging-estimated proton density fat fraction (MRI-PDFF) is an emerging biomarker for assessment of the beneficial effects of weight loss and improvement in glycemic control on liver lipid metabolism. In this protocol, a subpopulation of approximately 60 randomized subjects (~ 30 subject/treatment arm) who agree to participate (separate informed consent) will undergo MRI-PDFF measurement of subcutaneous and visceral adipose tissue, together with liver lipid content at baseline, at Week 52, and at Week 156.

Continuous Glucose Monitoring (CGM) is a useful technology (in addition to HbA1C) to qualitatively, as well as quantitatively, monitor the quality of 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). Co-administration of saxagliptin and dapagliflozin is expected to better demonstrate less variability in glycemic control and more time spent in the euglycemic range compared to treatment with glimepiride administration. In this protocol, a subpopulation of approximately 120 randomized subjects (~ 60 subject/treatment arm) who agree to participate (separate informed consent) will have CGM performed for periods of 7 days at baseline and at Week 52 of treatment.

2.2 Treatment Assignment

CCI [redacted]

Subjects entering the 52-week double-blinded 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) double-blind treatment arms in a 1:1 ratio:

- Blinded saxagliptin 5 mg and dapagliflozin 10 mg
- Titrated to blinded glimepiride 1 mg, 2 mg, 3 mg, 4 mg, or 6 mg

Randomization schedules for both subject treatment and containers will be generated and kept by BMS.

Subjects entering the 104-week long-term subject- and site-blinded treatment period

Following completion of the 52-week double-blinded treatment period, subjects eligible for the long-term subject- and site-blinded treatment period will be continued in their same randomization assignment based on their original randomization grouping. Subjects that were assigned to receive the blinded saxagliptin 5 mg and dapagliflozin 10 mg arm will continue to receive the blinded saxagliptin 5 mg and dapagliflozin 10 mg. Subjects that were assigned to receive the blinded glimepiride will continue to receive blinded glimepiride (at a dose of 1 mg, 2 mg, 3 mg, 4 mg, or 6 mg). At all study visits when study drug is dispensed, each subject will be assigned multiple container numbers by the IVRS. Container numbers will be assigned non-sequentially and will correspond to the numbers printed on the containers and bottles containing study drug, and will be recorded on the appropriate eCRF. The IVRS will be available 24 hours per day, 7 days a week.

2.3 Blinding and Unblinding

2.3.1 Blinding

The investigator, AstraZeneca personnel, and subjects will remain blinded to treatment allocation throughout the short-term, double-blind treatment period. The database used for the analysis of the short-term double-blind data of the study will be locked after all subjects have terminated the short-term double-blind treatment period of the study. The locked database will be unblinded for reporting purposes. To protect the integrity of the long-term treatment period of the studies, the subjects and investigators will not have access to the individual treatment assignments until the long-term treatment period has been completed.

For the duration of the double-blind treatment period (Day 1 to Week 52), the HbA1c and the urinary glucose values, including the urinary glucose:creatinine ratio, will be masked to the

sponsor and will not be available to the investigator. During the site- and subject-blinded treatment period, (after Week 52 and through Week 156), the above measurements will be unmasked to the sponsor. However, they still will not be available to the investigator except HbA1c values until after the study completion. HbA1c values will be provided to the investigator during the site- and subject-blinded treatment period.

2.3.2 Unblinding

Blinding is critical to the integrity of this clinical trial. However, in the event of a medical emergency or pregnancy, during which knowledge of the identity of the investigational product is critical to the subject's management, procedures are in place to have the blind broken for an individual subject. A listing of all subjects whose treatment is unblinded during the study will be included in the Clinical Study Report (CSR). A separate procedure is in place for unblinding in case of expedited safety reporting to regulatory authorities.

2.4 Protocol Amendments

Five protocol amendments have been processed and are incorporated into this SAP. This SAP is based on the Revised Protocol 04, which incorporates amendment 5; date 8 February 2018

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:

- Aspartate aminotransferase (AST) and/or alanine aminotransferase (ALT) > 3X upper limit of normal (ULN) and total bilirubin (TB) >2X 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
- Hepatic disorders leading to discontinuation from study treatment and/or death
- Liver laboratory abnormalities such as elevated AST and/or ALT with or without TB elevations

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

2.7 DKA Adjudication Committee

The diabetic ketoacidosis (DKA) Adjudication Committee will assess available information on each potential DKA event and will classify the event in accordance with the definitions in the DKA

Adjudication Charter T2DM. All potential events of DKA will be recorded in the eCRF and submitted to an independent DKA Adjudication Committee.

The DKA Adjudication Committee will be kept blinded to the study drug treatment received by each subject with a potential DKA event in the clinical study. A separate DKA Adjudication Manual will define and describe the procedures for the collection of DKA information, handling, adjudication criteria and reporting of these events/cases.

3 OBJECTIVES

The following sections list the primary, secondary, and exploratory objectives specified in the protocol.

3.1 Primary

To compare the mean change from baseline in HbA1c achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin at Week 52.

3.2 Secondary

Four secondary efficacy objectives for the short-term treatment period and two secondary efficacy objectives for the long-term treatment period are identified for consideration in this study, in addition to the primary objective.

Short-term:

- To compare the mean change from baseline in total body weight achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin at Week 52.
- To compare the proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0%, at Week 52 with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin.
- To compare the mean change from baseline in systolic blood pressure (SBP) achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin at Week 52.
- To compare the time to treatment intensification (addition of insulin or other glucose-lowering agent for rescue therapy or discontinuation for lack of glycemic control) achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin during the 52-week double-blind treatment period.

Long-term:

- To compare the time to treatment intensification (addition of insulin or other glucose-lowering agents for rescue therapy or discontinuation for lack of glycemic control) achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy

with metformin, compared to glimepiride added to current background therapy with metformin during the 156-week treatment period.

- To compare the proportion of subjects achieving therapeutic glycemic response, defined as HbA1c < 7.0%, at Week 156 with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin

3.3 Exploratory Objectives

Short-term:

- To assess the proportion of subjects achieving therapeutic glycemic response, defined as HbA1c < 7.0%, without any hypoglycemia at Week 52 achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin.
- To assess the proportion of subjects achieving therapeutic glycemic response, defined as HbA1c < 7.0%, without any hypoglycemia and weight gain at Week 52 achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin.
- To assess mean change from baseline in fasting plasma glucose (FPG) achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin at Week 52.
- To assess change from baseline of metabolic measures (glucagon, C-peptide, pro-insulin, insulin, iHOMA2) at Week 52 achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin.
- To assess change from baseline in average glucose values and post-prandial glucose values measured by 6-point self-monitored blood glucose (SMBG) profiles at Week 52 achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin.
- To assess change from baseline of urinary albumin to creatinine ratio at Week 52 achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin.
- To assess change from baseline of biomarkers (high-sensitivity C-reactive protein [hsCRP], N-terminal pro-brain natriuretic peptide [NT Pro-BNP]) at Week 52 of double-blind treatment achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin.
- To assess changes in patient-reported treatment satisfaction, quality of life, and barriers to medication adherence achieved with saxagliptin, in co-administration with dapagliflozin,

added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin at Week 52.

- To assess change from baseline in glycemic variability, as defined by mean amplitude of glycemic excursions (MAGE), in a sub-group using CGM achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin at Week 52 in the subpopulation of subjects who undergo CGM.
- To assess changes from baseline to Week 52 in visceral and subcutaneous adipose tissue volume and percent hepatic lipid content achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin in the subpopulation of subjects who undergo MRI-PDFF.

Long-term:

- To assess the mean change from baseline in HbA1c achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin at Week 156.
- To assess the time spent at or below HbA1c target (HbA1c < 7.0%) achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin during 156 weeks.
- To assess the proportion of patients achieving therapeutic glycemic response, defined as HbA1c < 7.0%, without any hypoglycemia at Week 156 achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin.
- To assess the proportion of patients achieving therapeutic glycemic response, defined as HbA1c < 7.0%, without any hypoglycemia and weight gain at Week 156 achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin.
- To assess mean change from baseline in FPG achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin at Week 156.
- To assess change from baseline of metabolic measures (glucagon, C-peptide, pro-insulin, insulin, iHOMA2) at Week 156 achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin.
- To assess change from baseline of urinary albumin to creatinine ratio at Week 156 achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin.
- To assess change from baseline of biomarkers (hsCRP, NT Pro-BNP) at Week 156 achieved with saxagliptin, in co-administration with dapagliflozin, added to current background

therapy with metformin, compared to glimepiride added to current background therapy with metformin.

- To compare the mean change from baseline in total body weight achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin at Week 156.
- To compare the mean change from baseline in systolic blood pressure achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin at Week 156.
- To assess changes in patient-reported treatment satisfaction, quality of life, and barriers to medication adherence achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin at Week 156.
- To assess changes from baseline to Week 122 in visceral and subcutaneous adipose tissue volume and percent hepatic lipid content achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin in the subpopulation of subjects who undergo MRI-PDFF.

3.4 Safety objectives:

- Confirmed hypoglycemia defined as: blood glucose ≤ 70 mg/dL (3.9 mmol/L) recorded in the hypoglycemia module of the CRF for saxagliptin co-administered with dapagliflozin vs glimepiride at 52 and 156 weeks of therapy.
- To evaluate the safety and tolerability of treatment achieved with saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, compared to glimepiride added to current background therapy with metformin at 52 and 156 weeks of therapy.

4 ENDPOINTS

4.1 Primary Endpoint

The primary endpoint for analysis is mean change from baseline to Week 52 in HbA1c.

4.2 Secondary Endpoints

The secondary efficacy endpoints for the short-term treatment period include:

- Mean change from baseline in total body weight at Week 52.
- Proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0%, at Week 52.
- Mean change from baseline in SBP at Week 52.
- Time to treatment intensification (addition of insulin or other glucose-lowering agent for rescue therapy or discontinuation for lack of glycemic control) during the 52-week double-blind treatment period.

The secondary efficacy endpoints for the long-term treatment period include:

- Time to treatment intensification (addition of insulin or other glucose-lowering agent for rescue therapy or discontinuation for lack of glycemic control) during the 156-week treatment period.
- Proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0%, at Week 156.

4.3 Exploratory Endpoints

The exploratory endpoints for the short-term treatment period include:

- Proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0%, without any hypoglycemia at Week 52.
- Proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0%, without any hypoglycemia and weight gain at Week 52.
- Mean change from baseline in FPG at Week 52.
- Mean change from baseline of metabolic measures (glucagon, C-peptide, pro-insulin, insulin, iHOMA2) at Week 52.
- Mean change from baseline in average glucose values and post-prandial glucose values measured by 6-point SMBG profiles at Week 52.
- Mean change from baseline in urinary albumin to creatinine ratio at Week 52.
- Mean change from baseline in biomarkers (hsCRP and NT Pro-BNP) at Week 52.
- Mean change from baseline in patient-reported treatment satisfaction, quality of life, and barriers to medication adherence at Week 52.
- In a sub-study: Mean change from baseline in glucose variability, as defined by mean amplitude of glycemic excursions (MAGE), using CGM at Week 52.
- In a sub-study:
  - Mean change from baseline in visceral and subcutaneous adipose tissue volume as assessed by MRI at Week 52 in the subpopulation of subjects who undergo MRI-PDFF.
  - Mean change from baseline in percent hepatic lipid content as assessed by MRI at Week 52 in the subpopulation of subjects who undergo MRI-PDFF.

The exploratory endpoints for the long-term treatment period include:

- Mean change from baseline in HbA1c at Week 156.
- Mean time spent at or below HbA1c target (HbA1c < 7.0%) during 156 weeks.
- Proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0%, without any hypoglycemia at Week 156.
- Proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0%, without any hypoglycemia and weight gain at Week 156.
- Mean change from baseline in FPG at Week 156.
- Mean change from baseline of metabolic measures (glucagon, C-peptide, pro-insulin, insulin, iHOMA2) at Week 156.
- Mean change from baseline in urinary albumin to creatinine ratio at Week 156.
- Mean change from baseline in biomarkers (hsCRP and NT Pro-BNP) at Week 156.

- Mean change from baseline in total body weight at Week 156.
- Mean change from baseline in SBP at Week 156.
- Mean change from baseline in patient-reported treatment satisfaction, quality of life, and barriers to medication adherence at Week 156.

- Mean change from baseline in visceral and subcutaneous adipose tissue volume and percent hepatic lipid content as assessed by MRI at Week 122 in the subpopulation of subjects who undergo MRI-PDFF.

4.4 Safety Endpoints

The safety endpoints include:

- The proportion of subjects with hypoglycemia events and the frequency and severity of the hypoglycemia events at Weeks 52 and Week 156
- The proportion of subjects experiencing adverse events (AE) and marked abnormalities in clinical laboratory tests at Week 52 and Week 156
- 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 at Week 52 and Week 156

5 SAMPLE SIZE AND POWER

The mean change in baseline in HbA1c at Week 52 will be assessed by comparing saxagliptin, in co-administration with dapagliflozin, added to current background therapy with metformin, versus the active comparator treatment group (glimepiride added to current background therapy with metformin).

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 using last observation carried forward (ANCOVA with 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.

Assuming a mean change from baseline in HbA1c difference effect of 0.35% for saxagliptin in co-administration with dapagliflozin plus metformin vs glimepiride plus metformin, a common standard deviation of 1.1%, and using a 2-sided significance level of 0.05 for the comparison, and assuming a 5% non-evaluability rate, then a sample size of 220 patients per arm will yield 90% power for the comparison of saxagliptin in co-administration with dapagliflozin plus metformin against glimepiride plus metformin.

6 STUDY PERIODS, TREATMENT REGIMENS AND POPULATIONS FOR ANALYSES

6.1 Study Periods

Study CV181-365 consists of 4 study periods:

1) **Screening period (Period A)**. This period starts with enrollment and ends on the start of lead-in period.
2) **Lead-in period (Period B)**. During this period, subjects will continue their diabetes management with metformin, diet and exercise. No placebo or study medication will be provided during the lead-in period.
3) **52-week short-term double-blind treatment period (Period C)**. On the Day 1 visit, subjects who meet all the protocol-specific enrollment and randomization criteria will be randomized.
4) **104-week long-term site and subject blind treatment period (Period D)**. After completing the short-term period, subjects will enter the 104-week site and subject blinded 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 early will follow the same visit schedule as subjects who remain on treatment.

6.2 Treatment Regimens

The "as randomized" treatment group is defined as the treatment group to which a subject was randomized at the start of the double-blind 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 and will be used for subject disposition and protocol deviation results for the ST analysis only .

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 double-blind 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.

A subset of the randomized subjects data set will be used in the short-term, and short-term and long-term analysis, defined as randomized subjects who consented to take part in the MRI-PDFF sub-study.

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.

The Evaluable Subject Data Set will only be used in the main efficacy result at Week 52 and no evaluable population will be summarized in the ST+LT CSR, although RPDs will be summarized for the ST+LT treatment period.

6.3.6 Treated Subjects Data Set

The Treated Subjects Data Set for the short-term treatment period and the short- and long-term treatment period will consist of all subjects who receive at least one dose of double-blind study medication during the short-term double-blind 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 short-term double-blind treatment and entered the long-term 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 short-term double-blind treatment period (even if the treatment they received in the short- or long-term treatment period was different).

7 STATISTICAL ANALYSES

7.1 General Methods

7.1.1 Definitions

7.1.1.1 Baseline Value

Unless otherwise stated, for the short-term, and short-term and long-term analyses, 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 short-term and long-term 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, or short-term and long-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.

Safety analyses include either subjects' data "prior to rescue medication", or "regardless of rescue medication". For analyses based on "prior to rescue medication", subjects' data up to and on the date of the first dose of rescue medication, subject to the prescribed cut off of the IP treatment

discontinuation (see Section 0), will be used. For analyses based on "regardless of rescue medication", subjects' data are excluded only after treatment discontinuation based on the prescribed cut off days (see Section 0).

Efficacy tables are summarized for subjects "prior to rescue and treatment discontinuation" and "regardless of rescue and treatment discontinuation". For analyses based on "prior to rescue and treatment discontinuation" subjects' data will be included up to the earliest of the start of rescue medication, or last dose of study medication including the cuts off allowances described in Section 0. For analyses "regardless of rescue and treatment discontinuation" all subjects' data will eligible for inclusion in the analyses.

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 ANCOVA, for Hb1Ac analysis of change from baseline, the measurement closest to the target date is assigned as the Week 52 measurement and will be used. If no Week 52 measurement is available (subject has discontinued before Week 52, or measurement not taken at Week 52, 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 52, their last post-baseline measurement taken prior to or on the date of the first dose of rescue medication will be used in the "prior to rescue medication" analyses.

For safety analyses based on "regardless of rescue medication", which exclude data after treatment discontinuation, all data will be eligible to be used until the prescribed cut off (see Section 0).

For urinary albumin to creatinine ratio (uACR) shift tables, if no week 52 data are available for a subject, then last available observation prior to week 52 is used.

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

The same methodology will be used in the ST+LT analyses at Week 156.

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 general model will be used:

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

(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,

α_k is the mean effect at time k

$(\alpha\ \tau)_{ik}$ is the interaction term between treatment group i and time k.

$(\alpha\ M_{baseline})_{ijk}$ 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.

If the models do not converge, then the spatial power covariance matrix will be used instead.

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.

Appropriate diagnostics will be performed to check model assumptions.

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.

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 treatment-by-baseline 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 for the primary efficacy endpoint will be assessed for subgroups, in the ST analysis, based on age group, female age group, gender, race, region, baseline HbA1c categories, and baseline eGFR categories, 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. Tests of the treatment-by-subgroup interaction will be assessed using contrasts of the treatment effect by subgroups at Week 52. 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 52.

7.1.3 Analysis of Covariance (ANCOVA)

In summaries of efficacy endpoints examining changes from baseline at Week 52 in the short-term analysis, and week 156 in the short-term and long-term analysis, 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:

$D_{t,ij} = intercept + \beta\,[Y_{0,ij}] + \tau_i + error_{ij}$

(Model 7.1.3.1)

where:

$D_{t,ij} = Y_{24,ij} - Y_{0,ij}$ = the Week 52 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 52 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*.

Intercept, β, 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.

Model assumptions will be checked.

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 n, mean, median, first quartile (Q1) and third quartile (Q3) and standard error (SE). Geometric mean and coefficient of variation (CV) will be also be provided for lipid panel results. 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 double-blind treatment period, or during the short-term and long-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. 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. 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) 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). For each treatment group, the probability of response is first modeled using a logistic regression model with baseline (e.g., baseline HbA1c) as a covariate.

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.

For composite response endpoints (glycemic control without hypoglycemia or weight gain), the proportion of subjects achieving targets will be summarized by treatment group without adjustment for baseline.

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 prior to Day 1 visit (between Week -1 and Day 1) and within a week of Week 52/Study Termination and Rescue or Early Treatment Termination. 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 post-prandial averages. The pre-prandial and post-prandial 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 pre-prandial and post-prandial 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.

Figure 7.1.3-1 is an example of calculation as described in the original publication of MAGE [3].

Figure 7.1.3-1: Blood glucose curve of a subject studied on intermediate-acting insulin regimen.

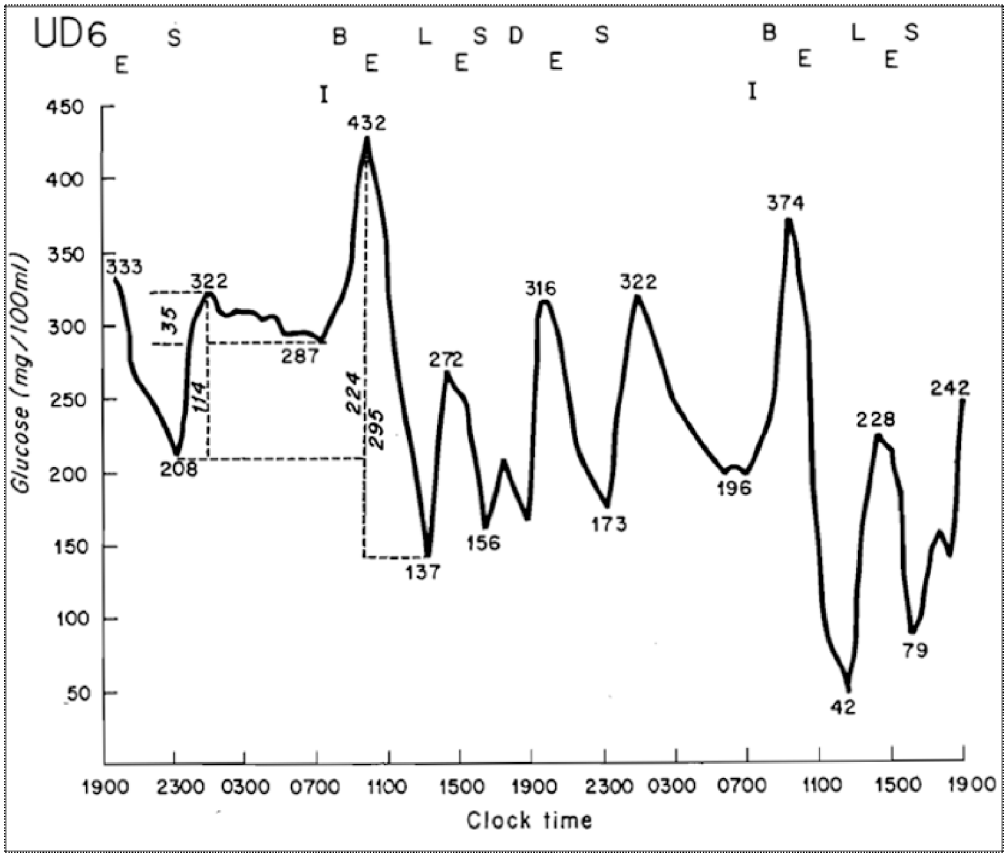


Figure from Service, et al., 1970 [3]

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 that despite that this 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 excursion during the 24 hours.

Phase V® is collecting the data 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 7.2-1. 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 ≥ 7.5% and ≤ 10.5%	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	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 double-blind/ ST + LT 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 double-blind/ST + LT 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	Randomized subjects who took no dose of metformin or dose of metformin outside dose range for ≥ 2 consecutive weeks in the short-term double-blind/ST + LT treatment period	Partial /Complete	Exclusion will start from the 14th consecutive day that metformin was either not taken or outside range of ≥ 1500. Subjects who were on treatment < 14 days and did not take metformin according to protocol will be excluded
8	Randomized subjects who receive no study medication or dosing for ≥ 2 consecutive weeks in the short-term double-blind/ST + LT 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 double-blind study medication will be excluded

Table 7.2-1: List of Relevant Protocol Deviations

RPD #	RPD Criteria	Exclusion Type	Exclusion Level
9	Randomized subjects who received incorrect study medication or dosing for ≥ 2 consecutive weeks in the short-term double-blind/ST + LT 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 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 double-blind/ST + LT treatment period.	Complete	

7.3 Study Population

7.3.1 Subject Disposition

The disposition of subjects for the screening period, lead-in period, and the 52-week short-term double-blind 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 double-blind treatment period will be tabulated and listed. The summary will be presented by randomized treatment group and overall.

The summary of status in the long-term treatment period 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, at the ST analysis. At the ST+LT analysis, the demographic and baseline characteristics will be summarized for the Randomized subjects in MRI-PDFF sub-study.

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, and hematuria (other disease characteristics, for ST+LT only) is listed in Table 7.3.2.D.

Table 7.3.2A: Demographic and Common Baseline Characteristics

Characteristic	Summarized as	Categories
Gender	Categorical	Male, Female
Age	Categorical and Continuous	< 65 yrs ≥ 65yrs - <75 ≥75 yrs
Female Age	Categorical	≤ 50 yrs > 50 yrs
Race	Categorical	White, Black or African American, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, Asian, Other
Ethnicity	Categorical	Hispanic/Latino, Non-Hispanic/Latino
Body Weight	Continuous	-
Height	Continuous	
Body Mass Index	Categorical and Continuous	< 25 kg/m^2 ≥ 25 kg/m^2 ≥ 27 kg/m^2 ≥ 30 kg/m^2
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

Table 7.3.2D: Hematuria (Other disease characteristics) (ST+LT only)

Characteristic	Summarized as	Categories
Hematuria n (%)	Categorical	Female (Yes, No) Male (Yes, No)

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 double-blind treatment period, at the ST analysis. At the ST+LT analysis the demographic and baseline characteristics will be summarized for the Randomized subjects in MRI-PDFF sub-study.

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 double-blind treatment period.

7.4 Extent of Exposure

7.4.1 Study and Rescue Medication

The extent of exposure to study medication during the 52-week short-term double-blind treatment period is defined as the difference between the last and the first dose of study medication of the short-term double-blind 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 double-blind treatment period and the short-term double-blind treatment period prior to rescue as well as regardless of 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, 169-252, 253-336, 337-371, $\geq$ 372 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). All rescue medication use during the 52-week short-term double-blind treatment period will be summarized and listed by treatment group.

The extent of exposure to study medication during the combined short- and long-term treatment periods is defined as the difference between the last dose of short-term double-blind or long-term treatment and the first dose of double-blind study medication plus 1 day. For a subject, when there are duplicate records of study medication exposure data within one of the study treatments dispensed (saxagliptin, dapagliflozin, or glimepiride) that have identical start and stop dates and number of pills taken, but different bottle/lot numbers, the average of the pill counts across that date range in the duplicate records will be used to count the number of pills taken for that study

treatment during that date range. The extent of exposure to study medication will be summarized using the Treated Subjects data set for the combined short- and long-term treatment period prior to rescue as well as regardless of rescue. The number and percent of subjects with an extent of exposure within pre-specified day ranges (1-90, 91-180, 181-360, 361-540, 541-720, 721-900, 901-1100 and > 1100 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 used during the combined short- and long-term 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 the short-term and combined short- and long-term treatment periods.

7.4.2 Concomitant Medications

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

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

Concomitant medications for both the short-term and combined short- and long-term treatment periods will be summarized using the Treated Subjects dataset by drug class and (generic) drug name. Summary tables, for allowed, as well as disallowed concomitant medications, by drug class and generic drug name will be generated.

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 the short-term and the combined short- and long-term treatment period for each treatment group. For each subject, percent compliance is defined as the total number of tablets taken divided by the total number of tablets that should have been taken, 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 double-blind treatment period dose recorded in the "Record of Study Medication" to the last short-term double-blind 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 double-blind treatment dose.

For the combined short- and long-term:

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 (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 treatment dose (following any adjustments for duplicate records as described in Section 7.4.1)..

7.5 Efficacy

7.5.1 Overall Efficacy Summary

All efficacy analyses will be performed using the Randomized Subjects data set. In addition, at the ST analysis, the primary efficacy analysis will be performed using the Evaluable Subjects data set if > 10% of subjects in any treatment group have relevant protocol deviations.

Unless otherwise specified, all analyses will be done using values prior to rescue/intensification of treatment or treatment discontinuation (plus a tolerance window after last dose as detailed in Section 0). For the endpoint of the proportion of subjects achieving a therapeutic glycemic response at Week 52/Week 156, all available data will be used, but subjects rescued prior to Week 52/Week 156 will be treated as non-responders.

The comparator (glimepiride plus metformin) will be tested against saxagliptin, in co-administration with dapagliflozin plus metformin, for the primary efficacy endpoint at the alpha = 0.05 level (two-sided). The secondary efficacy endpoints will then be tested sequentially, each comparison tested at the alpha = 0.05 level (two-sided). The following testing hierarchy (shown in order of the step-wise testing sequence) for treatment comparisons will be used:

- Mean change from baseline in HbA1c at Week 52 (primary endpoint).
- Mean change from baseline in total body weight at Week 52.
- Percent of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0%, at Week 52.
- Mean change from baseline in systolic blood pressure (SBP) at Week 52.
- Time to treatment intensification (addition of insulin or other glucose-lowering agent for rescue therapy or discontinuation for lack of glycemic control) during the 52-week double-blind treatment period.

If all the above five comparisons performed for the short-term analysis are statistically significant, the following two comparisons will be tested sequentially for the short-term plus long-term analysis:

- Time to treatment intensification (addition of insulin or other glucose-lowering agent for rescue therapy or discontinuation for lack of glycemic control) during the 156-week treatment period.
- Proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0%, at Week 156

Analyses of exploratory endpoints specifically designed for sub-studies (CGM and MRI-PDFF) 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 52 visit. The primary analysis of the change in HbA1c from baseline at Week 52 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 double-blind treatment period assessment. 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 (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 Section 7.1.2 in case of non-convergence of the preferred model or other issues. Data collected after the initiation of rescue medication will be excluded. Data collected outside the analysis window (defined in Section 8 of this SAP) after subjects discontinued double-blind 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 glimepiride will be calculated.

Sensitivity Analyses

The following sensitivity analyses will be carried out for the primary end point for the ST data/at week 52:

- 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 52 value is not available)
- Analyses to address missing values (described below in Section 7.5.4)
  - Analysis using multiple imputation return-to-baseline
  - Tipping point analysis

The following sensitivity analyses will be carried out for the ST+LT data/at Week 156 are: .

- Change from baseline in HbA1c at week 156, ANCOVA (LOCF) using the Randomized subject data set - completers not requiring glycemic rescue therapy, (defined as subjects who

took at least one dose of randomized study medication, completed the LT phase, and did not take any glycemic rescue therapy during the ST or LT treatment period)

- Change from baseline in HbA1c at week 156, regardless of rescue and treatment discontinuation, repeated measures analysis using the Randomized subject data set

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 mixed model repeated measures (MMRM) efficacy analyses, which assumes that data are missing at random (MAR), missing observations will not be imputed.

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. 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 return to its baseline value and not derive any benefit from treatment. 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.

Last observation carried forward (LOCF) will also be used to handle missing data as described in Section 7.1.1.3 for analyses such as sensitivity analyses of Hb1Ac (ANCOVA) at week 52 and 156.

7.5.3.1 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 52, 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 52. 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 as a fixed factor 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 Dapagliflozin + Saxagliptin + Metformin group to control group Glimipiride + Metformin. The above approach will be conducted using the Glimipiride + Metformin as the control group.

7.5.3.2 Tipping Point Analysis based on 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 to be 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 significant 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: Use the common Step 1 in Section 4.1.5.1 for partial imputation to get monotone missing 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 (Dapagliflozin + Saxagliptin + Metformin). The variables used as explanatory variables for imputation include Treatment, 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 is no longer rejected.

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

Step 4: Using a 2-sided Type I error level of 0.05, find the tipping point, i.e., the value of delta parameter for which the primary hypothesis is no longer rejected. In this study, we will compare Dapagliflozin + Saxagliptin + Metformin vs. Glimepiride + Metformin for each delta value. We will find the delta value at which the test of hypothesis is no longer rejected. A clinical interpretation about the plausibility of the assumptions underlying the tipping point will be provided.

7.5.4 Secondary Efficacy Analyses

In order to protect the overall type I error rate, the interpretation of the statistical significance of treatment comparisons for each secondary efficacy endpoint will be done using a step-wise procedure as outlined in Section 7.5.1

The ordered list of secondary endpoints for the sequential testing procedure is:

- Mean change from baseline in total body weight at Week 52.
- Percent of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0%, at Week 52.
- Mean change from baseline in SBP at Week 52.

- Time to treatment intensification (addition of insulin or other glucose-lowering agent for rescue therapy or discontinuation for lack of glycemic control) during the 52-week double-blind treatment period.
- Time to treatment intensification (addition of insulin or other glucose-lowering agent for rescue therapy or discontinuation for lack of glycemic control) during the 156-week treatment period.
- Proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0%, at Week 156

The analysis of change from baseline for total body weight and SBP will be performed using the same longitudinal repeated measures model as for the primary efficacy endpoint.

The proportion of subjects achieving a therapeutic glycemic response (defined as HbA1c < 7.0%) at Week 52 and Week 156 will be summarized by treatment group and compared between treatment groups using logistic regression, adjusting for baseline Hb1Ac. The 95% confidence intervals for the response rate within each treatment group, as well as for the difference in response rates between treatment groups, will be calculated with adjustment for baseline HbA1c. Subjects who are rescued, discontinued or have missed measurements at Week t (t = 52 and 156, respectively) will be considered as not achieving glycemic response.

Time to treatment intensification will be analyzed using a Cox proportional hazards model. Estimates of the hazard ratio and 95% confidence intervals will be provided. This analysis will be performed only when there are at least 10 events in each treatment group. Kaplan-Meier estimates will be calculated and plotted by treatment group. For the analysis during 52 / 156 weeks, all subjects will be censored at 52 / 156 week short-term / long-term treatment period if treatment intensification has not occurred by then. Subjects rescued at Week 52 / 156 visit will be counted as having an event for the analysis.

7.5.5 Exploratory Efficacy Analyses

No adjustments for multiplicity will be done for exploratory endpoints; only nominal p-values will be presented.

7.5.5.1 Proportion of Subjects Achieving Composite Response Targets

The proportion of subjects achieving composite response targets will be summarized by treatment group without adjustment for baseline for the following composite response endpoints:

- Proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0%, without any hypoglycemia at Week 52.
- Proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0%, without any hypoglycemia and weight gain at Week 52.
- Proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0%, without any hypoglycemia at Week 156.
- Proportion of subjects achieving a therapeutic glycemic response, defined as HbA1c < 7.0%, without any hypoglycemia and weight gain at Week 156.

7.5.5.2 Mean Change from Baseline at Week t

The analyses of change from baseline in exploratory endpoints including HbA1c, FPG, uACR, fasting plasma glucose, body weight, and SBP at Week 52 and Week 156 (if applicable) will be performed using the longitudinal repeated measures model as described in Section 7.1.2. The analysis of change from baseline for metabolic measures (glucagon, C-peptide, pro-insulin, insulin, iHOMA2), average glucose values and post-prandial glucose values measured by 6-point SMBG profiles (ST analysis only) will be performed using ANCOVA model as described in Section 7.1.3. The statistical model assumptions will be checked, especially because uACR results are often skewed.

7.5.5.3 Change from Baseline for Lipids

The change from baseline in Triglycerides, Total Cholesterol, LDL and HDL parameters during the short-term and long-term treatment periods will be summarized for each treatment group at visits 52 and 156 only. The data will be summarized by arithmetic mean, SD, geometric means and coefficient of variation (CV), minimum, maximum, Q1, Median and Q3 values.

7.5.5.4 Mean Change from baseline in endpoints for sub-study

In the CGM sub-study, CGM will be performed for periods of 7 days at baseline and at Week 52 of treatment. Mean amplitude of glucose excursions (MAGE) for a 24-hr period is calculated in the way defined in Section 7.1.3. The average of the 7 MAGE values at baseline or at Week 52 will be used for analyses. Analyses of the mean change from baseline in MAGE at Week 52 will be performed using the ANCOVA model based on LOCF data.

In the MRI-PDFF sub-study, the mean change from baseline in visceral and subcutaneous adipose tissue volume and percent hepatic lipid content as assessed by MRI at Week 52 will be performed using the ANCOVA model based on LOCF data.

The analysis of the mean change from baseline in visceral and subcutaneous adipose tissue volume and percent hepatic lipid content as assessed by MRI at the Week 122 CRF visit will be performed using the same longitudinal repeated measures model as described in Section 7.1.2.

MRI values from an earlier visit are used instead of values at Week 52/122 for patients who are rescued or prematurely discontinue prior to Week 52/122. MRIs from week 52/122 could have been scheduled up to 4 weeks after the Rescue or Premature Treatment Discontinuation Date.

The analysis of percent hepatic lipid content will be adjusted for baseline weight.

Liver fat fraction (%) will also be summarized in a shift tables from baseline to week 52 (in the ST summary), and weeks 52 and 122 (in the ST+LT summary), using categories of >=5.5% and <5.5%.

For MRI laboratory data (visceral adipose tissue, subcutaneous adipose tissue and liver fat function) obtained from day 2, up to and including the last short-term double-blind dosing date + 27 days (see Table8.4M) will be considered as obtained during the short-term double-blind treatment period. MRI laboratory data obtained from the day after the last study medication up to the last visit date of the follow-up period will be considered as obtained during the follow-up

period. MRI laboratory data obtained from the first day of long-term treatment up to including the last day of long-term medication (excluding assessments that were within 27 days from last day of ST double-blind treatment and were assigned to Week 52 analysis visit) will be considered as obtained during the combined short- and long-term treatment periods.

7.5.6 Summary of Primary Efficacy Endpoint within Subgroups

Treatment effects will be assessed for subgroups based on age group, female age group, gender, race, region, baseline HbA1c categories, and baseline eGFR categories 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. 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 52.

7.5.7 Pharmacokinetic Analyses

Not applicable

7.5.8 Biomarker Analyses

The change from baseline in C-Reactive protein and N-Terminal ProB-type Natriuretic Peptide parameters during the short-term and long-term treatment period will be summarized descriptively for each treatment group at week 52 and 156 only.

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, 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. The outcomes research analyses will be summarized as a separate report by Phase V®.

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, hypoglycemia, 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 52-week double-blind period (at the time of the primary endpoint analysis) and then again at the final database lock (156 weeks) on the combined short- and long-term treatment periods.

Additional sensitivity analyses for adverse events and laboratory marked abnormalities will be performed excluding data after rescue for the short-term double-blind treatment period. The primary analyses of events of hypoglycemia will be performed excluding data after rescue.

7.6.1 Adverse Events

Adverse events analyses will be performed for the short-term and combined short- and long-term 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 8.8.

In summaries by SOC and PT, AEs will be sorted by international order for system organ class (SOC) and by decreasing frequency of preferred term within each SOC for Dapagliflozin + Saxagliptin + Metformin group. In summaries by PT, AEs will be sorted by decreasing frequency within PT for Dapagliflozin + Saxagliptin + Metformin group.

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.

Extra details of surgical or spontaneous amputation are captured on CRF pages separate to the AE pages. Amputation AE/SAE will be included in the summaries of AE/SAEs. In addition, separate summaries will be provided including AEs leading to surgical or spontaneous amputation, an overview of adverse events leading to surgical or spontaneous/non-surgical amputations, and contributing factors and conditions that triggered any surgical or spontaneous/non-surgical amputations.

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 double-blind treatment period. 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 double-blind treatment period. For the analyses of the short-term double-blind treatment period, the summary of AEs and analyses by SOC and PT will be performed using the data prior to rescue as well as the data regardless of rescue. 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 double-blind treatment up to and including 4 days and 30 days respectively, after the last dose date in the short-term double-blind 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 double-blind treatment period.

In addition, the following summaries will be provided for the short-term double-blind 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, ≥ 65, and ≥ 75 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 the combined short- and long-term treatment period. AEs and SAEs with an onset from Day 1 of short-term double-blind treatment up to and including 4 days and 30 days respectively, after the last dose date in the short- or

long-term treatment period will be included in summaries of the combined short- and long-term 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 used to identify deaths for 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 the short-term treatment period and listing of all deaths that occur during the study will be produced for the combined short- and long-term treatment period.

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 double-blind treatment up to and including 30 days after the last dose date in the short-term double-blind 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 double-blind treatment period.

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- or long-term treatment period will be considered as occurring during the combined short- and long-term treatment periods.

SAEs occurring during the short-term or combined short- and long-term treatment periods 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.

At the time of the ST+LT analysis two SAE listings will be produced: A listing of all SAEs for the short-term treatment period will be produced, displaying all SAEs (including pre-treatment events) that occurred prior to or on the start date of long-term treatment period, as well as listing of all SAEs for the combined short- and long-term treatment period 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 double-blind 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 during the short-term double-blind treatment period 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.

AEs with an onset during the combined short- and long-term 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 the combined short- and long-term treatment period, no upper cutoff day windows (i.e., 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 other 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 double-blind treatment up to and including 4 days and 30 days, respectively, after the last dose date in the short-term double-blind 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 double-blind treatment period.

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- or long-term treatment period will be considered as occurring during the combined short- and long-term 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 double-blind treatment up to and including 4 (30 days in case of SAEs) days after the last dose date in the short-term double-blind 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 double-blind treatment period. Hypoglycemic events with an onset from Day 1 of the short-term double-blind treatment up to and including 4 (30 days in case of SAEs) 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."[2]

Documented symptomatic hypoglycemia: **"**An event during which typical symptoms of hypoglycemia are accompanied by a measured plasma glucose concentration ≤ 70mg/dl (3.9mmol/l)."[2]

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."[2]

It is not medically possible for an event to be considered to be both asymptomatic hypoglycemia as well as severe hypoglycemia. In the ST+LT analysis, if an event is classified as both severe and asymptomatic, it will be considered only to be severe hypoglycemia in the summary tables. If an event is classified as both severe and as documented symptomatic hypoglycemia, even if unlikely but not deemed medically implausible, the event will be included in the summary tables under both classifications.

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."[2]

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 patients 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."[2]

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. 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 double-blind treatment period, and the short-term and long-term treatment period. . For short-term double-blind period analyses, the only upper cutoff date is 4 days (30 days in case of SAEs) following the date of the last dose of short-term double-blind study medication or the start of the long-term treatment period, whichever comes first. For the short-term and long-term treatment period, the upper cutoff date is 4 (30 days in case of SAEs) days following the date of the last dose of study medication.

Hypoglycemic events with an onset during the short-term double-blind treatment period and leading to discontinuation of study medication will be summarized by treatment group. A similar summary will be produced for the short-term and long-term treatment period. When summarizing hypoglycemic events leading to discontinuation, no upper cutoff day windows are applied.

A summary of the incidence of confirmed hypoglycemia, defined as blood glucose $\leq$ 70mg/dL (3.9 mmol/L) recorded in the hypoglycemia module of the CRF, will be provided. Additionally, a summary will be provided for subjects who have at least one episode of confirmed hypoglycemia.

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 double-blind treatment period up to the start date of the long-term treatment period, and separately for the short-term and long-term treatment period.

7.6.5.2 Confirmed Adjudicated Cardiovascular Adverse Events

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

In this study, only events of heart failure requiring hospitalization will be adjudicated.

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 this study, the following hepatic events will be adjudicated:

- AST and/or ALT >3X ULN and TB >2X ULN (within 14 days of the AST and/or ALT elevation)

- AST and/or ALT >10X ULN
- Hepatic event AE/SAE occurring within 30 days before subject death

In addition, a listing of all adjudicated hepatic adverse events by subjects will be provided.

7.6.5.4 Confirmed Adjudicated DKA Events

The number and percentage of subjects with confirmed DKA events (as determined by the adjudication committee) will be summarized by preferred term (as listed in Appendix B of the DKA Adjudication Charter v3.0 dated 17 July 2018) and treatment group.

7.6.5.5 Other AEOSIs

Other 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 listing of AEOSI will include all data regardless of use of rescue medication:

- AEs of genital infection by recurrence
- AEs of genital infection by gender
- AEs of Urinary Tract Infection (UTI) by recurrence
- AEs of UTI by gender
- AEs of severe cutaneous adverse reactions (SCAR)
- AEs of renal impairment/failure
- AEs of volume depletion
- AEs of fracture
- AEs of liver injury
- AEs of hypersensitivity reactions
- AEs of decreased lymphocyte count
- AEs of pancreatitis
- AEs of Breast, bladder, prostate and pancreatic cancer
- AEs of cardiac failure
- AEs of decreased thrombocyte count

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 short- and long-term 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 (Glucagon, proinsulin, hematology and chemistry assessments, UACR, C-Reactive Protein (biomarker), N Terminal ProB Type Natriuretic Peptide

(biomarker) 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 double-blind 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 double-blind 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 short-term treatment up to and including 4 days (30 days for liver function laboratory tests) after the last dose of short- or long-term medication will be considered as obtained during the combined short- and long-term 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 3 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.

Because the mechanism of action of dapagliflozin is known to impact urine glucose levels, urinalysis results of glucose will not be summarized. All laboratory evaluations performed by central laboratories will be included in summary tables,. All laboratory tables for the CSRs will be produced in both conventional 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 double-blind treatment period and the combined short- and long-term treatment periods will be summarized by treatment group. In the short-term double-blind 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 short-term treatment period and the combined short- and long-term 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 x (Scr)-1.154 x(Age)-0.203 x (0.742 if female) x (1.210 if African-American)
- creatinine kinase (creatine phosphokinase) (CK)
- creatinine, serum (S_{cr})
- electrolytes - sodium, potassium, chloride, magnesium and calcium
- total protein, serum
- inorganic phosphorus
- urinary albumin to creatinine ratio
- creatinine clearance

Change from baseline in categorical variables will be provided using shift tables. For urinary albumin to creatinine ratio, the shift table will be derived from laboratory assessments after the date of first dose of short-term double-blind study medication up to and including 4 days following the date of last dose of short-term and long-term study medication, or up to the end date of the short-term and long-term treatment period, whichever comes first. Week 156 value in the shift table is chosen for Week 156 using the day range of 1003 to last day of long-term treatment as in SAP table8.4H, regardless of rescue medication, and up to and including 4 days following the date of last dose of short-term + long term study medication. If no week 156 value available then last available post-baseline observation prior to week 156 should be used (imputed) for Week 156 in the shift table (i.e., LOCF)

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 double-blind 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 double-blind treatment period.

Vital signs data obtained after the start of study medication dosing up to and including 4 days after the last short- or long-term dosing date will be considered as obtained during the combined short- and long-term 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) for BMI, waist circumference, sitting SBP, DBP and HR, only.

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 52 and at Week 156 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 short-term and the combined short- and long-term treatment periods.

8 CONVENTIONS

Data collected outside the analysis window (defined in Section 8.4) after subjects discontinued double-blinded 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 non-missing 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.
- For MRI visits, values from an earlier visit will be used instead of values at Week 52/122 for patients who were rescued or prematurely discontinue prior to Week 52/122. MRIs from week 52/122 could have been scheduled up to 4 weeks after the Rescue or Premature Treatment Discontinuation Date.

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 double-blind study drug, observations may not contribute to summaries/analyses if they are measured after the last dose of the double-blind study medication given during the treatment periods, as indicated below:

For efficacy parameters:

- Lipids, glucagon, proinsulin, uACR, biomarkers (C-reactive protein, N-terminal ProB-type Natriuretic Peptide) and SBP will be summarized only if measured on or prior to the 4th day after the last study drug treatment date.
- FPG, fasting C-peptide, fasting insulin, HOMA assessments 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, BMI, waist circumference 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, including hypoglycemia, reported as 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 30th day after the last study drug treatment date.
- All other safety events (non-serious AEs, non-serious hypoglycemia, etc.) and measurements (safety lab, vital signs, etc.) will be included in the summaries only if occurred/measured on or prior to the 4th day after the last study drug treatment date.

8.4 Longitudinal Assessments

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

Visit windows are specified below.

Table 8.4 A: Vital Signs and FPG Visit Windows for the Short-term Analyses

Visit	Treatment Period	Target Day	Day Range
Week 3	ST	22	2-35
Week 6	ST	43	36-64
Week 12	ST	85	65-99
Week 16	ST	113	100-155
Week 28	ST	197	156-239
Week 40	ST	281	240-323
Week 52	ST	365	324 to last day of ST double-blind treatment

Table 8.4 B: Vital Signs (except BMI and Waist Circumference), FPG Visit Windows for the Short and Long-term Analyses

Visit	Treatment Period	Target Day	Day Range
Week 3	ST	22	2-35
Week 6	ST	43	36-64
Week 12	ST	85	65-99
Week 16	ST	113	100-155
Week 28	ST	197	156-239
Week 40	ST	281	240-323
Week 52	ST	365	324 to last day of ST double-blind treatment
Week 65	LT	456	First day of long-term treatment -501
Week 78	LT	547	502-592
Week 91	LT	638	593-683
Week 104	LT	729	684-774
Week 117	LT	820	775-865

Table 8.4 B: Vital Signs (except BMI and Waist Circumference), FPG Visit Windows for the Short and Long-term Analyses

Visit	Treatment Period	Target Day	Day Range
Week 130	LT	911	866-956
Week 143	LT	1002	957-1047
Week 156	LT	1093	1048 to last day of long-term treatment

Table 8.4 C: BMI - Visit Windows for the Short and Long-term Analyses

Visit	Treatment Period	Target Day	Day Range
Week 28	ST	197	2-281
Week 52	ST	365	282 to last day of ST double-blind treatment
Week 65	LT	456	First day of long-term treatment -501
Week 78	LT	547	502-592
Week 91	LT	638	593-683
Week 104	LT	729	684-774
Week 117	LT	820	775-865
Week 130	LT	911	866-956
Week 143	LT	1002	957-1047
Week 156	LT	1093	1048 to last day of long-term treatment

Table 8.4 D: Waist Circumference - Visit Windows for the Short and Long-term Analyses

Visit	Treatment Period	Target Day	Day Range
Week 28	ST	197	2-281
Week 52	ST	365	282 to last day of ST double-blind treatment
Week 65	LT	456	First day of long-term treatment -638
Week 117	LT	820	639-956
Week 156	LT	1093	957 to last day of long-term treatment

Table 8.4 E: HbA1c- Visit Windows for the Short-term Analyses

Visit	Treatment Period	Target Day	Day Range
Week 12	ST	85	2-99
Week 16	ST	113	100-155
Week 28	ST	197	156-239

Visit	Treatment Period	Target Day	Day Range
Week 40	ST	281	240-323
Week 52	ST	365	324 to last day of ST double-blind treatment

Table 8.4 F: HbA1c-- Visit Windows for the Short- and Long-term Analyses

Visit	Treatment Period	Target Day	Day Range
Week 12	ST	85	2-99
Week 16	ST	113	100-155
Week 28	ST	197	156-239
Week 40	ST	281	240-323
Week 52	ST	365	324 to last day of ST double-blind treatment
Week 65	LT	456	First day of long-term treatment -501
Week 78	LT	547	502-592
Week 91	LT	638	593-683
Week 104	LT	729	684-774
Week 117	LT	820	775-865
Week 130	LT	911	866-956
Week 143	LT	1002	957-1047
Week 156	LT	1093	1048 to last of LT treatment

Table 8.4 G: Safety Lab and Lipid Panel (except HbA1c, FPG, Biomarkers and HOMA) and uACR - Visit Windows for the Short-term Analyses

Visit	Treatment Period	Target Day	Day Range
Week 6	ST	43	2-64
Week 12	ST	85	65-141
Week 28	ST	197	142-281
Week 52	ST	365	282 to last day of ST double-blind treatment

uACR data collected at Weeks 28 and 52 also used these visits, target days and day ranges for the MMRM analysis; lipid panel data collected at Week 52 also used these visits, target days and day ranges.

Table 8.4 H: Safety Lab and Lipid Panel (except HbA1c, FPG, Biomarkers and HOMA) and uACR - Visit Windows for the Short and Long-term Analyses

Visit	Treatment Period	Target Day	Day Range
Week 6	ST	43	2-64

Table 8.4 H: Safety Lab and Lipid Panel (except HbA1c, FPG, Biomarkers and HOMA) and uACR - Visit Windows for the Short and Long-term Analyses

Visit	Treatment Period	Target Day	Day Range
Week 12	ST	85	65-141
Week 28	ST	197	142-281
Week 52	ST	365	282 to last day of ST double-blind treatment
Week 78	LT	547	First day of long-term treatment -638
Week 104	LT	729	639-820
Week 130	LT	911	821-1002
Week 156	LT	1093	1003 to last day of long-term treatment

uACR data collected at Weeks 28, 52, 117 and 156 also used these visits, target days and day ranges for the MMRM analysis; lipid panel data collected at Weeks 52 and 156 also used these visits, target days and day ranges.

Table 8.4 I: HOMA, Biomarkers - Visit Windows for the Short-term Analyses

Visit	Treatment Period	Target Day	Day Range
Week 52	ST	365	Day 2 to last day of ST double-blind treatment

Table 8.4 J: HOMA, Biomarkers – Visit Windows for the Short and Long-term Analyses

Visit	Treatment Period	Target Day	Day Range
Week 52	ST	365	Day 2 to last day of ST double-blind treatment
Week 156	LT	1093	First day of long-term treatment to last day of long-term treatment

Table 8.4 K: ECG - Visit Windows for the Short-term Analyses

Visit	Treatment Period	Target Day	Day Range
Week 52	ST	365	Day 2 to last day of ST double-blind treatment

Table 8.4 L: ECG – Visit Windows for the Short and Long-term Analyses

Visit	Treatment Period	Target Day	Day Range
Week 52	ST	365	Day 2 to last day of ST double-blind treatment
Week 104	LT	729	First day of long-term treatment - 911
Week 156	LT	1093	912 - last day of long-term treatment

Table 8.4 M: MRI - Visit Windows for the Short-term Analyses

Visit	Treatment Period	Target Day	Day Range
Week 52	ST	365	Day 2 to (last day of ST double-blind treatment + 27 days)

Table 8.4 N: MRI – Visit Windows for the Short and Long-term Analyses

Visit	Treatment Period	Target Day	Day Range
Week 52	ST	365	Day 2 to (last day of ST double-blind treatment + 27 days)
Week 122	LT	855	First day of long-term treatment to last day of long-term treatment (exclude assessments that were within 27 days from last day of ST double-blind treatment and were assigned to Week 52 analysis visit)

8.5 Post-Treatment Efficacy Observations

While short-term double-blind 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 double-blind study medication (up to and including the start of long-term treatment period, where applicable) as indicated below:

- Lipids, glucagon, proinsulin, uACR, biomarkers (C-reactive protein, N-terminal ProB-type Natriuretic Peptide) and SBP will be summarized only if measured on or prior to the 4th day after the last study drug treatment date.
- FPG, fasting C-peptide, fasting insulin, HOMA assessments 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 short- and long-term 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

8.9 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.

8.10 Percent Compliance Calculation

See Section 7.4.3.

8.11 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 double-blind period will be included in the short-term CSR while all the analyses specified for the combined short- and long-term treatment period will be included in the Final CSR.

10 CHANGES OF ANALYSIS FROM PROTOCOL

Changes from Protocol prior to short-term data base lock are summarized in the following table:

	Changes	Reason for changes
	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	To avoid error in summarizing data based on glucose values which were not collected using consistent methods.

Changes	Reason for changes
irrespective of the method using which glucose data were collected.	
MI sensitivity analyses for primary endpoint were added.	To be consistent with other saxagliptin plus dapagliflozin studies.

11 REFERENCES

- 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.
- 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.
- Service, F.J., Mean amplitude of glycemic excursions, a measure of diabetic instability. Diabètes. 1970, Sep ; 19, 644-55.
- Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. Journal of American Statistical Association. 1958; 53:457-481.
- Greenwood M. The natural duration of cancer. Reports on public health and medical subjects. 1926; 33:1-26.
- 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.

12 SUMMARY OF CUTOFFS FOR ANALYSES WHICH EXCLUDE RESULTS AFTER TREATMENT DISCONTINUATION

Assessment collected more than x days after last dose in the ST+LT treatment period are excluded from analyses		
Number of days		Assessment/measurement
1	Efficacy	fasting plasma glucose fasting C-peptide fasting insulin spot urinary glucose/creatinine ratio HOMA assessments Beta cell function for C-Peptide Beta cell function for insulin Insulin sensitivity for C-peptide Insulin sensitivity for insulin Insulin rstn for C-peptide Insulin rstn for insulin
4	Efficacy	SBP glucagon proinsulin urinary albumin to creatinine ratio biomarkers C-reactive protein N-terminal ProB-type Natriuretic Peptide lipids (Total-C, LDL-C, HDL-C, TG)
4	Safety	non-serious adverse events including non-serious hypoglycemic events vital signs hematology labs clinical chemistry labs
8	Efficacy	HbA1c body weight BMI waist circumference
30	Safety	SAEs (including hypoglycemic reported as SAEs) liver function tests ALT AST alkaline phosphatase total bilirubin

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)
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 CV181-365 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[3] is given in Sections 2.1.1 – 2.1.3 below. A double asterisk indicates those scales and subscales intentionally omitted from protocol CV181-365.

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 CV181365:

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 CV181-365 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
Calcium	mg/dL	< 7.5 mg/dL	≥ 1 mg/dL from ULN and ≥ 0.5 mg/dL from PreRx CA
Magnesium	mEq/L	< 1 mEq/L	> 4 mEq/L
PO4 (Phosphate)	mg/dL	Age 17-65: ≤ 1.8 mg/dL Age ≥ 66: ≤ 2.1 mg/dL	Age 17-65: ≥ 5.6 mg/dL Age ≥ 66: ≥ 5.1 mg/dL
Urine			
UACR (Urinary Albumin to	mg/g		> 1800 mg/g

Elevated AT (ALT and/or AST) and Total Bilirubin

The following three criteria will be summarized in examination of elevated AT (ALT and/or

AST) and total bilirubin:

- (AST or ALT > 3X ULN) and (Bilirubin > 1.5X ULN within 14 days on or after AT elevation)
- (AST or ALT > 3X ULN) and (Bilirubin > 2X ULN within 14 days on or after AT elevation)
- (AST or ALT > 3X ULN) and ([Bilirubin > 2X ULN and no ALP ≥ 2X ULN) within 14 days on or after AT elevation])

APPENDIX 4 GEOGRAPHIC REGIONS

Geographic Region	Countries
North America	United States
Latin America	Mexico
Europe	Russian Federation, Poland, Sweden, Germany United Kingdom, Czech Republic, Hungary, Romania