AstraZeneca

Statistical Analysis Plan
Short-term double-blind treatment period

Study Code	MB102230
Edition Number	2.0
Date	22Sep2017

A Multicenter, Randomized, Double-Blind, Placebo-controlled, Parallel Group, Phase 3 Study to Evaluate the Efficacy and Safety of Dapagliflozin as an Add-on to Insulin Therapy in Subjects with Type 1 Diabetes Mellitus

A Multicenter, Randomized, Double-Blind, Placebo-controlled, Parallel Group, Phase 3 Study to Evaluate the Efficacy and Safety of Dapagliflozin as an Add-on to Insulin Therapy in Subjects with Type 1 Diabetes Mellitus

Study Statistician

PPD

PPD

10/2/2017
Date

A Multicenter, Randomized, Double-Blind, Placebo-controlled, Parallel Group, Phase 3 Study to Evaluate the Efficacy and Safety of Dapagliflozin as an Add-on to Insulin Therapy in Subjects with Type 1 Diabetes Mellitus

Global Product Statistician PPD PPD

10/2/2017
Date

TABLE OF CONTENTS PAGE

TITLE PAGE 1
SIGNATURE OF STUDY STATISTICIAN 2
SIGNATURE OF GLOBAL PRODUCT STATISTICIAN 3
TABLE OF CONTENTS 4
LIST OF ABBREVIATIONS 9
AMENDMENT HISTORY 12
1. STUDY DETAILS 13
1.1 Study objectives 13
Primary Objective 13
Secondary Objectives: 13
Efficacy 13
Safety 14
Other/Exploratory Objectives: 14
1.2 Study design 15
1.3 Number of subjects 16
2. ANALYSIS SETS 17
2.1 Definition of analysis sets 17
2.1.1 Enrolled subjects dataset 17
2.1.2 Lead-in subjects dataset 17
2.1.3 Full analysis dataset 17
2.1.4 Evaluable subjects dataset 17
2.1.5 Treated subjects dataset 17
2.2 Violations and deviations 18
3. PRIMARY AND SECONDARY VARIABLES 19
3.1 Primary efficacy variable 19
3.2 Secondary efficacy variables 19
3.3 Safety variables 20
3.4 Exploratory variables 20
4. ANALYSIS METHODS 21
4.1 General principles 21
4.1.1 Objectives and hypotheses 21

4.1.2 Definitions....22
4.1.2.1 Short-term treatment period, long-term extension period, baseline value....22
4.1.2.2 Change from baseline....23
4.1.2.3 Percent change from baseline....23
4.1.2.4 Insulin up-titration more than 25% than baseline....25
4.1.2.5 Last Observation Carried Forward (LOCF)....25
4.1.2.6 Average daily value of continuous glucose monitoring system readings....26
4.1.2.7 Percent of continuous glucose monitoring system readings falling within a range....26
4.1.2.8 Standard deviation of continuous glucose monitoring system readings....26
4.1.2.9 Mean daily glucose of 6-point self monitored blood glucose....26
4.1.2.10 Post-prandial glucose values of continuous glucose monitoring readings and of self monitored blood glucose (SMBG)....27
4.1.2.11 Mean Amplitude of Glucose Excursions (MAGE) of continuous glucose monitoring system readings....27
4.1.3 Descriptive summaries of continuous variables....29
4.1.4 Descriptive summaries of categorical variables....29
4.1.5 Analysis of Covariance (ANCOVA)....29
4.1.5.1 ANCOVA model for change from baseline....29
4.1.5.2 ANCOVA model for percent change from baseline....30
4.1.6 Longitudinal repeated measures analysis....31
4.1.6.1 Longitudinal repeated measures analysis for outcomes with single measurement per week t....31
4.1.6.2 Longitudinal repeated measures analysis for outcomes with multiple measurement per week t....32
4.1.6.3 Longitudinal repeated measures analysis model assumption assessment....32
4.1.7 Kaplan-Meier curve and estimates for time-to-event analyses....33
4.1.8 Proportion of subjects with a pre-defined characteristic at week t....34
4.2 Analysis methods....34
4.2.1 Study population....34
4.2.1.1 Subject disposition....34
4.2.1.2 Demographic and other baseline characteristics....34
4.2.1.3 Specific and general disease histories....37
4.2.2 Extent of exposure....38
4.2.2.1 Study medication....38
4.2.2.2 Interruption of study medication....38
4.2.2.3 Current and concomitant medications....38
4.2.2.4 Measurements of treatment compliance....39
4.2.3 Efficacy Analysis....39
4.2.3.1 Primary efficacy analysis....39
4.2.3.2 Secondary and exploratory efficacy analysis....43
Percent change from baseline in total daily insulin dose....43
Percent change from baseline in body weight....44
Change from baseline in continuous glucose monitoring (CGM) parameters....45

Proportion of subjects achieving glycemic response 47
Change from baseline in fasting plasma glucose 47
Change from baseline in 6-point self monitored blood glucose (SMBG) parameters. 48
Change from baseline in seated systolic blood pressure 48
Time to insulin up-titration more than 25% from baseline 49
4.2.3.3 Outcomes Research Analyses 49
Health status measured with EQ-5D-3L questionnaire 49
Treatment Satisfaction measured with Diabetes Treatment Satisfaction Questionnaire status questionnaire 49
Hypoglycemia Fear Survey II Worry Subscale 50
4.2.4 Safety 50
4.2.4.1 Adverse events 50
4.2.4.2 Adverse events leading to discontinuation of study medication 53
4.2.4.3 Confirmed adjudicated cardiovascular adverse events 53
4.2.4.4 Adverse events of special interest 53
4.2.4.5 Laboratory evaluation 59
4.2.4.6 Vital signs 61
4.2.4.7 Physical examination 61
4.2.4.8 Electrocardiograms 62
4.2.4.9 Pregnancy test results 62
5. INTERIM ANALYSES 62
6. CONVENTIONS 62
6.1 Duration of type 1 diabetes 62
6.2 Missing and multiple measurements 63
6.3 Longitudinal assessments 63
6.4 Post-treatment observations 65
6.5 Assignment of doses to adverse events and laboratory assessments 65
6.5.1 Treatment at onset of an adverse event 67
6.5.2 Treatment at the time of a laboratory assessment 67
6.6 Concomitant medications 67
6.7 Counting rules for adverse events 67
6.7.1 At subject level 67
6.7.2 At event level 68
6.8 Fasting state 68
6.9 Percent compliance calculation 68
6.10 Strip sign for selected laboratory data 69
6.11 United States Conventional Units and Standard International Units for Laboratory Data 69

6.12 Capillary glucose readings of plasma glucose values and whole blood glucose values 69
6.13 Laboratory evaluations 69
7. CHANGES OF ANALYSIS FROM PROTOCOL 69
8. REFERENCES 70
9. APPENDIX 71
9.1 List of significant protocol deviations 71
9.2 Geographic regions 74
9.3 Units and marked abnormality criteria for safety laboratory variables 74
9.4 Summary of updates 77

LIST OF TABLES

Table 1 List of Relevant Protocol Deviations 18
Table 2 Formulae Used to Transform Back onto the Original Scale the Estimates of Percent Change from Baseline Analyzed as Geometric Mean 23
Table 3 Contrast Coefficients for the treatment-by-baseline interaction for week 24 estimates 32
Table 4 Demographic and Baseline Characteristics 35
..... 35
Table 5 Diabetes-Related Baseline Characteristics 36
Table 6 Baseline Renal Function Characteristics 37
Table 7 Subgroup Analyses of Primary Efficacy Endpoint for HbA1c 41
Table 8 Contrast Coefficients for the Example Gender-by-treatment Interaction 43
Table 9 Visit Windows for Short-term Double-blind Treatment Period Up to Week 24 63
Table 10 Visit Windows for Short-term Double-blind Treatment Period Up to Week 24 for CGM parameters, Total daily Insulin and 6-point SMBG 64

LIST OF FIGURES

Figure 1 Study Schematic 16

Figure 2 Blood glucose curve of a subject studied on intermediate-acting insulin regimen. 28

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
A1C	Glycosylated Hemoglobin
ACE-I	Angiotensin converting enzyme inhibitor
ADA	American Diabetes Association
AE	Adverse event
AST	Aspartate aminotransferase
ALT	Alanine aminotransferase
ALP	Alkaline phosphatase
ANCOVA	Analysis of covariance
ARB	Angiotensin II receptor blocker
BMI	Body Mass Index
BP	Blood pressure
BUN	Blood urea nitrogen
C	Cystatin
CEC	Clinical Event Committee
CGM	Continuous Glucose Monitoring
CHF	Congestive Heart Failure
CK	Creatine kinase
CrCl	Creatinine clearance
CRF	Case Report Form
CSII	Continuous subcutaneous insulin infusion
CVD	Cardiovascular disease
DBP	Diastolic blood pressure
DKA	Diabetic Ketoacidosis
DTSQ	Diabetes Treatment Satisfaction Questionnaire
DTSQs	Diabetes Treatment Satisfaction Questionnaire status version
DTSQc	Diabetes Treatment Satisfaction Questionnaire change version
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
eGFR	Estimated glomerular filtration rate
EQ-5D-3L	Euro Quality of life 5 dimensions 3 levels

Abbreviation or special term	Explanation
FFA	Free fatty acids
FPG	Fasting plasma glucose
HCO3	Bicarbonate
HCT	Hematocrit
HDL-C	High density lipoprotein-cholesterol
HFS-II	Hypoglycemia Fear Survey II
HIV	Human immunodeficiency virus
LDL-C	Low density lipoprotein-cholesterol
LOCF	Last observation carried forward
MA	Marked abnormality value
MAGE	Mean amplitude of glucose excursion
MAR	Missing at random
MCH	Mean cell hemoglobin
MCHC	Erythrocyte mean corpuscular hemoglobin concentration
MCV	Erythrocyte mean corpuscular volume
MDI	Multiple daily injections
MDRD equation	The formula is named after the Modification of Diet in Renal Disease study in the USA
MedDRA	Medical Dictionary for Regulatory Activities
MNAR	Missing not at random
MPDG	Mean post-prandial daily glucose
NYHA	New York Heart Association
OGTT	Oral Glucose Tolerance Test
PO4	Phosphate
PT	Preferred Term
PTH	Parathyroid hormone
RBC count	Red blood cell count
RDW	RBC distribution width-CV
RPD	Relevant protocol deviation
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SBP	Systolic blood pressure

Abbreviation or special term	Explanation
SCr	Serum creatinine
SD	Standard deviation
SI	International System of Units
SMBG	Self-monitored blood glucose
SMQ	Standard MedDRA Query
SOC	System Organ Class
SU	Sulphonyl urea
TB	Total bilirubin
TC	Total cholesterol
TG	Triglycerides
TZD	Thiazolidinedione
T1DM	Type 1 Diabetes Mellitus
UACR	Urine albumin-to-creatinine ratio
ULN	Upper limit of normal
WBC count	White blood cell count
WOCBP	Woman of childbearing potential

AMENDMENT HISTORY

Date	Brief description of change
	Original issue based on MB102 T1D Core Statistical Analysis Plan

1. STUDY DETAILS

The present Statistical Analysis Plan (SAP) refers to the 24-week double-blind treatment period (short-term) of this clinical study.

1.1 Study objectives

Primary Objective

The primary objective of this study is to compare the change from baseline in HbA1c after 24 weeks of double-blinded treatment with dapagliflozin 5 mg or 10 mg plus adjustable insulin versus placebo plus adjustable insulin.

Secondary Objectives:

Efficacy

Six secondary efficacy objectives are identified for special consideration in this study, in addition to the primary objective:

- Compare the percent change from baseline in total daily insulin dose with dapagliflozin 5 mg or 10 mg plus adjustable insulin versus placebo plus adjustable insulin after 24 weeks of double-blinded treatment
- Compare the percent change from baseline in body weight with dapagliflozin 5 mg or 10 mg plus adjustable insulin versus placebo plus adjustable insulin after 24 weeks of double-blinded treatment
- Compare the change from baseline in the mean value of 24-hour glucose readings obtained from continuous glucose monitoring (CGM) with dapagliflozin 5 mg or 10 mg plus adjustable insulin versus placebo plus adjustable insulin after 24 weeks of double-blinded treatment
- Compare the change from baseline in mean amplitude of glucose excursion (MAGE) of 24-hour glucose readings obtained from CGM with dapagliflozin 5 mg or 10 mg plus adjustable insulin versus placebo plus adjustable insulin after 24 weeks of double-blinded treatment
- Compare the change from baseline in the percent of 24-hour glucose readings obtained from CGM that falls within the target range of > 70 mg/dL and ≤ 180 mg/dL with dapagliflozin 5 mg or 10 mg plus adjustable insulin versus placebo plus adjustable insulin after 24 weeks of double-blinded treatment
- Compare dapagliflozin 5 mg or 10 mg plus adjustable insulin versus placebo plus adjustable insulin for the proportion of subjects achieving an HbA1c reduction from baseline to Week 24 visit ≥ 0.5% without severe hypoglycemia events

Safety

- To assess the proportion of subjects with hypoglycemia events and the frequency and severity of the hypoglycemia events with dapagliflozin 5 mg or 10 mg plus adjustable insulin versus placebo plus adjustable insulin
- To evaluate the safety and tolerability by assessment of adverse events (AE), vital signs, diabetic ketoacidosis (DKA) events, physical examination findings, ECGs, and laboratory values.

Other/Exploratory Objectives:

1) Assess the proportion of subjects with HbA1c reduction of at least 0.5% (≥ 0.5%) from baseline with dapagliflozin 5 mg or 10 mg plus adjustable insulin versus placebo plus adjustable insulin after 24 weeks of double-blinded treatment
2) Assess the proportion of subjects with HbA1c < 7.0% with dapagliflozin 5 mg or 10 mg plus adjustable insulin versus placebo plus adjustable insulin after 24 weeks of double-blinded treatment
3) Assess the change from baseline in fasting plasma glucose (FPG) with dapagliflozin 5 mg or 10 mg plus adjustable insulin versus placebo plus adjustable insulin after 24 weeks of double-blinded treatment
4) Assess the change from baseline in the percent of 24-hour glucose readings obtained from CGM that is within the hypoglycemic range of ≤ 70 mg/dL with dapagliflozin 5 mg or 10 mg plus adjustable insulin versus placebo plus adjustable insulin after 24 weeks of double-blinded treatment
5) Assess the change from baseline in seated systolic blood pressure with dapagliflozin 5 mg or 10 mg plus adjustable insulin versus placebo plus adjustable insulin after 24 weeks of double-blinded treatment among subjects with hypertension at baseline, defined as seated SBP ≥ 140 mmHg and/or seated DBP ≥ 90 mmHg.
6) Assess the change from baseline in average glucose values measured by 6-point self monitor blood glucose (SMBG) with dapagliflozin 5 mg or 10 mg plus adjustable insulin versus placebo plus adjustable insulin after 24 weeks of double-blinded treatment
7) Assess the change from baseline in postprandial glucose values measured by 6-point SMBG with dapagliflozin 5 mg or 10 mg plus adjustable insulin versus placebo plus adjustable insulin after 24 weeks of double-blinded treatment
8) Assess the change from baseline in postprandial glucose values measured by CGM with dapagliflozin 5 mg or 10 mg plus adjustable insulin versus placebo plus adjustable insulin after 24 weeks of double-blinded treatment
9) Assess the change from baseline in standard deviation of 24-hour glucose readings obtained from CGM with dapagliflozin 5 mg or 10 mg plus adjustable insulin versus placebo plus adjustable insulin after 24 weeks of double-blinded treatment
10) Assess the change from baseline to Week 24 in health status, as measured by the EQ-5D-3L questionnaire, between dapagliflozin 5 mg or 10 mg plus adjustable insulin versus placebo plus adjustable insulin after 24 weeks of double-blinded treatment.

11) Assess the changes from baseline to Week 24 in the summary score for treatment satisfaction and scores for perceived frequency of hyperglycemia and perceived frequency of hypoglycemia, respectively, as measured by the Diabetes Treatment Satisfaction Questionnaire status (DTSQs) between the dapagliflozin 5 mg or 10 mg plus adjustable insulin versus placebo plus adjustable insulin after 24 weeks of double-blinded treatment
12) Compare the number of non-severe hypoglycemic events per subject in those achieving an HbA1c reduction of $\geq$0.5% on dapagliflozin 5 mg or 10 mg plus adjustable insulin versus placebo plus adjustable insulin from baseline to Week 24 visit.
13) Assess the change from baseline to Week 24 in the summary score of fear of hypoglycemia as measured by the Hypoglycemia Fear Survey (HFS-II) Worry Subscale between the dapagliflozin 5 mg or 10 mg plus adjustable insulin versus placebo plus adjustable insulin after 24 weeks of double-blinded treatment.

1.2 Study design

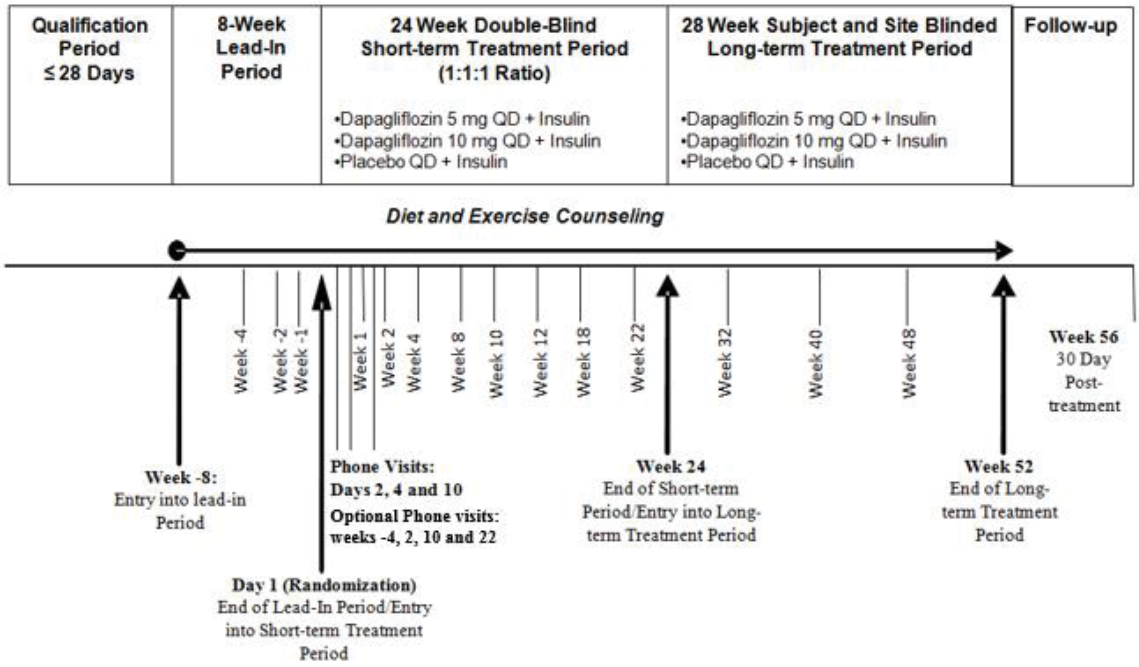
This trial is a randomized, double-blinded, three-arm, parallel-group, placebo-controlled, multicenter trial to evaluate the efficacy and safety of dapagliflozin, when added to ongoing insulin therapy, in subjects with T1DM and inadequate glycemic control.

Potential subjects will be assessed for eligibility criteria at the screening visit. Eligible subjects will enter an 8-week lead-in period in order to optimize their diabetes management based on individual subject challenge to glycemic control (including hyperglycemia, hypoglycemia, erratic meal/exercise patterns, as defined by the investigator) and to assess the variability in blood glucose profiles and frequency of hypoglycemic episodes at baseline. No placebo or study medication will be provided during the lead-in period. On the Day 1 visit, subjects who meet all the protocol-specific enrollment and randomization criteria will be randomized into one of the three blinded treatment arms (dapagliflozin 5 mg QD, dapagliflozin 10 mg QD, or placebo QD) in a 1:1:1 ratio. Randomization will be stratified by the following factors to ensure equal representation across all treatment groups: 1) current use of CGM (this refers to an unblinded/personal device which may already be in use before using the blinded device as part of the study) categorized as yes vs. no; 2) method of insulin administration at baseline (multiple daily injections (MDI) three or more injections per day vs. continuous subcutaneous insulin infusion - CSII); and 3) baseline A1C $\geq$ 7.5% and < 9.0% vs. $\geq$ 9.0% and $\leq$ 10.5%.

Subjects will receive dapagliflozin 5 mg QD, 10 mg QD, or placebo QD for 24 weeks, followed by a 28-week subject and site blinded long-term treatment period. Besides study medications, subjects will be treated with MDI injections (three or more injections per day of basal and prandial insulin) or CSII. It is recommended that subjects reduce their daily insulin dose by up to 20% for both basal and bolus insulins after first dose of study drug on Day 1 following randomization to reduce risk of hypoglycemia although in some cases it may be necessary to reduce insulin (particularly basal insulin) in advance of study drug administration. It is at the discretion of the investigator to determine the extent to which to reduce the insulin doses. Throughout the study (from the beginning of the lead-in period to the end of the long-term treatment period), insulin dose may be adjusted as deemed appropriate to

be consistent with good medical practice according to SMBG readings, local guidance and individual circumstances. Subjects will be monitored closely with regard to safety parameters, including vital signs, safety laboratory tests, ECGs, and adverse events.

Figure 1 Study Schematic



Notes:

Eligible subjects should have the Day 1 randomization visit within 56 ± 5 days of the Week -8 entry into lead-in period visit.

Subjects will receive the last dose of study medications on Week 52 visit and they will have the last follow-up visit on Week 56 visit (30 days after Week 52 visit).

1.3 Number of subjects

The primary endpoint is the change from baseline in HbA1c to Week 24 visit.

With 243 subjects per treatment group with both baseline and at least one post baseline measurement, there is approximately 90% power to detect a difference in means of 0.35% between each dapagliflozin treatment group and placebo at the two-sided 0.0262 significance level (based on Dunnett and Tamhane step-up procedure) (Dunnett and Tamhane 1992), assuming a standard deviation (SD) of 1.1%. Assuming that 5% of subjects do not have a post-baseline assessment, a total of 768 subjects (256 subjects per treatment arm) need to be randomized to each dapagliflozin or placebo group in 1:1:1 ratio. Due to regulatory requirements, approximately 160 subjects from Japan will be included (of the total population of 768 subjects).

A mean reduction in the primary endpoint of change from baseline in HbA1c of 0.35% was used to determine sample size as this represents a clinically meaningful improvement in glycemic control under the current study design with adjustable background insulin dose. A lesser degree of reduction in HbA1c in combination with other beneficial outcomes in secondary endpoints may also be clinically important.

2. ANALYSIS SETS

2.1 Definition of analysis sets

2.1.1 Enrolled subjects dataset

The Enrolled Subjects Dataset includes all subjects who signed informed consent.

2.1.2 Lead-in subjects dataset

The Lead-in Subjects Dataset includes all subjects who have at least one vital sign measurement during the lead-in period.

2.1.3 Full analysis dataset

The full analysis dataset, aka, full analysis set, will consist of all randomized subjects who took at least one dose of double-blind study medication during the short-term double-blind period. Whenever using the full analysis dataset, 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).

2.1.4 Evaluable subjects dataset

The evaluable subjects dataset, aka, per-protocol analysis set, will be a subset of the full analysis dataset through exclusion of subjects with relevant protocol deviations. Relevant protocol deviations are those that have the potential to impact the results of the primary analysis. Detailed exclusion criteria for the evaluable subjects dataset are specified in section 2.2 All decisions to exclude subjects from the full analysis dataset to form the evaluable subjects dataset will be made prior to the unblinding of the study.

The per-protocol analysis set will be used to analyze the primary endpoint only if more than 10% of the full analysis set subjects in any treatment group have relevant protocol deviations, as defined in Section 2.2.

2.1.5 Treated subjects dataset

The treated subjects dataset, aka, safety analysis set, consists of all subjects who received at least one dose of double-blind study medication during the short-term double-blind treatment period. The treated subjects dataset would include any subject who accidentally received double-blind study medication but was not randomized in the study. These subjects will be included in treated subject dataset analyses based on the treatment they received. For randomized subjects, all analyses using the treated subject dataset will be presented by 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 study period). In this case, the safety data for those subjects will be presented by the treatment actually received. In the case where a subject never received the treatment as assigned by randomization, then the safety data for that subject will be presented by the first treatment received.

2.2 Violations and deviations

Protocol deviations will be reviewed in a blinded fashion by the study team prior to database lock. All decisions to exclude subjects and/or data from the Full Analysis Set to form the Per-Protocol Analysis Set will be made prior to the unblinding of the study and agreed by the study team.

Subjects who deviate from protocol conditions (e.g., important inclusion/exclusion criteria) will be reported as having major (or significant) protocol deviations. Major protocol deviations that are determined to have the potential to affect the primary efficacy results are deemed Relevant Protocol Deviations (RPDs). A list of relevant protocol deviation criteria, along with the consequent handling of data should those deviations occur, is given in Table 1. Minor (excluding relevant) protocol deviations are listed in Appendix 9.1.

Table 1 List of Relevant Protocol Deviations

Number	RPD criteria	Exclusion Level
1	Randomized Subjects without Type 1 diabetes or central laboratory HbA1c obtained at randomization not within ± 0.2% of the protocol-specific HbA1c range	Complete exclusion
2	Randomized Subjects not satisfying the target population antihyperglycemic therapy requirement (insulin use > 12 months, unchanged method (≥ 3 injections/day or CSII) and total insulin dose of ≥ 0.3U/kg/day for at least 3 months prior to the screening visit).	Complete exclusion
3	Randomized Subjects that used antihyperglycemic medication (other than protocol allowed medication and study medication), for 14 or more consecutive days during the short-term double-blind treatment-period.	Partial exclusion (All efficacy data after the 14th day of administration non-protocol required antihyperglycemic medication will be excluded)
4	Randomized Subjects that were treated with any systemic corticosteroid therapy for ≥ 5 consecutive days during the short-term double-blind treatment period.	Partial exclusion (All efficacy data after the 5th day of systemic corticosteroid therapy will be excluded)

Table 1 List of Relevant Protocol Deviations

Number	RPD criteria	Exclusion Level
5	Randomized subjects who do not take any double-blind study medication assigned at randomization randomized for ≥ 2 consecutive weeks	Partial exclusion. (All efficacy data after the 14th day of interruption will be excluded.)
6	Randomized subjects with abnormal free T4 values at enrollment	Complete exclusion
7	Randomized subjects with history of hemoglobinopathy, with the exception of sickle cell trait (SA) or thalassemia minor; or chronic or recurrent hemolysis.	Complete exclusion
8	Randomized subjects less than 80% or more than 120% compliant with double-blind treatment regimen.	Complete exclusion

Subjects having relevant protocol deviations will be summarized by treatment group and overall. Subjects with any protocol deviations will be listed.

3. PRIMARY AND SECONDARY VARIABLES

3.1 Primary efficacy variable

The primary efficacy variable is the change in HbA1c from baseline to week 24.

The primary efficacy variable will be calculated as outlined in Section 4.1.2.2.

3.2 Secondary efficacy variables

The secondary efficacy variables are:

1) Percent change from baseline to the Week 24 visit in total daily insulin dose
2) Percent change from baseline to the Week 24 in body weight
3) Change from baseline to the Week 24 visit in the mean value of 24-hour glucose readings obtained from CGM
4) Change from baseline to the Week 24 visit in mean amplitude of glucose excursion (MAGE) of 24-hour glucose readings obtained from CGM
5) Change from baseline to the Week 24 visit in the percent of 24-hour glucose readings obtained from CGM that falls within the range of > 70 mg/dL and ≤ 180 mg/dL
6) Proportion of subjects achieving an HbA1c reduction from baseline to the Week 24 visit (LOCF) ≥ 0.5% and without severe hypoglycemia events

Percent change from baseline and change from baseline for a secondary efficacy variable will be calculated as outlined in Sections 4.1.2.2 and 4.1.2.3. Proportions will be analyzed as outlined in Section 4.1.8.

3.3 Safety variables

The safety evaluations will include analyses of:

- AEs
- Laboratory parameters (including Calculated creatinine clearance (CrCl))
- ECG
- Vital signs (pulse, BP)
- Hypoglycemic events
- Diabetic ketoacidosis events
- Physical examination findings

For calculating the creatinine clearance the method of Cockcroft and Gault will be used:

$$CrCl = \frac{(140 - age) \times (weight\ in\ kg)}{serum\ creatinine\ (mg/dL) \times 72} \quad (\times 0.85 \text{ if female})$$

3.4 Exploratory variables

The efficacy exploratory endpoints of the study are:

1) Proportion of subjects with HbA1c reduction from baseline to the Week 24 visit (LOCF) of at least 0.5% (≥ 0.5%)
2) Proportion of subjects with HBA1C < 7.0% at the Week 24 visit (LOCF) of double-blinded treatment
3) Change from baseline to the Week 24 visit in FPG
4) Change from baseline to the Week 24 visit in the percent of 24-hour glucose readings obtained from CGM that falls within the hypoglycemic range of ≤ 70 mg/dL.
5) Change from baseline to the Week 24 visit in seated systolic blood pressure among subjects with hypertension at baseline, defined as seated SBP ≥ 140 mmHg and/or seated DBP ≥ 90 mmHg.
6) Change from baseline to the Week 24 visit in average glucose values measured by 6-point SMBG.
7) Change from baseline to the Week 24 visit in postprandial glucose values measured by 6-point SMBG.

8) Change from baseline to the Week 24 visit in postprandial glucose values measured by CGM
9) Change from baseline to the Week 24 visit in the standard deviation of 24-hour glucose readings measured by CGM
10) Change from baseline to the Week 24 visit in each of the 5 dimensions, as well as the health status index of the EQ-5D-3L questionnaire.
11) Changes from baseline to the Week 24 visit in the summary score for treatment satisfaction and scores for perceived frequency of hyperglycemia and perceived frequency of hypoglycemia, respectively, as measured by the Diabetes Treatment Satisfaction Questionnaire status (DTSQs)
12) Number of non-severe hypoglycemic events per subject in those achieving a HbA1c reduction of ≥ 0.5% from baseline to the Week 24 visit (LOCF)
13) Changes from baseline to Week 24 in the summary score of the HFS-II Worry Subscale

4. ANALYSIS METHODS

4.1 General principles

4.1.1 Objectives and hypotheses

Research Hypothesis: After 24 weeks of oral administration of double-blind treatment, the change from baseline in A1C (HbA1c) level with dapagliflozin plus adjustable insulin is greater than placebo plus adjustable insulin in subjects with type 1 diabetes who have inadequate glycemic control.

The primary objective of this study is to compare the change from baseline in HbA1c after 24 weeks of double-blinded treatment with dapagliflozin 5 mg or 10 mg plus adjustable insulin versus placebo plus adjustable insulin. The primary estimand for the primary endpoint is treatment difference at Week 24 if subjects did not discontinue randomized treatment. In order to maintain an overall Type I error rate of 5% for the endpoint, a Dunnett and Tamhane step-up procedure will be used, which allows for the correlation of 0.5 between the standard normal deviate for each comparison (Dunnett and Tamhane 1992). Applying the Dunnett and Tamhane procedure, statistical significance will be declared for both doses at the two-sided 5% level if the two-sided p-values from both pairwise comparisons are smaller than 5%. If the larger p-value among the two pairwise comparisons is greater than 5% and the smaller p-value is below 2.62%, then statistical significance will be declared for the latter comparison.

Secondary objectives are to compare the effects of each dose of dapagliflozin versus placebo given as add-on therapy to insulin after a 24-week double-blind treatment period by evaluation of the endpoints outlined in section 1.1

Statistical tests for secondary efficacy endpoints will be only conducted if there is a statistically significant difference in the primary endpoint for both pairwise comparisons, i.e., dapagliflozin 5mg vs. placebo and dapagliflozin 10mg vs. placebo. The same Dunnett and Tamhane step-up procedure (Dunnett and Tamhane 1992) will be applied to each secondary

efficacy endpoint. Secondary efficacy endpoints will be tested in the order that they appear in the objectives section of the protocol. Statistical tests will be only performed for a given secondary endpoint if both comparisons for a proceeding secondary endpoint are significant. Otherwise, the testing procedure will stop at the secondary endpoint that does not reach statistical significance.

4.1.2 Definitions

4.1.2.1 Short-term treatment period, long-term extension period, baseline value

The study consists of the following study periods:

1) The qualification period. This period starts with enrollment and ends on the start of the lead-in period.

2) The 8-week lead-in period. During this period, subjects will optimize their diabetes management based on individual subject challenge to glycemic control. No placebo or study medication will be provided during the lead-in period.

3) The 24-week short-term double-blind treatment period which will be the timeframe for the primary efficacy analysis. On the Day 1 visit, subjects who meet all the protocol-specific enrollment and randomization criteria will be randomized. It will be recommended that subjects reduce their daily insulin dose by 20% for both basal and bolus insulin on Day 1 following randomization to reduce the risk of hypoglycemia.

4) The 28-week long-term treatment period where additional safety and efficacy measurements will be obtained. After completing the short-term period, subjects will enter the 28-week subject- and site-blinded long-term treatment period.

5) The follow-up period. After completing the short-term and long-term period, subjects will enter the 4 week follow-up period.

For each subject, the baseline value of a parameter is defined as the last assessment of that parameter on or prior to the date/time of the first dose of the double-blind study medication.

For measurements referring to 24-hour collection periods (i.e. daily insulin dose), all available data prior to the date of the first dose of the double-blind study medication will be considered as falling in the baseline analysis window.

For measurements referring to multiple data points over the course of a day (i.e. 6-point SMBG measurement parameters and CGM parameters), all available data on or prior to the date/time of the first dose of the double-blind study medication will be considered as falling in the baseline analysis window.

For measurements referring to 24-hour collection periods (i.e. daily insulin dose) and those referring to multiple data collection points over a day (i.e. 6-point SMBG measurement parameters and CGM parameters), for which several values are considered as falling in the

baseline analysis window, the baseline value will be identified using the algorithm outlined in Section 6.3 in order to define the baseline value.

4.1.2.2 Change from baseline

Change from baseline to any Week t in double-blind treatment period is defined as follows:

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

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.

The "Week t" to which a measurement belongs is determined using the conventions described in Section 6.3.

4.1.2.3 Percent change from baseline

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

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

Where

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

The "Week t" to which a measurement belongs is determined using the conventions described in Section 6.3.

For analyses of parameters in terms of percent change from baseline to Week t, values will first be transformed to natural logarithms (Ln) and the results will be expressed as geometric mean percent changes from baseline. Analysis will be performed using the logarithms of the post-baseline to baseline ratios. Subsequently, the estimates from the analysis will be back-transformed to original values for reporting in the tables using the formulae detailed in Table 2

Table 2 Formulae Used to Transform Back onto the Original Scale the Estimates of Percent Change from Baseline Analyzed as Geometric Mean

Quantity	Computation method
Geometric mean of the Week t to baseline ratio	Exp(mean change from baseline in natural logarithm)
Mean percent change from baseline	100 x [exp(mean change from baseline in natural logarithm*) – 1]

Table 2 Formulae Used to Transform Back onto the Original Scale the Estimates of Percent Change from Baseline Analyzed as Geometric Mean

Quantity	Computation method
Standard error of mean percent change from baseline	100 x exp(mean change from baseline in natural logarithm) x standard error of mean change from baseline in natural logarithm* – or, equivalently – 100 x Geometric mean of the Week t to baseline ratio x standard error of mean change from baseline in natural logarithm*
Lower confidence limit for mean percent change from baseline	100 x [exp(lower confidence limit for mean change from baseline in natural logarithm*) – 1]
Upper confidence limit for mean percent change from baseline	100 x [exp(upper confidence limit for mean change from baseline in natural logarithm*) – 1]
Adjusted geometric mean of the Week t to baseline ratio	exp(Adjusted mean change from baseline in natural logarithm*)
Adjusted mean percent change from baseline	100 x [exp(Adjusted mean change from baseline in natural logarithm*) – 1]
Standard error of adjusted mean percent change from baseline	100 x exp(Adjusted mean change from baseline in natural logarithm) x standard error of mean change from baseline in natural logarithm* – or, equivalently – 100 x adjusted geometric mean of the Week t to baseline ratio x standard error of mean change from baseline in natural logarithm*
Lower confidence limit for adjusted mean percent change from baseline	100 x [exp(lower confidence limit for adjusted mean change from baseline in natural logarithm*) – 1]
Upper confidence limit for adjusted mean percent change from baseline	100 x [exp(upper confidence limit for adjusted mean change from baseline in natural logarithm*) – 1]
Adjusted geometric mean of the Week t to baseline ratio achieved with each dapagliflozin treatment arm relative to that achieved with Control, expressed as a percent difference. Please note that for the SAS output, a shorter text will be used: Adjusted GM of Week t/Baseline for each dapagliflozin treatment arm relative to Control, in % difference.	100 x (((adjusted mean percent change for dapagliflozin treatment arm +100)/(adjusted mean percent change for Control +100)) – 1) – or, equivalently – 100 x (exp(difference in adjusted mean change from baseline between dapagliflozin treatment arm and Control in natural logarithm*) – 1)

*Change in natural logarithm refers to a calculated difference between two values after performing a natural logarithmic transformation on each.

4.1.2.4 Insulin up-titration more than 25% than baseline

Information on insulin dosing is collected in two ways over the course of the study. For some two-week periods, total daily bolus and total basal doses are reported on a daily basis, separately for each type of insulin. For the rest of study duration, from the start of lead-in to end of study, the information of total (the sum of basal and bolus) daily insulin dose is reported on a weekly basis and it corresponds to the minimum and maximum value of total insulin administration on a day during the weekly interval.

For the purpose of analysis, for those days that the total daily bolus and total basal doses are reported separately, the total daily insulin dose will be calculated by summing up the bolus and basal dose for the same day. If either the bolus or basal dose for the same day are missing, then the total daily will not be calculated for that day. Subsequently, these data will be grouped into 7-days periods and the minimum and maximum value will be calculated for each of these periods in analogy to the reporting for the weekly intervals.

The calculation of the insulin up-titration will be based on the midpoints between maximum and minimum value for each reported weekly interval period. All days within a week period will be considered having the same midpoint value.

The starting time point of the up-titration is considered to be the first day of the first 14-day window during the study if following conditions are met: 1). All of non-missing daily midpoint values in a 14 consecutive days window are 25% higher than the baseline value; 2) number of missing values is less than 7 in a 14-days window; and 3) there are less than 5 consecutive days having missing values in a 14-days window. The 14- day windowing will be shifted by one day each time, starting from the first dose date.

The baseline value for this calculation will be defined as the midpoint between maximum and minimum value in 7 days with non-missing records prior to the date of the first dose of the double-blind study medication (see section 4.1.2.1). The midpoint derived from weekly records in 7 days prior to the first dose date will be considered as a baseline value if the baseline value derived from daily records is missing. This value will be used to compare on treatment values to define insulin up-titration more than 25%.

4.1.2.5 Last Observation Carried Forward (LOCF)

In the analysis of change (or percent change) from baseline, as well as response endpoint at Week *t* LOCF, the Week *t* measurement will be used. If no Week *t* measurement is available (subject has discontinued before Week *t,* or the subject has not discontinued, but the measurement was not taken at Week *t*), the last post-baseline measurement prior to Week t will be used. Unless specified otherwise, all data, regardless of insulin up-titration, will be used for the calculation of LOCF values.

For the sensitivity analysis, in the case that data from subjects after the first post-baseline up-titration of insulin by more than 25% from baseline would be excluded from analysis, the last

post-baseline measurement taken prior to or on the date of the first instance the up-titration is met will be used in analysis.

4.1.2.6 Average daily value of continuous glucose monitoring system readings

The daily average for each subject will be calculated as the area under the curve (AUC), using the trapezoidal rule, divided by total duration of data collection, using the formula below:

$$\frac{\sum_{i=1}^{n}\left(\frac{\left(M_i + M_{i+1}\right)}{2}\cdot\left(T_{i+1} - T_i\right)\right)}{\left(T_n - T_1\right)}$$

Where M_i is the CGM reading at time T_i and M_{i+1} is the subsequent non-missing reading measured at time T_{i+1}. T_1 and T_n are the time of the first and last non-missing CGM reading during the 24-hour period.

4.1.2.7 Percent of continuous glucose monitoring system readings falling within a range

Percent of CGM readings falling within a range over a 24-hour period will be calculated as the number of readings during the 24-hour period within the range divided by the total number of available readings during the same 24-hour period.

4.1.2.8 Standard deviation of continuous glucose monitoring system readings

The 24-hour standard deviation of CGM readings will be calculated as the standard deviation of available readings within the 24-hour period.

4.1.2.9 Mean daily glucose of 6-point self monitored blood glucose

Subjects must perform 6-point blood glucose monitoring (before meal, 2 hours post meal) at home for any 3 days within one week prior to Day 1 visit, Week 12 and Week 24/Study Termination or Early Treatment Termination. Meals are considered to be breakfast, lunch and dinner. Preprandial and postprandial measurements will be considered eligible for analysis if they occur within 15 minutes prior to the start time of each meal and from 1.5 to 2.5 hours after the start of the meal time, respectively. The pre- and post-prandial meal glucose measurements obtained outside the time windows (within 15 minutes prior to the meal, 1.5 to 2.5 hours after the meal) will be excluded in the summaries and analyses.

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

be calculated as (pre-breakfast + post-breakfast)/ 2. The average value for pre-dinner is excluded for the calculation since the average value for the post-dinner is missing.

4.1.2.10 Post-prandial glucose values of continuous glucose monitoring readings and of self-monitored blood glucose (SMBG)

The "start of the meal time" will be collected for breakfast, lunch and dinner for the last seven days of CGM collection. During the same periods subjects must perform 6-point finger stick blood glucose monitoring. The post-prandial glucose values of CGM and SMBG readings will be those values that occur 2 hours after the start time of breakfast, lunch and dinner. Postprandial measurements will be considered eligible for analysis if they occur from 1.5 to 2.5 hours after the start of the meal time. The post-prandial measurements obtained outside the time windows (1.5 to 2.5 hours after the meal) will be excluded in the summaries and analyses. For CGM, if there is more than one eligible value, then the value used for analysis will be selected as described in Section 6.3, where the target value will be the value at exactly 2 hours after start time of each meal.

For either CGM or SMBG parameters, glucose concentrations at each of the 3 post-prandial time points will be averaged, over the days of data collection, to derive the mean glucose concentrations at each of the 3 time points. The mean post-prandial daily glucose (MPDG) will then be calculated as the average over the average glucose concentrations at the 3 time points. The average of all post-prandial glucose values during a day will be used for the analysis of change in post-prandial glucose values independently of whether there is a corresponding preprandial value available.

4.1.2.11 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*" (Service et al 1970). Therefore, it corresponds to the mean of absolute differences between consecutive maxima and minima, as long as these differences are greater than one standard deviation for the same 24-hour period.

Here an example of calculation as described in the original publication of MAGE in (Service et al 1970).

Figure 2 **Blood glucose curve of a subject studied on intermediate-acting insulin regimen.**

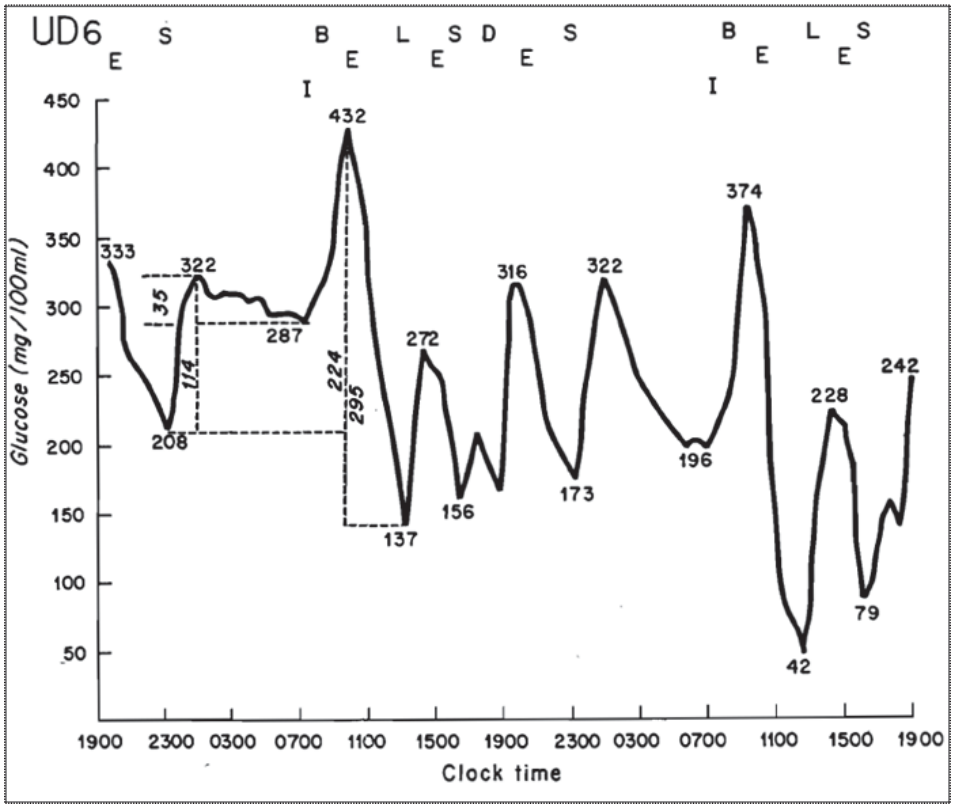


Figure from (Service et al 1970)

As stated in (Service et al 1970), the calculation of MAGE for subject in Figure 2 "*is as follows: the first excursion, 333 to 208mg/100mL, is 125 mg/100mL, exceeding the value of one standard deviation; 62. The first blood glucose increase, from 208 to 322mg/100mL, is 114mg/100ml [...] and also exceeds one standard deviation; however, the subsequent blood glucose decrease of 35 mg/100mL (from 322 to 287mg/100mL) is less than one standard deviation and so this excursion is not counted. The next blood glucose increase, from 208 to 432 (224mg/100mL), and decrease, from 432 to 137mg/100mL (295 mg/100mL), 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 272mg/100mL which corresponds to an excursion of 135 mg/100mL. The next excursion is 116mg/100mL to a nadir at 156mg/100mL, followed by an excursion of 160mg/100mL to a peak at 316mg/100mL. The next nadir is at 173mg/100mL corresponding to an excursion of 143mg/100mL and it is followed by a peak at 322mg/100mL and a respective excursion of 149mg/100mL and by a nadir at 196mg/100mL and a respective excursion of 126mg/100mL. The next excursion is 178mg/100mL to a peak at 374mg/100mL,

followed by one of 332mg/100mL to a nadir at 42mg/100mL and one of 186mg/100mL to the subsequent peak of 228mg/100mL. Finally, the last couple of excursions are 149mg/100mL to the next nadir at 79mg/100mL and 163mg/100mL to the final peak point at 242mg/100mL. Therefore, MAGE for this subject profile corresponds to the arithmetic mean of all excursions and equals approximately 177mg/100mL.

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.

4.1.3 Descriptive summaries of continuous variables

Descriptive summaries of continuous variables in terms of change or percent change from baseline values will be provided by treatment group and overall, if applicable.

4.1.4 Descriptive summaries of categorical variables

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

4.1.5 Analysis of Covariance (ANCOVA)

4.1.5.1 ANCOVA model for change from baseline

In summaries of efficacy endpoints examining changes from baseline at Week t or Week t LOCF, ANCOVA of the differences between post-baseline and baseline measurements will be performed, with treatment group and randomization stratification factors (i.e. one term for each combination of all stratification factors) as effects and the baseline measurement as a covariate. The following ANCOVA model will be used:

$$D_{t,ijk} = intercept + \beta\,[Y_{0,ijk}\,] + \tau_i + s_k + error_{ijk} \text{ (Model 4.1.5.1)}$$

where

- $D_{t,ij} = Y_{t,ijk} - Y_{0,ijk}$ = the Week t or Week t LOCF change from baseline of subject j in treatment group i (as defined in Section 4.1.2.2 and Section 6.3 on conventions),
- $Y_{0,ijk}$ is the baseline measurement of subject j in treatment group i,
- $Y_{t,ijk}$ is the Week t or Week t LOCF 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.

- s_k is the mean effect of randomization stratification factor group k (i.e. category k of the variable with categories for all possible combinations of stratification factor groups).
- *Intercept*, β and τ_i, s_k 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.

4.1.5.2 ANCOVA model for percent change from baseline

For analyses of parameters in terms of percent change from baseline to Week t or Week t LOCF, values will first be transformed to logarithms and the results will be expressed as geometric mean percent changes from baseline. ANCOVA of the logarithms of the post-baseline to baseline ratios will be performed. Treatment group and randomization stratification factors (i.e. one term for each combination of all stratification factors) will be considered as a fixed effect. The natural logarithm (Ln) of the baseline measurement will also be included as a covariate. The following ANCOVA model will be used:

$$Z_{t,ijk} = intercept + \beta\,[Ln(Y_{0,ijk})] + \tau_i + s_k + error_{ijk} \quad (Model4.1.5.2)$$

where

- $Z_{t,ij} = Ln(Y_{t,ijk}) - Ln(Y_{0,ijk})$ = the Week t or Week t LOCF change from baseline in the natural logarithm of subject j in treatment group i,
- $Y_{0,ijk}$ is the baseline measurement of subject j in treatment group i,
- $Y_{t,ijk}$ is the Week t or Week t LOCF measurement of subject j in treatment group i,
- β is the slope of $Z_{t,ij}$ regressing on the baseline measurement and,
- τ_i is the mean effect of treatment group i.
- s_k is the mean effect of randomization stratification factor group k (i.e. category k of the variable with categories for all possible combinations of stratification factor groups).
- *Intercept*, β and τ_i, s_k are unknown parameters to be estimated from the data.

The model will provide least squares mean estimates and 2-sided 95% confidence intervals for (geometric) mean percent changes within and (when warranted) between treatments. Where applicable, the t-statistic 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.

Table 2 details the formulae that will be used to transform back the results from model 4.1.5.2 to obtain the values reported in the tables.

4.1.6 Longitudinal repeated measures analysis

4.1.6.1 Longitudinal repeated measures analysis for outcomes with single measurement per week t

A longitudinal repeated measures analysis using 'direct likelihood' will be performed. The SAS procedure PROC MIXED will be used.

The dependent variable will be the change from baseline to each Week t included in the model for efficacy endpoints examining changes from baseline. For analyses of parameters in terms of percent change from baseline at Week t, values will first be transformed to logarithms and the results will be expressed as geometric mean percent changes from baseline. The dependent variable will be the natural logarithms (Ln) of the post-baseline to baseline ratios. The model estimates will be back transformed to original values using the formulae detailed in Table 2.

The preferred model will include the fixed categorical effects of treatment, week, randomization stratification factors (i.e. one term for each combination of all stratification factors) and treatment-by-week interaction, as well as the continuous fixed covariates of baseline measurement and baseline measurement-by-week interaction. For parameters analyzed as percent change from baseline, the natural logarithm of the baseline values will be used in the above model specification. 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:

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

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

The model will provide least-squares mean estimates, standard errors and 2-sided 95% confidence intervals for mean change at all time points within and between treatments. Where applicable, for specific weeks, the t-statistic 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.

In case that is specified, an additional effect will be added in the mixed model to indicate those values after baseline up-titration more than 25% from baseline.

4.1.6.2 Longitudinal repeated measures analysis for outcomes with multiple measurement per week t

Some efficacy parameters (i.e. parameters from CGM or parameters from 6-point SMBG or daily insulin dose) are measured daily during a visit window.

Within each visit window (including baseline and post-baseline visit windows as outlined in Section 6.3), the average for all available data points will be calculated and this will constitute the baseline or Week t value (as applicable).

Analysis will be performed with a longitudinal repeated measurements model as outlined in Section 4.1.6.1.

4.1.6.3 Longitudinal repeated measures analysis model assumption assessment

The assumptions underlying the longitudinal repeated measures analysis model will be checked, if specified in Section 4.2.3. This section also details the steps to follow in case these assumptions would not be satisfied.

Assessment of Treatment-by-Baseline Interaction:

Treatment-by-baseline interaction will be assessed for the Week 24 estimates. In the model defined in 4.1.6.1 and 4.1.6.2, the interaction will be tested by including the treatment-by-baseline interaction.

The following contrast statements will be used of testing for interaction.

- Contrast coefficients for all weeks but Week 24 will be zero.
- Contrast coefficients at Week 24 are presented in Table 3.

Table 3 Contrast Coefficients for the treatment-by-baseline interaction for week 24 estimates

Treatment by baseline interaction	Contrast Coefficients
Dapa 5mg by baseline value	0.5
Dapa 10mg by baseline value	0.5
Placebo by baseline value	-1

The test for interaction will be performed at the 0.10 level of significance. If significant, the interaction will be assessed as qualitative or quantitative. Assessment of the interaction type will be based on regression lines plotted for each treatment group. The intercepts and slopes for these regression plots will be obtained from the analysis model including the interaction term. The intercepts will be estimated by the least squares means. The abscissa of the plots will range from the minimum baseline value from all subjects included in the analysis to the maximum baseline value.

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

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

Assessment of Treatment-by-combination of Stratification Factor Interaction:

Treatment-by-combination of stratification factors interaction for estimates at Week 24 will be assessed using method described in Section 4.2.3.1 on primary analysis within subgroups. In the model defined in 4.1.5.1and 4.1.5.2, the interactions will be tested by including the treatment-by-combination of stratification factor interaction and appropriate contrast statement for testing the interaction in the week 24 estimates. The tests for interaction will be performed at the 0.10 level of significance.

Outlier Detection:

Specifically, an outlier is defined as an observation with a residual that is more than three times the interquartile range above the 75^{th} percentile or below the 25^{th} percentile. Outliers will be identified based on diagnostics performed for the primary efficacy outcome measure using standardized residuals from the longitudinal repeated measurement model. These diagnostics will be based on the box plot. If outliers are present, then additional sensitivity analyses will be performed with the outliers excluded, in order to assess their impact on the results.

Adjustments for baseline imbalances:

Adjustments for baseline imbalances between treatment groups in the longitudinal repeated measures model may be performed as a sensitivity analysis: baseline characteristics for which a clinically important imbalance exists between the treatments will be identified. If such imbalances exist, an adjustment will be performed by including the corresponding baseline characteristics as additional covariates in the longitudinal repeated measures model. This will allow the assessment of the importance of these imbalances on the treatment group comparisons.

4.1.7 Kaplan-Meier curve and estimates for time-to-event analyses

Kaplan-Meier plots (Kaplan and Meier 1958) of time to event variables will be displayed by treatment group. Unless otherwise specified, the plot will be presented only when there are at least 5 events in one treatment group. Additionally, a table will accompany the plot and will display the Kaplan-Meier estimates of the cumulative proportion (with 95% CI calculated based on Greenwood's method when applicable) (Greenwood 1926) 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.

4.1.8 Proportion of subjects with a pre-defined characteristic at week t

The proportion of subjects with a pre-defined characteristic at Week t will be analyzed using a logistic regression model with adjustments for baseline values and randomization strata when there are at least 5 responders on average by treatment group. This pre-defined characteristic could correspond to a continuous variable dichotomized into a categorical endpoint. The odds ratio in response rate between each of the dapagliflozin treatment groups and the control group will be displayed along with the 2-sided 95% confidence intervals. P-values will be calculated if applicable. When there are less than 5 responders on average by treatment group, the unadjusted (and difference) proportions, exact 95% confidence interval, and p-values from the Fisher's exact test (when applicable) will be provided.

4.2 Analysis methods

4.2.1 Study population

4.2.1.1 Subject disposition

The disposition of subjects for the pre-randomization period and the short-term double-blind treatment period will be summarized.

The status in the pre-randomization period will be summarized based on all subjects enrolled (who signed informed consent). That includes all subjects who entered lead-in period, subjects who completed lead-in period, reasons for treatment discontinuation from lead-in period, subjects who were randomized, subjects who were not randomized, and reasons for not being randomized. The summary of status in the short-term double-blind treatment period will include all randomized subjects, and be presented by randomized treatment group. This summary will include subjects completing and discontinuing treatment during the short-term double-blind treatment period with reasons for treatment discontinuation from the short-term double-blind treatment period, and their status (completing the short-term treatment, continuing to long-term treatment, continuing to short-term off-treatment, long-term off-treatment and 30 day follow up) after double-blind treatment period . This summary will be repeated for Japanese subjects.

Subjects enrolled, randomized, and treated will be summarized by country and study site.

4.2.1.2 Demographic and other baseline characteristics

Demographic and other baseline characteristics, including diabetes-related characteristics and renal function characteristics will be summarized by treatment group and overall, using the Full Analysis Set and the Per Protocol Analysis Set if deemed necessary.

Demographic and baseline characteristics are listed in Table 4. Diabetes related baseline characteristics are listed in Table 5. Renal function baseline characteristics are listed in Table 6.

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 (i.e., each denominator includes the number of subjects with missing/unknown values for the characteristic). This summary will be repeated for Japanese subjects.

Table 4 Demographic and Baseline Characteristics

Characteristic	Summarized as	Categories
Gender	Categorical	Male, Female
Age	Categorical and Continuous	< 65yrs ≥ 65yrs and <75yrs ≥ 75yrs
Age	Categorical	<35yrs ≥35yrs and <50yrs ≥50yrs
Female Age	Categorical	≤ 50yrs > 50yrs
Race	Categorical	White, Black or African American, Asian, Other
Country	Categorical	Japan Other
Ethnicity	Categorical	Hispanic/Latino, Non Hispanic/Latino
Body Weight	Continuous	-
Body Mass Index	Categorical and Continuous	≤ 23kg/m^2 > 23kg/m^2 - ≤ 25kg/m^2 > 25kg/m^2 - ≤ 27kg/m^2 > 27kg/m^2 - ≤ 30kg/m^2 > 30kg/m^2
Geographic Region	Categorical	As defined in Appendix 9.2

Table 5 Diabetes-Related Baseline Characteristics

Duration of Type 1 Diabetes	Categorical and Continuous	< 10yrs ≥ 10 and ≤ 20yrs > 20yrs
HbA1c	Categorical and Continuous	< 8% ≥ 8 and < 9% ≥ 9%
FPG	Continuous	-
Fasting C-peptide	Continuous	-
Fasting C-peptide	Categorical	< 0.05ng/mL(lower limit of detection) ≥ 0.05ng/mL
Method of insulin administration	Categorical	Multiple Daily Injections Continuous Subcutaneous Insulin Infusion
Use of continuous glucose monitoring (CGM)	Categorical	Yes No
Baseline HbA1c	Categorical	≥ 7.5% and < 9.0% ≥ 9.0% and ≤ 10.5%

Combinations of randomization stratification factors	Categorical	Baseline HbA1c ≥ 7.5% and < 9.0% - Multiple Daily Injections - with CGM use Baseline HbA1c ≥ 7.5% and < 9.0% - Multiple Daily Injections - without CGM use Baseline HbA1c ≥ 7.5% and < 9.0% - Continuous Subcutaneous Insulin Infusion - with CGM use Baseline HbA1c ≥ 7.5% and < 9.0% - Continuous Subcutaneous Insulin Infusion - without CGM use Baseline HbA1c ≥ 9.0% and ≤ 10.5% - Multiple Daily Injections - with CGM use Baseline HbA1c ≥ 9.0% and ≤ 10.5% - Multiple Daily Injections - without CGM use Baseline HbA1c ≥ 9.0% and ≤ 10.5% - Continuous Subcutaneous Insulin Infusion - with CGM use Baseline HbA1c ≥ 9.0% and ≤ 10.5% - Continuous Subcutaneous Insulin Infusion - without CGM use

Table 6 Baseline Renal Function Characteristics

Characteristic	Summarized as	Categories
eGFR (MDRD)	Categorical	< 60mL/min/1.73m2 ≥ 60 and < 90mL/min/1.73m2 ≥ 90mL/min/1.73m2
Calculated Creatinine Clearance using the Cockcroft and Gault method	Categorical	< 60mL/min ≥ 60 and < 90mL ≥ 90mL/min

4.2.1.3 Specific and general disease histories

The number (percent) of subjects with type 1 diabetes, diabetes-related disease histories will be summarized by treatment group and overall using the Full Analysis Set.

The number (percent) of subjects with general medical history findings will also be summarized by treatment group and overall using the Full Analysis Set.

4.2.2 Extent of exposure

4.2.2.1 Study medication

The extent of exposure to double-blind study medication during the 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 double-blind study medication will be summarized using the Treated Subjects data set for the short-term double-blind treatment period, where the number and percent of subjects with an extent of exposure within pre-specified day ranges will be presented by treatment group. The categorization for the short-term double-blind treatment period of 24 weeks is 1-7, 8-30, 31-60, 61-90, 91-120, 121-180, > 180 days. The mean (SD), median and range of the number of days of exposure will also be presented.

In addition, the exposure in terms of total patient-years will be calculated by treatment group using the sum of the exposure to study medication of all subjects (in years) in a treatment group.

4.2.2.2 Interruption of study medication

Interruption of study drug is indicated by 0 tablets of study medication recorded on the CRF.

4.2.2.3 Current and concomitant medications

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

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

Current medications are defined as medications with a start date prior to the first day of short-term double-blind treatment period and without a stop date prior to the consent date, i.e. current medication will be any medication with at least 1 dose taken on or after the day of consent date up to the day prior to the first dose of study medication.

Concomitant medication during the short-term double-blind treatment period is defined as a medication with either

- a recorded medication start date falling within the short-term double-blind treatment period, or

- a recorded medication start date prior to the first day of study medication during the short-term double-blind treatment period without any recorded medication stop date prior to the start of the short-term double-blind treatment period.

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

Missing and partial date handling of start and stop dates of previous, current and concomitant medications, is described in Section 6.6. The World Health Organization (WHO) dictionary is used to code non-study medications.

4.2.2.4 Measurements of treatment compliance

Percent treatment compliance is calculated during the short-term double-blind treatment period for each treatment group. For each subject, percent compliance is defined as the number of tablets taken divided by the number of tablets that should have been taken. A subject is considered compliant if percent compliance is ≥ 80% and ≤ 120%. The number and percent of subjects compliant during the short-term double-blind treatment period will be displayed for the Safety Analysis Set by treatment group.

Compliance will be calculated for the double-blind study medication (dapagliflozin or placebo, respectively).

Details in calculating percent compliance are specified in Section 6.9.

4.2.3 Efficacy Analysis

This section addresses separately the analyses to be conducted on the primary efficacy variable, on the secondary and exploratory efficacy variables.

4.2.3.1 Primary efficacy analysis

The primary endpoint is the change in HbA1c from baseline to Week 24. The comparison between each dapagliflozin group and placebo are of interest for this study.

The primary endpoint is the change in HbA1c from baseline to Week 24. The primary estimand for the treatment difference at Week 24 if subjects did not discontinue randomized treatment. The primary analysis of the change in HbA1c from baseline to Week 24 will be based on a longitudinal repeated measures analysis using 'direct likelihood' (see Section 4.1.6.1). The analysis will include all available data up to Week 24 or until premature discontinuation of randomized treatment, whichever occurs first. The analysis will be conducted regardless of insulin up-titration.

The model will use subjects in the primary efficacy dataset (i.e., Full Analysis Set) who have a baseline assessment and any post-baseline double-blind treatment period assessment.

For subjects whose post-randomization insulin dose is up-titrated by more than 25% of insulin relative to baseline, measurements obtained after the first instance the 25% up-titration criteria

is met will correspond to rescue and be included in the primary analysis but excluded in one of sensitivity analyses.

Point estimates and 95% confidence intervals for the mean change within each treatment group as well as the difference in mean change between each dapagliflozin treatment group and placebo will be calculated. Student t-statistics corresponding to the Type III sums of squares for the differences in the least squares means between each dapagliflozin group and placebo at Week 24 will be calculated.

To assess the robustness of the primary efficacy analysis for the change in HbA1c from baseline to Week 24, additional sensitivity analyses will be carried out using the following approaches:

- A longitudinal repeated measures analysis, similar to the primary analysis, excluding measurements obtained after the first instance the 25% up-titration criteria is met.
- ANCOVA analysis including subjects in the Full Analysis Set who have a baseline assessment and at least one post-baseline assessment using LOCF methodology for Week 24 (see Section 4.1.5.1).
- ANCOVA analysis including subjects in the Full Analysis Set who have a baseline assessment and a Week 24 assessment (i.e. a "completers" analysis, but without LOCF methodology) (see Section 4.1.5.1).
- A longitudinal repeated measures analysis, similar to the primary analysis, but including subjects in the Per Protocol Analysis Set who have a baseline assessment and any post-baseline assessment. This analysis will be performed only if > 10% of the subjects in any treatment group in the Full Analysis Set have exclusion of the primary efficacy variable data which may have been affected by protocol deviations.
- The ITT estimand, i.e., treatment difference at Week 24 regardless of treatment discontinuation, will be evaluated as the secondary estimand of the primary endpoint, using a longitudinal repeated measures analysis using 'direct likelihood' (see Section 4.1.6.1). The analysis of the ITT estimand will include all available data up to Week 24, regardless of insulin up-titration and regardless of premature discontinuation of randomized treatment.
- A de facto ITT estimand, i.e. treatment difference at Week 24 regardless of whether treatment was actually taken, will be evaluated as another estimand of the primary endpoint. This analysis will include all randomized subjects and all available HbA1c measurements, regardless of treatment adherence. For the subjects who discontinue treatment early and do not have HbA1c measurements at Week 24, their Week 24 HbA1c values will be imputed based on the data from those retrieved subjects in the same treatment group, i.e. those who discontinue treatment early but have HbA1c measurements at Week 24, using a multiple imputation approach. If the number of retrieved subjects is insufficient to inform the imputation, the missing data will instead be imputed based on placebo data only, using a multiple imputation approach. The change from baseline to Week 24 in HbA1c will be analyzed using an ANCOVA model with factors of treatment group and randomization stratification factor (i.e. 1 term for each combination of all stratification factors) and covariate for baseline HbA1c.

The longitudinal repeated measures model is based on a missing at random (MAR) assumption with respect to missing data. The proportion of missing data by visit will be summarized and reasons (i.e. missing visit only, discontinued treatment, or discontinuation from study, including by reason for treatment or study discontinuation) will be assessed to verify the validity of this assumption. If a substantial amount of missing data is observed or if imbalances occur amongst the treatment groups, sensitivity analyses based on a missing not at random assumption (MNAR) assumption will be conducted, as appropriate and feasible based on the observed data. Various methods such as multiple imputation, placebo-based imputation, or tipping point analysis may be considered.

Primary analysis within subgroups

The primary efficacy endpoint of HbA1c will be summarized for the subgroups defined on the basis of the categorized variables listed in Table 7, and by stratification factors (also listed in Table 7).

Table 7 Subgroup Analyses of Primary Efficacy Endpoint for HbA1c

Group variable	Subgroups
Race group	Asian Black or African American White Other
Country	Japan Other
Gender	Female Male
Ethnicity	Hispanic/Latino Non- Hispanic/Latino
Baseline A1C	Categories in Table 5
Subject age	< 35, ≥ 35 and <50, and ≥ 50yrs
Geographic region	See Appendix 9.2
Method of insulin administration	Multiple Daily Injections Continuous Subcutaneous Insulin Infusion
Use of continuous glucose monitoring (CGM)	Yes No
Baseline HbA1c	≥ 7.5% and < 9.0% ≥ 9.0% and ≤ 10.5%

Table 7 Subgroup Analyses of Primary Efficacy Endpoint for HbA1c

Group variable	Subgroups
Baseline Body Mass Index	$\leq 23 kg/m^2$
	$> 23 kg/m^2$ - $\leq 25 kg/m^2$
	$> 25 kg/m^2$ - $\leq 27 kg/m^2$
	$> 27 kg/m^2$ - $\leq 30 kg/m^2$
	$> 30 kg/m^2$

If the value of the group variable cannot be determined, the subject will be excluded from the corresponding subgroup analysis. For each subgroup analysis, if, in any treatment group, the number of subjects is less than 10 for a subgroup, then the summary for that subgroup will be displayed but will not be included in tests for subgroup-by-treatment interaction when the test is based on categorical subgroups (when the test is based on a continuous variable, the subject data will be included).

The subgroup by treatment interaction will be assessed for the primary efficacy endpoint using the longitudinal repeated measures analysis model as described in Section 4.1.6.1 with subgroup, subgroup-by-time, and subgroup-by-time-by-treatment group interaction as three additional effects. In cases that the subgroup is one of the stratification factors, then the combination of stratification factors, as described in section 4.1.6.1 will be based on the two other stratification factors.

When multiple experimental groups are included in a trial, the usual test for subgroup-by-treatment interaction considers all types of interaction without regard for the clinical importance. For example, the following two observed interactions would provide equal evidence for interaction, although the second scenario is usually considered more clinically relevant:

3) larger treatment effect (experimental minus placebo) in males than females for the low dose group, and smaller treatment effect in males than females for the high dose group
4) larger treatment effect in males than females for both dose groups.

Because the usual interaction test gives equal weight to many types of interactions, it has less power to detect the interactions of the greatest clinical relevance. Therefore, a focused interaction test will be performed that examines deviations in the average effect of the experimental groups (relative to control) across subgroups.

Within the framework of a longitudinal repeated measures model, an orthogonal interaction contrast will be used to test the average difference from control across dapagliflozin treatment groups between subgroups. For example, with two dapagliflozin treatment groups and a control group in one study and the group variable is gender, then the contrast coefficients will be as in the following table.

Table 8 Contrast Coefficients for the Example Gender-by-treatment Interaction

Gender	Male	Male	Male	Female	Female	Female
Treatment group	Dapa 5mg	Dapa 10mg	Control	Dapa 5mg	Dapa 10mg	Control
Contrast coefficient for all weeks but Week 24	0	0	0	0	0	0
Contrast coefficient at Week 24	-1/2	-1/2	1	1/2	1/2	-1

For groups with K ($K \geq 2$) subgroups (e.g., race), the subgroup-by-treatment interaction will be tested using a single (K-1)-degree-of-freedom hypothesis. For HbA1c subgroup analysis (for categories as defined in Table 5), the continuous value will be utilized to assess the subgroup by treatment interaction. The categories stated in Table 8 will be used to assess the subgroup by treatment interaction for race, gender, ethnicity, geographic region and female age. For age categories, only < 35, ≥ 35 and <50, and ≥ 50yrs of age will be used to assess the interaction. For the simultaneous (combination) randomization stratification factors analysis (i.e. method of insulin administration, use of CGM, baseline HbA1c) the model will also include one term for each combination of all stratification factor as in primary efficacy analysis. The adjusted mean changes from baseline, SE's, and 95% CIs for each subgroup and the nominal p-value for subgroup by treatment interaction at Week 24 will be calculated.

4.2.3.2 Secondary and exploratory efficacy analysis

Percent change from baseline in total daily insulin dose

The first secondary endpoint is the percent change in total daily insulin dose from baseline to Week 24.

Insulin data are collected on a daily basis over 2-week periods corresponding to distinct study visits/weeks. Each data collection corresponds to insulin doses over a 24-hour period (see implications for baseline and study week definitions in Section 4.1.2.1 and Section 6.3). Information is entered on CRF regarding the total daily bolus and basal insulin administered. The total daily insulin dose is the sum of the total daily bolus and total daily basal insulin doses. If information is not available on basal or bolus insulin dose on a specific day, then the total will not be calculated for that day and it will be considered as missing for the analysis.

In between those 2-week periods of total daily insulin dose data collection, insulin is reported as minimum and maximum values of total daily insulin administration over weekly intervals. The data from these weekly intervals will not be used for the analysis of secondary efficacy endpoint described in this section.

The baseline is defined as mean value of non-missing daily records in 7 days prior to the date of the first dose of the double-blind study medication for this analysis. The midpoint derived

from weekly records in 7 days prior to the first dose date will be considered as a baseline value if the baseline value derived from daily records is missing.

The analysis of the percent change (using logarithmic transformation for the endpoint in the model) in total daily insulin from baseline to Week 24 will be based on a longitudinal repeated measures analysis using 'direct likelihood' as outlined in Section 4.1.6.2.

The model will use subjects in the primary efficacy dataset (i.e., Full Analysis Set) who have at least one baseline assessment and any post-baseline double-blind treatment period insulin dose available.

Point estimates and 2-sided 95% confidence intervals for the mean percent change within each treatment group, as well as the difference in mean percent change between each dapagliflozin treatment group and placebo, will be calculated. Student t-statistics corresponding to the Type III sums of squares for the differences in the least squares means between each dapagliflozin group and placebo at Week 24 will be calculated.

The same analyses will be repeated for the total daily basal and total daily bolus insulin doses. In cases that either of this information is missing on a specific day, then this information will be considered as missing for these two analyses. No p-value will be estimated for these two analyses.

Percent change from baseline in body weight

The second secondary endpoint is the percent change in body weight from baseline to Week 24.

Body weight is collected at selected visits during the short-term period with a single value at each visit.

The analysis of the percent change in body weight from baseline at Week 24 will be based on a longitudinal repeated measures analysis using 'direct likelihood' as outlined in Section 4.1.6.1.

The model will use subjects in the primary efficacy dataset (i.e., Full Analysis Set) who have at least one baseline assessment and any post-baseline double-blind treatment period assessment.

Point estimates and 2-sided 95% confidence intervals for the mean percent change within each treatment group, as well as the difference in mean percent change between each dapagliflozin treatment group and placebo will be calculated. T-statistics corresponding to the Type III sums of squares for the differences in the least squares means between each dapagliflozin group and placebo at Week 24 will be calculated.

The same analysis will be repeated for BMI. No p-value will be estimated for this analysis.

Change from baseline in continuous glucose monitoring (CGM) parameters

A CGM sensor will be inserted subcutaneously on protocol specified study days to allow the collection of information on subject's interstitial glucose level over 2-week intervals at Week -2 to Day 1, Week 10 to Week 12, and Week 22 to Week 24. The system records data approximately every 5 minutes. The data will remain blinded to the subject and to the investigators during the recording and will be downloaded into a data file.

The sensor will be changed after 7 consecutive days of data collection for safety reasons. Each new insertion is followed by a calibration period of a two hours period according to the manufacturer's instructions. Measurements during the calibration period are not considered valid for analysis.

All analysis endpoints will be based on the 24-hour readings for every subject. One value per analysis endpoint will be calculated for every subject for each 24-hour period. These values will be then further summarized to calculate the change from baseline in the respective endpoints.

Glucose values between 40mg/dL – 400mg/dL (2.2mmol – 22.2mmol/L) are the approved and validated glucose range for the Dexcom G4 Platinum CGM device that will be used for this study. As the CGM system will report glucose values < 40mg/dL as "LOW" and glucose values > 400mg/dL as "HIGH", glucose values reported as "LOW" and "HIGH" will be entered the database as 39mg/dL (2.2mmol/L) and 401mg/dL (22.3mmol/L) respectively. Conversion to standard international units will be performed as per BMS standards conversion factor.

The 24-hour period is defined based on the first available time point with a valid CGM glucose reading, for each sensor wear period within each visit period (i.e. each 14 day interval). For each 24-hour interval, data will be used for analysis as long as there are available at least 70% of the glucose readings and any time distance between any two consecutive readings is less than 3 hours. The 70% threshold is chosen in order to avoid an excessive exclusion of subject data due to missingness, balanced with the risk that too low a threshold may result in bias due to potential selectivity of data capture. A large CGM data gap may result in important information being not captured, especially if this occurs around meal time. A low threshold for accepting missingness would result in potentially excluding valuable data from analysis, while a high threshold may compromise the validity of the data. Since periodic calibration of the device will result in 2-hour gaps, a maximum gap of 3 hours was chosen as a reasonable cut-off criterion for the analysis.

In terms of analysis, the percent of the readings during each 24-hour period that fall under, above or between specific ranges provide information of deviations from some clinically relevant targets (Clarke et al 2009). Similarly the mean 24-hour of glucose readings is a "straightforward descriptor of overall glycemic control"(Clarke et al 2009). From an analysis point of view, these are considered as measures of location of the distribution of the glucose readings.

In terms of measures of the variability of glucose measures, the 24-hour standard deviation of glucose readings for each subject will be calculated. This is considered as a measure of "glucose instability"(Service et al 1970). The standard deviation quantifies the dispersion of glucose readings, but does not differentiate between major or minor deviations of glucose readings from the 24-hour average glucose values (Monnier et al 2008). The major flunctuations are considered clinically relevant for dysglycemia, since they are linked to oxidation stress (Monnier et al 2008). For this reason, the 24-hour mean amplitude of glycemic excrusions (MAGE) will also be calculated for each subject for every 24-hour interval. As opposed to the standard deviation that takes into account both minor and major flunctuations, MAGE emphasises on major glucose fluctuations (Monnier et al 2008).

See Section 4.1.2.6 for the calculation of 24-hour daily average CGM values as area under the curve (AUC), Section 4.1.2.7 for the calcuation of the percent of CGM reading falling withing a specified range during the 24-hour period, Section 4.1.2.8 for the 24-hour standard deviation of CGM readings, Section 4.1.2.10 for the postprandial glucose values of CGM readings (exploratory measure), and Section 4.1.2.11 for the calculation of MAGE as described by Service et al.(Service et al 1970).

Change from baseline in the mean value of 24-hour glucose readings, the percent of 24-hour glucose readings that falls within the target range of > 70mg/dL and ≤ 180 mg/dL and the mean amplitude of glucose excursion of 24-hour glucose readings, respectively, is considered as secondary efficacy endpoints, while the standard deviation of 24-hour glucose readings, the percent of 24-hour glucose readings that falls within the range of ≤ 70mg/dL, and the postprandial glucose values derived from CGM readings are considered as exploratory measures. Since the same analysis methods will be used for all of these methods, they are described here.

Each CGM data collection corresponds to data that will be summarized as a 24-hour period (see implications for baseline definitions at Section 4.1.2.1). Subsequently all 24-hour period values within a 14-days interval will include one value for each visit in summaries as specified in section 6.3.

The analysis of the change in each of the CGM parameters from baseline to Week 24 will be based on a longitudinal repeated measures analysis using 'direct likelihood' as outlined in Section 4.1.6.2.

The model will use subjects in the primary efficacy dataset (i.e., Full Analysis Set) who have at least one baseline assessment and any post-baseline double-blind treatment period assessment.

Point estimates and 2-sided 95% confidence intervals for the mean change within each treatment group, as well as the difference in mean change between each dapagliflozin treatment group and placebo, will be calculated. Only for the secondary endpoints, t-statistic corresponding to the Type III sums of squares for the differences in the least squares means

between each dapagliflozin group and placebo at Week 24 will be calculated. No p-value will be estimated for the exploratory analyses.

Proportion of subjects achieving glycemic response

The proportion of subjects having an HbA1c reduction from baseline to Week 24 (LOCF) ≥ 0.5% and without severe hypoglycemia events is a secondary combined endpoint.

The proportion will be analyzed using the methodology outlined in Section 4.1.8. The odds ratio and its 2-sided 95% confidence interval between each dapagliflozin treatment group and placebo, will be calculated. The analysis will use subjects in the primary efficacy dataset (i.e., Full Analysis Set) who have at least one baseline assessment and any post-baseline double-blind treatment period assessment.

The proportion of subjects having an HbA1c reduction from baseline to Week 24 LOCF ≥ 0.5% and an HbA1c Week 24 LOCF value < 7% are separate exploratory endpoints; therefore no statistical testing will be performed to compare each dapagliflozin group and placebo.

Number of non-severe hypoglycemic events per subject in those achieving a HbA1c reduction of ≥ 0.5% from baseline to the week 24 visit (LOCF) is an exploratory endpoint; therefore no statistical testing will be performed to compare each dapagliflozin group and placebo.

Analyses will be analyzed using the methodology outlined in Section 4.1.8 using the randomized subjects dataset and the LOCF visit. No p-value will be calculated. The analysis will use subjects in the primary efficacy dataset (i.e., Full Analysis Set) who have at least one baseline assessment and any post-baseline double-blind treatment period assessment.

Change from baseline in fasting plasma glucose

The change in fasting plasma glucose from baseline to Week 24 is an exploratory endpoint; therefore no statistical testing will be performed to compare each dapagliflozin group and placebo.

Fasting plasma glucose is collected at selected visits during the short-term double-blind treatment period with a single value at each visit.

The analysis of the change in fasting plasma glucose from baseline to Week 24 will be based on a longitudinal repeated measures analysis using ‘direct likelihood’ as outlined in Section 4.1.6.1.

The model will use subjects in the primary efficacy dataset (i.e., Full Analysis Set) who have at least one baseline assessment and any post-baseline double-blind treatment period assessment.

Point estimates and 2-sided 95% confidence intervals for the mean change within each treatment group, as well as the difference in mean change between each dapagliflozin treatment group and placebo, will be calculated.

Change from baseline in 6-point self monitored blood glucose (SMBG) parameters.

The change average glucose values and postprandial glucose values measured by 6-point SMBG from baseline to Week 24 are exploratory endpoints; therefore no statistical testing will be performed to compare each dapagliflozin group and placebo.

See Section 4.1.2.9 for the calculation of the average daily glucose value of 6-point SMBG, and Section 4.1.2.10 for the postprandial glucose values of 6-point SMBG.

The 6-point SMBG parameters should include capillary blood glucose values. In cases that there are capillary plasma glucose values reported, they will be converted to capillary blood equivalent values as described in section 6.12.

The analysis of the change in average glucose values from baseline to Week 24 will be based on a longitudinal repeated measures analysis using 'direct likelihood' as outlined in Section 4.1.6.2.

The analysis of the change in postprandial glucose values from baseline to Week 24 will be based on a longitudinal repeated measures analysis using 'direct likelihood' as outlined in Section 4.1.6.2.

The model will use subjects in the primary efficacy dataset (i.e., Full Analysis Set) who have at least one baseline assessment and any post-baseline double-blind treatment period assessment.

Point estimates and 2-sided 95% confidence intervals for the mean change within each treatment group, as well as the difference in mean change between each dapagliflozin treatment group and placebo, will be calculated.

Change from baseline in seated systolic blood pressure

The change in seated systolic blood pressure from baseline to Week 24 is an exploratory endpoint; therefore no statistical testing will be performed to compare each dapagliflozin group and placebo in subjects with hypertension at baseline.

Seated systolic blood pressure is collected at selected visits during the short-term double-blind treatment period with a single value at each visit.

The analysis of the change in seated systolic blood pressure from baseline to Week 24 will be based on a longitudinal repeated measures analysis using 'direct likelihood' as outlined in Section 4.1.6.1.

The model will be restricted to data from subjects in the primary efficacy dataset (i.e., Full Analysis Set) who have at least one baseline assessment and any post-baseline double-blind treatment period assessment and who have hypertension at baseline (defined as seated SBP ≥ 140mmHg and/or seated DBP ≥ 90mmHg).

Point estimates and 2-sided 95% confidence intervals for the mean change within each treatment group, as well as the difference in mean change between each dapagliflozin treatment group and placebo, will be calculated.

Time to insulin up-titration more than 25% from baseline

Time to the first post-randomization insulin up-titration by more than 25% relative to the baseline insulin dose will be presented by treatment group using a Kaplan-Meier curve using methodology described in Section 4.1.7. A plot will be presented only when there are at least 5 subjects meeting the criteria in one treatment group. An accompanying table of the cumulative proportion of subjects whose insulin was up-titrated by more than 25% from baseline at specific time points during the short-term double-blind treatment period will be produced.

4.2.3.3 Outcomes Research Analyses

Health status measured with EQ-5D-3L questionnaire

Health status is measured at baseline and Week 24 of the short-term double-blinded treatment period by the EQ-5D-3L questionnaire. The EQ-5D-3L is a generic, preference-based utility questionnaire and consists of two parts, the EQ visual analogue scale (VAS) and the EQ-5D-3L index (Kind 1996). The EQ VAS ranges from 0 = worst possible health to 100 = best possible health. The EQ-5D-3L index is a five-dimension questionnaire. The dimensions consist of mobility, self-care, usual activity, pain/discomfort and anxiety/depression. Each item has three levels: no problems, some problems and severe problems.

The change to from baseline to Week 24 LOCF in the EQ VAS and in EQ-5D-3L index will be analyzed separately, using ANCOVA models as outlined in Section 4.1.5.1. Point estimates and 2-sided 95% confidence intervals for the mean change within each treatment group, as well as the difference in mean change between each dapagliflozin treatment group and placebo, will be calculated. No p-values will be generated.

Proportions of subjects with assessments at baseline, Week 24 (LOCF), as well as a shift in the classification of the 5 dimensions of EQ-5D from baseline to Week 24 (LOCF), will be summarized.

All above mentioned analyses will use subjects in the primary efficacy dataset (i.e., Full Analysis Set) who have at least one baseline assessment and any post-baseline double-blind treatment period assessment.

Treatment Satisfaction measured with Diabetes Treatment Satisfaction Questionnaire status questionnaire

The DTSQ has been developed to assess patient's satisfaction with treatment and perception of change in hyper- and hypoglycaemia (Bradley 1994). The DTSQ status version (DTSQs) has 8 items. Each item has seven levels.

The change from baseline to Week 24 LOCF in overall DTSQs score and in individual item scores (which include scores for perceived frequency of hyperglycemia, for perceived frequency of hypoglycemia, for satisfaction with treatment, for convenience with treatment, for flexibility with treatment, for satisfaction with understanding of diabetes, for willingness to recommend treatment to someone else and for satisfaction to continue with current treatment) will be analyzed separately, using ANCOVA models as outlined in Section 4.1.5.1. Point estimates and 2-sided 95% confidence intervals for the mean change within each treatment group, as well as the difference in mean change between each dapagliflozin treatment group and placebo, will be calculated. No p-values will be generated.

All above mentioned analyses will use subjects in the primary efficacy dataset (i.e., Full Analysis Set) who have at least one baseline assessment and any post-baseline double-blind treatment period assessment.

Hypoglycemia Fear Survey II Worry Subscale

The HFS-II Worry Subscale is an 18-item questionnaire, in which responses are made on a 5-point Likert scale (where 0 = Never and 4 = Always). The HFS-II Worry Subscale has a score ranging from 0 to 72, with higher scores indicating increased fear of hypoglycemia.

The change from baseline to Week 24 LOCF in HFS-II score will be analyzed, using ANCOVA models as outlined in Section 4.1.5.1. Point estimates and 2-sided 95% confidence intervals for the mean change within each treatment group, as well as the difference in mean change between each dapagliflozin treatment group and placebo, will be calculated. No p-values will be generated.

All above mentioned analyses will use subjects in the primary efficacy dataset (i.e., Full Analysis Set) who have at least one baseline assessment and any post-baseline double-blind treatment period assessment.

4.2.4 Safety

This safety section describes safety analyses that will be conducted for the short-term double-blind treatment period.

The Treated Subjects data set will be used in all summaries of safety data in the short-term double-blind treatment period.

4.2.4.1 Adverse events

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.

No statistical test will be performed to compare adverse event rates between treatment groups. This policy was adopted to recognize the lack of power and the potential for misleading

interpretation based on repeated statistical tests which increase the family wise Type I error level.

Counting rules for adverse events are described in Section 6.7.

In summaries by SOC and PT, AEs will be sorted by decreasing frequency within each SOC and PT according to the highest Dapagliflozin dose group. In summaries by PT, AEs will be sorted by decreasing frequency within PT according to the highest dapagliflozin dose 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 All Adverse Events). Separate summaries will be provided including hypoglycemia events reported on that special CRF pages (see Section Hypoglycemia).

All Adverse Events

All adverse events (serious and non-serious, excluding hypoglycemic and DKA events that are not reported as SAEs) with onset during the short-term double-blind treatment period will be summarized by SOC, PT, and treatment group, for both the primary and sensitivity safety analyses. In addition, a subject listing of all reported adverse events will be produced, displaying all events (including pre-treatment events) that occurred prior to the start date of long-term treatment period. All adverse events (serious and non-serious) including all hypoglycemic events will also be summarized by treatment group.

AEs and SAEs with an onset from Day 1 of the short-term double-blind treatment period 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.

In addition, number of subjects reporting any of the following during the double-blind treatment period will be summarized for all treated subjects:

- at least one adverse event
- at least one related adverse event
- deaths
- at least one SAE
- at least one related SAE
- SAE leading to discontinuation of study medication
- AE leading to discontinuation of study medication
- At least one SAE of hypoglycemia
- Hypoglycemia SAE leading to discontinuation of study medication

- At least one SAE of DKA
- DKA SAE leading to discontinuation of study medication

This summary will be repeated for Japanese subjects.

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):

- Proportion of subjects with adverse events by SOC and PT in subgroups of subjects defined by age category (<35 years, >=35 and <50 years, and >=50 years), gender, and race.
- Most common adverse events by PT and treatment group (i.e., reported by ≥ 5% of subjects in any treatment group),
- Adverse events by SOC, PT, intensity and treatment group,
- Adverse events related to study medication by SOC, PT, and treatment group.

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 according to standard sponsor algorithms 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 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. This data will be presented as the rate per 100 years of patient exposure. Exposure to study medication will be calculated according to approved standard sponsor algorithms as well. As an example, if 5 patients report 7 unique episodes of headache and had a combined cumulative exposure of 20 years to study medication, the incidence rate is reported as 7 / 20 * (100) or 35 cases per 100 patient years of exposure.

Additionally, a table 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. No formal comparisons are made between treatments. No formal statistical testing will be performed, only summary statistics are provided.

For the purpose of clinical trial transparency, listing of most common non-serious AEs will be produced for posting at public websites (e.g., NIH's www.ClinialTrials.gov).

Deaths

All deaths recorded on the status page, the AE page, or SAE page (with a death date, cause of death, outcome or SAE categorization present) of the CRF will be considered a death in the analyses. Any deaths that occur during the study will be described in depth as narrative in the

CSR. A listing of all deaths that occur prior to the start date of long-term treatment period will be produced.

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 the short-term double-blind treatment period 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 occurring during the short-term double-blind treatment period will be summarized by SOC, PT, and treatment group for both the primary and sensitivity safety analyses. In addition, the proportion of subjects with related SAEs will be presented by SOC, PT, and treatment group. A listing of all SAEs will be produced, displaying all SAEs (including pre-treatment events) that occurred prior to the start date of long-term treatment period.

4.2.4.2 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 SOC, PT, and treatment group for both the primary and sensitivity safety analyses. This summary will include hypoglycemia and DKA events that are reported as SAEs. When summarizing AEs leading to discontinuation, no upper cutoff day windows (i.e. 4 days and 30 days from last dosing date in the short-term double-blind treatment period for AEs and SAEs respectively) are applied. For short-term double-blind period analyses, the only upper cutoff date is the last date of the short-term double-blind treatment period.

In addition, a subject listing of discontinuation due to AEs will be provided, displaying all events that led to discontinuation with an onset date prior to and including the last date of short-term double-blind treatment period, if any.

4.2.4.3 Confirmed adjudicated cardiovascular adverse events

Cardiovascular events will be independently adjudicated by a Clinical Event Committee (CEC) blinded to the treatment of the subjects in accordance with a dedicated Clinical Event Committee Charter/Manual of Operations: Dapagliflozin Program. All adjudicated cardiovascular events will be summarized by treatment group.

4.2.4.4 Adverse events of special interest

Separate summaries will be provided for the following AEs of special interest. To identify each type of AE of special interest in this section, a list of PTs will be selected before database lock and unblinding of the database. Unless otherwise specified, AEs and SAEs of special interest with an onset from Day 1 of the short-term double-blind treatment period 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.

Hepatic adverse events

To facilitate monitoring of liver safety in subjects treated with dapagliflozin, a list of PTs will be selected before database lock of the study to form the Standard MedDRA Query (SMQ) for events of hepatic disorders. The number and percentage of subjects with events of hepatic disorders will be summarized by PT and treatment group in the short-term double-blind treatment period.

Genital Infection

The definition of events of genital infection is based on the dapagliflozin predefined list of events.

The number and percentage of subjects with events of genital infections will be summarized by PT and treatment group in the short-term double-blind treatment period. The number and percentage of subjects with events of genital infections will also be summarized by gender and treatment group in the short-term double-blind treatment period.

The number and percentage of subjects with events of genital infection will be summarized for the subgroups defined on the basis of categorized variables including the frequency of events per subject (recurrence; 1, 2, 3 or > 3), antimicrobial treatment (yes, no or unknown), and additional treatment (yes, no or unknown).

The number and percentage of subjects with events of genital infections qualifying as an SAE will be summarized by PT and treatment group in the short-term double-blind treatment period.

The number and percentage of subjects with events of genital infections leading to discontinuation of double-blind study medication will be summarized by PT and treatment group in the short-term double-blind treatment period.

Time to first event of genital infection will be presented by treatment group using a Kaplan-Meier curve using methodology described in Section 4.1.7. A plot will be presented only when there are at least 5 events in one treatment group. An accompanying table of the cumulative proportion of subjects with time to first genital infections event at specific time points during the short-term double-blind treatment period will be produced.

Urinary-tract Infection

The definitions of events of urinary-tract infection and its subset of kidney infections are based on the dapagliflozin predefined list of events.

The number and percentage of subjects with events of urinary-tract infection and with kidney infections will be summarized by PT and treatment group in the short-term double-blind

treatment period. The number and percentage of subjects with urinary-tract infection will also be summarized by gender and treatment group in the short-term double-blind treatment period.

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 the frequency of events per subject (recurrence; 1, 2, 3 or > 3), antimicrobial treatment (yes, no, or unknown), additional treatment (yes, no, or unknown).

The number and percentage of subjects with events of urinary-tract infection qualifying as an SAE will be summarized by preferred term and treatment group in the short-term double-blind treatment period.

The number and percentage of subjects with events of urinary-tract infection leading to discontinuation of double-blind study medication will be summarized by PT and treatment group in the short-term double-blind treatment period.

Time to first event of urinary-tract infection will be presented by treatment group using a Kaplan-Meier curve using methodology described in Section 4.1.7. A plot will be presented only when there are at least 5 events in one treatment group. An accompanying table of the cumulative proportion of subjects with time to first urinary-tract infection event at specific time points during the short-term double-blind treatment period will be produced.

A listing of subjects who undergo urine culture will be provided.

Diabetic Ketoacidosis

A DKA Adjudication Committee, blinded to the treatment of the subjects, will independently adjudicate all the DKA events reported by the investigators and those identified based on pre-defined algorithm during the study period. A separate Adjudication Manual will define and describe the procedure for the handling, reporting, and classification of these cases.

DKA events with an onset from Day 1 of the lead-in period up to and including 1 day prior to start of short-term double blind treatment period will be considered as occurring during lead-in period, whereas DKA events with an onset from Day 1 of the short-term double-blind treatment period up to and including 4 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. The proportion of subjects with DKA events will be tabulated by treatment group in the lead-in period and the short-term double-blind treatment period.

SAEs of DKA with an onset during the short-term double-blind treatment period and leading to discontinuation of study medication will be summarized by treatment group. When summarizing DKA events leading to discontinuation, no upper cutoff day windows are applied.

The number and percentage of subjects with DKA events will be summarized by frequency of events per subject (recurrence; 1, 2, 3, > 3). The total number of events by treatment group will be summarized for the lead in period and the short-term double-blind treatment period.

The exposure adjusted incidence rate of definite DKA events (including recurrences) will be summarized by monthly intervals in order to assess any over time changes in the incidence rate.

Time to first DKA event will be presented by treatment group using a Kaplan-Meier curve using methodology described in Section 4.1.7. A plot will be presented only when there are at least 5 events in one treatment group. An accompanying table of the cumulative proportion of subjects with time to first DKA event at specific time points during the short-term double-blind treatment period will be produced.

A listing of subjects will be produced and it will display all confirmed adjudicated DKA events with an onset date/time from the start date/time of double-blind treatment period up to the start date of the long-term treatment period.

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 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.

Hypoglycemic events will be categorized using the following classes following ADA recommendations (Workgroup on Hypoglycemia, American Diabetes Association 2005):

1) **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." (Workgroup on Hypoglycemia, American Diabetes Association 2005)
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)." (Workgroup on Hypoglycemia, American Diabetes Association 2005)
3) **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." (Workgroup on Hypoglycemia, American Diabetes Association 2005)

4) **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." (Workgroup on Hypoglycemia, American Diabetes Association 2005)
5) **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." (Workgroup on Hypoglycemia, American Diabetes Association 2005)

The above mentioned classification refers to plasma glucose measurements. In cases where whole blood glucose measurements are reported, they will be converted to plasma equivalent as described in section 6.12.

All analyses of hypoglycemic events will be performed both overall and by class of events as per ADA categorization (Workgroup on Hypoglycemia, American Diabetes Association 2005) presented above.

The proportion of subjects with hypoglycemic events will be tabulated by treatment group in the short-term double-blind treatment period, both regardless of insulin up-titration and, separately, including data only prior to the first post-randomization insulin up-titration by more than 25% relative to baseline.

SAEs of hypoglycemia with an onset during the short-term double-blind treatment period and those leading to discontinuation of study medication will be summarized by treatment group. When summarizing hypoglycemic events leading to discontinuation no upper cutoff day windows are applied. For short-term double-blind period analyses the only upper cutoff date is the start date of the long-term treatment period.

Time to hypoglycemic events leading to discontinuation of study medication will be presented by treatment group using a Kaplan-Meier curve using methodology described in Section 4.1.7. A plot will be presented only when there are at least 5 events in one treatment group. An accompanying table of the cumulative proportion of subjects with time to hypoglycemic

events leading to discontinuation of study medication at specific time points during the short-term double-blind treatment period will be produced.

The number and percentage of subjects with hypoglycemic events will be summarized by frequency of events (recurrence; for severe hypoglycemia: 1, 2, ≥ 3; for documented symptomatic hypoglycemia: 1-3, 4-6, or ≥ 7, for asymptomatic, probable symptomatic, relative hypoglycemia and all hypoglycemic events independently of ADA categorization: 1-5, 6-9, or ≥ 10). The total number of events by treatment group will be summarized for the short-term double-blind treatment period.

The exposure-adjusted incidence rate of hypoglycemic events (including recurrences) will be summarized by monthly intervals in order to assess any over time changes in the incidence rate.

In order to explore any differences in the distribution of hypoglycemic events by randomization strata categories, the proportion of subjects with hypoglycemic events will be summarized for the whole duration of the double-blind treatment period by randomization stratification subgroups. Additionally, this analysis will be repeated by subgroups of change in HbA1c from baseline to Week 24 LOCF (i.e. reduction ≥ 0.5% vs. reduction < 0.5%) and HbA1c week 24 LOCF value (i.e. week 24 LOCF value < 7% vs. week 24 LOCF value ≥ 7%).

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.

Hypotension/Dehydration/Hypovolemia

The definitions of hypotension/dehydration/hypovolemia events are based on the dapagliflozin predefined list of events.

The number and percentage of subjects with hypotension/dehydration/hypovolemia events will be summarized by PT and treatment group in the short-term double-blind treatment period.

Renal Impairment/Failure

The definitions of renal impairment/failure events are based on the dapagliflozin predefined list of events.

The number and percentage of subjects with renal impairment/failure events will be summarized by PT and treatment group in the short-term double-blind treatment period.

Fractures

The definitions of events of fracture are based on the dapagliflozin predefined list of events.

The number and percentage of subjects with events of fracture will be summarized by PT and treatment group in the short-term double-blind treatment period.

Hypersensitivity

The definition of hypersensitivity is based on the SMQ of hypersensitivity narrow list.

The number and percentage of subjects with hypersensitivity will be summarized by PT and treatment group in the short-term double-blind treatment period.

4.2.4.5 Laboratory evaluation

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

For liver safety, a summary of proportion of subjects with elevated liver test including elevated aminotransferases (AT; alanine aminotransferase [ALT] and/or aspartate aminotransferase [AST]) and total bilirubin (see Appendix 3 for definition) will be provided. In addition, a summary of proportion of subjects with elevated liver tests and/or reported AE of hepatic disorder will also be provided. For those summaries, liver test data obtained after the start of study medication dosing up to and including 30 days after the last short-term double-blind dosing date (or up to the start of the long-term treatment period, whichever comes first) will be considered as obtained during the short-term double-blind treatment period.

For self-monitored blood ketones that are not associated with an event sent for DKA adjudication, number of subjects with at least one self-monitored blood ketone result and number of subjects with at least one self-monitored blood ketone result elevated (defined as beta-hydroxybutyrate > 3.0 mmol/L) will be summarized by treatment group. In addition, the proportion of subjects with at least one self-monitored blood ketone result elevated will be analyzed using a logistic regression procedure, adjusting for baseline HbA1c, percent of insulin reduction at start of treatment from baseline, age, gender and randomization strata.

Marked laboratory abnormalities

Laboratory abnormalities will be evaluated based on marked abnormality (MA) values. The pre-defined criteria for marked abnormalities are detailed in Appendix 9.3. 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 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.

Additionally, for each subject with a MA for a parameter, all the subject's values of that parameter over the short-term double-blind treatment period and the follow-up period, where applicable, will be listed.

Change from baseline and percent change from baseline for selected laboratory parameters over time

All analyses of laboratory data will use observed data regardless of insulin up-titration that occurs post-randomization. Visit windows are provided in Section 6.3 in order to link each laboratory test to a scheduled visit. Change from baseline during the short-term double-blind treatment period for selected laboratory parameters will be summarized by treatment group using descriptive statistics:

- hematocrit
- hemoglobin
- platelet count
- white blood cell (WBC) count
- Red blood cell (RBC) count
- total bilirubin
- alanine aminotransferase (ALT)
- alkaline phosphatase
- aspartate aminotransferase (AST)
- blood urea nitrogen (BUN)
- Measured creatinine clearance or estimated creatinine clearance (if measured creatinine clearance is not available) using the method of Cockcroft and Gault (change):

$$\mathrm{CrCl} = \frac{(140 - \mathrm{age}) \cdot (\mathrm{weight\,in\,kg.}) \cdot (0.85\,\mathrm{for\,females})}{\mathrm{serum\,creatinine(mg/dL)} \cdot 72}$$

- creatine kinase (creatine phosphokinase) (CK)
- creatinine, serum (S_{cr})
- electrolytes - sodium, potassium, bicarbonate, chloride, magnesium and calcium
- total protein, serum
- albumin
- inorganic phosphorus
- albumin to creatinine ratio

- glucose:creatinine ratio
- spot urine for glucose
- triglycerides
- low-density lipoprotein cholesterol
- high-density lipoprotein cholesterol
- total cholesterol
- free fatty acid
- uric acid

Additional laboratory data summaries

A 24-week short-term double-blind treatment period, the following summaries will be provided.

Shift Tables for Electrolytes (Sodium, Potassium, Calcium, Phosphate, and Magnesium) Categories

Shift tables of Treated Subjects Data Set subjects with electrolytes values in categories of low, normal, and high (based on normal range of central laboratory) will be summarized by treatment group using the highest (for sodium, calcium, phosphate, and magnesium) and lowest (for sodium, potassium, and calcium) values (regardless of post-randomization insulin up-titration) obtained during the short-term double-blind treatment period.

Shift Tables for Urine Albumin Excretion Categories

Shift tables of Treated Subjects Data Set subjects with urine albumin to creatinine ratios in 0 - < 30mg/g (normoalbuminuria), 30 - < 300mg/g (microalbuminuria), and ≥ 300mg/g (macroalbuminuria) will be summarized by treatment group using the Week 24 values (the last observation regardless of post-randomization insulin up-titration prior to Week 24 will be used if no Week 24 measurement is available).

4.2.4.6 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 the start of the long-term treatment period, whichever comes first, will be considered as obtained during the short-term double-blind treatment period.

Visit windows are provided in Section 6.3 in order to link each vital sign measurement to a scheduled visit. Measured values and changes from baseline in vital sign measurements will be summarized by treatment group at each scheduled visit using descriptive statistics (using available data regardless of insulin up-titration for subjects in Treated Subjects Data Set).

4.2.4.7 Physical examination

All physical examination data and abnormalities will be listed. Shift tables from baseline to end of short-term double blind treatment period of normal vs. abnormal assessments will be

provided (the last observation regardless of rescue prior to week 24 will be used if no week 24 measurement is available).

4.2.4.8 Electrocardiograms

The normality/abnormality of the ECG tracing, as determined by the investigator, will be summarized using frequency tables on number of subjects who have a normal/abnormal ECG tracing at Week 24 of the short-term double-blind treatment period, overall and by the ECG tracing at baseline. When the data at Week 24 is not available for a discontinued subject, then the last observation before discontinuation of that subject (regardless of insulin up-titration) will be used for summary.

4.2.4.9 Pregnancy test results

By-subject listing of pregnancy test results will be provided using Safety Analysis Set.

5. INTERIM ANALYSES

No interim analysis is planned for this study.

6. CONVENTIONS

6.1 Duration of type 1 diabetes

Duration of T1DM is calculated as the number of years from the T1DM 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 T1DM 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 T1DM.
- Missing day, month and year: No imputations will occur, and the subject will be excluded from all summaries related to duration of T1DM.
- 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.

6.2 Missing and multiple measurements

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

For all analyses and summaries of efficacy, safety, outcome research measures, missing values will not be imputed, except those specified to be analyzed using LOCF methodology, that are imputed by the LOCF.

For HFS-II worry subscale, an item mean score will be used to impute the summary score if more than 75% of items have responses (i.e., at least 14 out of 18 items have responses). Instead of dividing by the entire subscale to compute the item mean, only divide by the number of questions answered. The item mean score is converted to summary score by multiplying 18. If less than 14 items are available, the summary score will be counted as missing.

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. The selection of laboratory result for analysis for this subject will follow AZ global standard.

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.

6.3 Longitudinal assessments

Day 1 for the short-term double-blind treatment period is the start date of double-blind treatment medication. The observation closest to the target day (and time where applicable) is the measurement used in the analysis for each scheduled visit.

The following visit window will apply for all laboratory parameters, vital signs and physical measurements in the short-term double-blind treatment period up to 24 weeks.

Table 9 Visit Windows for Short-term Double-blind Treatment Period Up to Week 24

Visit	Target Day	Day Range
Week 1	8	Post-baseline -11*
Week 2	15	12-22*
Week 4	29	23-43*
Week 8	57	44-71*
Week 12	85	72-106*
Week 18	127	107-148*

Table 9 Visit Windows for Short-term Double-blind Treatment Period Up to Week 24

Visit	Target Day	Day Range
Week 24	169	149 - Last day of double-blind treatment + 8 day for HbA1c, body weight, BMI, patient reported outcomes or 149 - Last day of double-blind treatment +1 day for insulin dose, CGM parameters, FPG, SMBG

* on or before the last day of double-blind treatment.

Table 10 Visit Windows for Short-term Double-blind Treatment Period Up to Week 24 for CGM parameters, Total daily Insulin and 6-point SMBG

Visit	Target Day	Day Range
Week 2	15	1-40*
Week 12	85	60-110*
Week 24	169	140- Last day of double-blind treatment

* on or before the last day of double-blind treatment.

The use of post double-blind treatment observations is specified in Section 6.4.

For all safety laboratory parameters, the use of post-double-blind treatment observations is specified in Section 4.2.4.5.

For laboratory and non-laboratory parameters, if a subject has more than one measurement included within a window, the assessment closest to the target day will be used. In case of ties between observations located on different sides of the target day, the earlier assessment will be used. In case of ties located on the same side of the target day (i.e. more than one value for the same day but different time), the value with the earlier entry date/time will be used.

For measurements referring to 24-hour collection periods (i.e. daily insulin dose) and those referring to multiple data collection points over multiple days (i.e. 6-point measurement parameters and CGM parameters) more than one value will be included in the analysis window. The value for analysis for the visit window will be calculated as the daily average of all available values for analysis for each parameter falling in the analysis window.

For the shift tables for nominal ECG parameter of normal/abnormal the following rules apply. If more than one ECG measurement is performed within a specified time window, an abnormal ECG value will be chosen over a normal ECG value. If more than two abnormal or

two normal ECG values are within the same time window, the one closest to the target date will be used.

6.4 Post-treatment observations

While short-term double-blind treatment period (up to the start of long-term treatment period) efficacy observations will be listed regardless of whether the subject was taking blinded 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 the start of long-term treatment period) as indicated below:HbA1c, body weight, BMI, patient reported outcomes, will be summarized/analyzed only if measured on or prior to the 8th day after the last double-blind drug treatment date. Please note that the analysis of the ITT estimand for HbA1c will include all available HbA1c data up to Week 24, regardless of premature discontinuation of randomized treatment.

Insulin dose, CGM parameters, FPG, SMBG will be summarized only if measured on or prior to the first day after the last double-blind drug treatment date.

6.5 Assignment of doses to adverse events and laboratory assessments

In case of missing dates and/or times, 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 and times as recorded on the CRF. In this context,

- Unless otherwise specified in study-specific SAP, 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.
- the 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.
- if a dosing time is missing for a starting dose of study drug in a study period or a dose change, then it is defaulted to 9:00 if there are AM and BID dosing in the study, and it is defaulted to 19:00 if there is only PM dosing in the study.
- if there's a gap in the dosing information, then the dosing time corresponding to the start date of the gap, is defaulted to 00:00.

If a dosing time is missing for a last stop dosing date for the study drug in a study period, then it is defaulted to 23:59 when there is no study medication in the next study period, or is defaulted to 1 minute before the starting dose time of the study medication in the next study period when available.

6.5.1 Treatment at onset of an adverse event

Both onset date and (if requested) onset time of an AE will be compared to the dosing information. Treatment will then be the dose at the last observation for a subject in the dosing file where the AE onset date/time is not earlier than the dose date/time.

6.5.2 Treatment at the time of a laboratory assessment

Laboratory draw date and time will both be compared to the dosing information. Treatment at the time of the laboratory assessment will then be the dose at the last observation for a subject in the dosing file where the laboratory date/time is not earlier than the dose date/time.

6.6 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.

6.7 Counting rules for adverse events

6.7.1 At subject level

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 and time
  a) When assessing relationship to study medication, relationship as reported by the investigator, events will be reported as related or not to study medication. Related events will take precedence over unrelated events in determining the event to include in summary tables.
  b) More intense events will take precedence over less intense events in determining the event to include in summary tables.
  c) Earlier onset date-time events will take precedence over late onset date-time 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 and time

6.7.2 At event level

At event level, each unique AE record will be counted. Unique AE record can be obtained by collapsing all AE records following a standard algorithm. In addition to frequency summary, exposure adjusted summary will also be provided where the overall exposure of a subject is calculated from first dose day to the last day of short-term treatment regardless of whether a subject has had an event or not.

6.8 Fasting state

Lipids parameters listed in the protocol include TG, LDL-C, HDL-C, total-C and FFA. For TG, calculated LDLC and FFA, only data collected in fasting state will be used for analysis. For direct LDL-C, HDL-C and total-C, data collected in both fasting and non-fasting state will be used for analysis.

For FPG and fasting C-peptide, only assessments documented with the subject in fasting state will be summarized and listed.

6.9 Percent compliance calculation

Percent compliance for the double-blind study medication during the double-blind treatment period is defined for each subject as the number of tablets taken, divided by the number of tablets that should have been taken.

The number of tablets that should have been taken is calculated as 1 + the number of days from the first double-blind treatment period dose recorded in the “Record of Study Medication” to the last 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 double-blind treatment dose.

6.10 Strip sign for selected laboratory data

For selected laboratory test values that have been received with an operator sign as a part of the result (>, ≥, <, or ≤), 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.

The applicable impacted laboratory tests and applicable operator signs will be identified during the course of the study and study monitoring. For laboratory parameters not included in any statistical analyses, operator signs will not be stripped and the value will counted as missing.

6.11 United States Conventional Units and Standard International Units for Laboratory Data

Unless otherwise specified, all analyses of laboratory data (independently of being safety or efficacy) will be performed in both United States Conventional Units (US) and Standard International Units (SI). Conversion between US and SI, and vice versa, will be performed using the BMS standard conversion factors.

Analysis of HbA1c and CrCl (Cockcroft and Gault) will be performed only in United States Conventional Units as these units are accepted by Health Authorities in other countries.

6.12 Capillary glucose readings of plasma glucose values and whole blood glucose values

Capillary Glucose reading values in plasma are typically about 12% higher than readings based on whole blood. When necessary, the capillary whole blood glucose reading values will be converted to capillary plasma equivalent glucose reading values by multiplying them with the conversion factor of 1.12. Respectively, when necessary, the capillary plasma glucose reading values will be converted to capillary whole blood glucose reading equivalent values by dividing them with the same conversion factor.

6.13 Laboratory evaluations

All laboratory evaluations performed by central laboratories and local laboratories that are included in the database will be included in summary tables and in the listings.

7. CHANGES OF ANALYSIS FROM PROTOCOL

1) Self-monitored blood ketones that are not associated with an event sent for DKA adjudication will be analysed.

8. REFERENCES

Bradley 1994
Bradley, C 1994, 'The Diabetes Treatment Satisfaction Questionnaire: DTSQ'. in Bradley C (Ed) (1994), Handbook of Psychology and Diabetes: a guide to psychological measurement in diabetes research and practice. Harwood Academic Publishers.

Clarke et al 2009
Clarke, William and Kovatchev, Boris. Statistical tools to analyze continuous glucose monitor data. Suppl 1, Jun 2009, Diabetes technology & therapeutics, Vol. 11, pp. S45-54.

Dunnett and Tamhane 1992
Dunnett CW and Tamhane AC. A step-up multiple test procedure. J Am Stat Assoc 1992;87:162-170.

Greenwood 1926
Greenwood M. The natural duration of cancer. Reports on public health and medical subjects. 1926; 33:1-26.

Kaplan and Meier 1958
Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. Journal of American Statistical Association. 1958; 53:457-481.

Kind 1996
P.Kind. The EuroQol instrument: An index of health related quality of life. Quality of life and PharmacoEconomics in clinical trials. Second edition, edited by B.Spiker. Lippincott-Raven Publishers. Philadelphia 1996.

Monnier et al 2008
Monnier, Louis, Colette, Claude and Owens, David R. Glycemic variability: the third component of the dysglycemia in diabetes. Is it important? How to measure it? 6, Nov 2008, Journal of diabetes science and technology, Vol. 2, pp. 1094-100.

Service et al 1970
Service, F J, et al. Mean amplitude of glycemic excursions, a measure of diabetic instability. 9, Sep 1970, Diabetes, Vol. 19, pp. 644-55.

Workgroup on Hypoglycemia, American Diabetes Association 2005
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.

9. APPENDIX

9.1 List of significant protocol deviations

Number	Significant Deviation
1	Randomized subjects with no signed and dated written informed consent.
2	Randomized subjects without a diagnosis of T1DM defined as central laboratory test of C-peptide < 0.7 ng/ml.
3	Re-enrolled randomized subjects who have discontinued the study as a screen or lead-in failure
4	Randomized with a screening visit central laboratory HbA1c < 7.7% or > 11%.
5	Randomized subjects with BMI < 18.5 kg/m2.
6	Randomized subjects with age < 18 or > 75 years of age.
7	Randomized subjects (women) without a negative serum or urine pregnancy test (minimum sensitivity 25 IU/L or equivalent units of HCG) within 24 hours prior to the start of investigational product.
8	Randomized subjects (women) who are breastfeeding.
9	Randomized subjects (women) of child bearing potential who are unwilling or unable to use an acceptable method to avoid pregnancy for the entire study period.
10	Randomized subjects with history of T2DM, maturity onset diabetes of young (MODY), pancreatic surgery or chronic pancreatitis.
11	Randomized subjects with previous use of dapagliflozin and/or any other SGLT-2 inhibitors
12	Randomized subjects with any use of metformin and/or thiazolidinediones within 2 months prior to the Screening visit.
13	Randomized subjects with any use of any non-insulin antihyperglycemic agent within 1 month prior to the screening visit.
14	Randomized subjects with history of diabetic ketoacidosis (DKA) requiring medical intervention (eg, emergency room visit and/or hospitalization) within 1 month prior to the screening visit.
15	Randomized subjects with history of hospital admission for glycemic control (either hyperglycemia or hypoglycemia) within 1 month prior to the screening visit.
16	Randomized subjects with protocol defined frequent episodes of severe hypoglycemia within 1 month prior to the screening visit.
17	Randomized subjects with symptoms of poorly controlled diabetes, including marked polyuria and polydipsia with >10% weight loss during the three months prior to screening, or other signs and symptoms.
18	Randomized subjects with history of Addison’s disease or chronic adrenal insufficiency.
19	Randomized subjects with history of diabetes insipidus.
20	Randomized subjects with myocardial infarction within 6 months of the enrollment visit.

Number	Significant Deviation
21	Randomized subjects with cardiac surgery or revascularization (coronary artery bypass surgery /percutaneous transluminal coronary angioplasty) within 6 months of the enrollment visit.
22	Randomized subjects with Unstable angina within 6 months of the enrollment visit.
23	Randomized subjects with Unstable congestive heart failure (CHF) within 6 months of the enrollment visit.
24	Randomized subjects with CHF New York Heart Association (NYHA) Class III or IV within 6 months of the enrollment visit.
25	Randomized subjects with transient ischemic attack (TIA) or significant cerebrovascular disease within 6 months of the enrollment visit.
26	Randomized subjects with unstable or previously undiagnosed arrhythmia within 6 months of the enrollment visit.
27	Randomized subjects with history of unstable or rapidly progressing renal disease
28	Randomized subjects with history of congenital renal glucosuria
29	Randomized subjects with history of renal allograft
30	Randomized subjects with significant protocol defined hepatic diseases.
31	Randomized subjects with known immunocompromised status.
32	Randomized subjects with donation of blood, blood products to a blood bank, blood transfusion, or participation in a clinical study requiring withdrawal of > 400 mL of blood > 8 wks prior screening
33	Randomized subjects with malignancy within 5 years of the screening visit (with the exception of treated basal cell or treated squamous cell carcinoma)
34	Randomized subjects with history of bladder cancer
35	Randomized subjects with history of radiation therapy to the lower abdomen or pelvis at any time
36	Randomized subjects with aspartate aminotransferase (AST) > 3X upper limit of normal.
37	Randomized subjects with alanine aminotransferase (ALT) > 3X ULN.
38	Randomized subjects with serum total bilirubin (TB) > 2.0x ULN.
39	Randomized subjects with Calculated Creatinine Clearance <60 ml/min/1.73m2.
40	Randomized subjects with hemoglobin ≤ 11.0 g/dL (110 g/L) for men; hemoglobin ≤ 10.0 g/dL (100 g/L) for women.
41	Randomized subjects with positive for hepatitis B surface antigen or anti-hepatitis C virus antibody.
42	Randomized subjects with allergies or contraindication to the contents of dapagliflozin tablets or insulin.
43	Randomized subjects (women) who are pregnant.
44	Randomized subjects who are prisoners or are involuntarily incarcerated.
45	Randomized subjects with subjects who are compulsorily detained for treatment of either a psychiatric or physical (e.g., infectious disease) illness.

Number	Significant Deviation
46	Randomized subjects on a commercial weight loss program with ongoing weight loss, or on an intensive exercise program.
47	Randomized subjects who history of bariatric surgery or lap-band procedure within 12 months prior to screening.
48	Randomized subjects with replacement or chronic systemic corticosteroid therapy, defined as any dose of systemic corticosteroid taken for > 4 weeks within 3 months prior to the Day 1 visit.
49	Randomized subjects with any unstable endocrine, psychiatric or rheumatic disorders as judged by the Investigator.
50	Randomized subjects who are volume depleted or at risk for volume depletion due to co-existing conditions or concomitant medications and cannot carefully monitor their volume status.
51	Randomized subjects who are, in judgment of Investigator, unlikely to comply with protocol, or to correctly self administer MDI and/or insulin pump, or with severe medical or psychological condition.
52	Randomized subjects with any condition which, in the judgment of the Investigator, may render the subject unable to complete the study or which may pose a significant risk to the subject.
53	Randomized subjects who are currently abusing alcohol or other drugs or have done so within the last 6 months prior to the Day 1 visit.
54	Randomized subjects who are a participating investigator, study coordinator, employee of an investigator or immediate family member of any of the aforementioned.
55	Randomized subjects who are employee of BMS, AstraZeneca (AZ), or their relatives.
56	Randomized subjects administrated of any other investigational drug within 30 days of the screening visit.
57	Randomized subjects with clinical conditions, clinically significant abnormalities, after screening and prior to randomization which, in the Investigator's judgment, should preclude treatment.
58	Randomized Subjects that used antihyperglycemic medication (other than protocol allowed medication and study medication) for any duration during short-term double-blind treatment-period.
59	Randomized subjects administered any weight loss medication.
60	Randomized subjects administered acetaminophen (paracetamol) containing medications while using the Dexcom CGM device, and 24 hours prior to insertion.
61	Randomized subjects not discontinued from treatment period despite withdrawal of informed consent.
62	Randomized subjects not discontinued from treatment despite adverse event, laboratory abnormality, intercurrent illness where study continuation is not in the best interest of subject per Investigator
63	Randomized subjects not discontinued from treatment despite unblinding a subject for any reason (emergency or non-emergency)

Number	Significant Deviation
64	Randomized subjects not discontinued from treatment period despite pregnancy.
65	Randomized subjects not discontinued from treatment period despite termination of the study by the Sponsor.
66	Randomized subjects not discontinued from treatment period despite loss of ability to freely provide consent through imprisonment or involuntary incarceration for treatment of illness.
67	Randomized subjects not discontinued from treatment period despite Creatinine Clearance < 30 ml/min.
68	Randomized subjects not discontinued from treatment period despite at least one protocol-defined hypoglycemia episode that leads to a loss of consciousness and/or seizure as determined by Investigator
69	Randomized subjects not discontinued from treatment period despite change of insulin administration method (eg, switch from MDI to CSII or vice versa).
70	Randomized subjects not discontinued from treatment period despite initial and repeat central laboratory tests meet protocol defined criteria for ALT, AST, TB.
71	Randomized subjects administered study medication from an incorrect kit number (dispensing error).
72	Randomized subjects use of any GLP-1 receptor agonist within the protocol specified time frame prior to the screening visit.

The above significant protocol deviations will be listed in a separate report.

9.2 Geographic regions

Geographic Region	Countries/Regions
North America	Canada, United States
Latin America	Argentina, Chile
Europe	Russia, Germany, Bulgaria Poland, Netherland , Sweden, Switzerland United Kingdom
Asia/Pacific	Japan

9.3 Units and marked abnormality criteria for safety laboratory variables

Clinical laboratory variables will be summarized and listed using the units listed here. 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.

Conventional (US) Units:

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		> 1.5X ULN if PreRx ≤ ULN; > 2X 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

Clinical Laboratory variables	Units	Marked Abnormality Criteria	
		Low	High
Magnesium	mEq/L	< 1 mEq/L	> 4 mEq/L
PO4 (Phosphate)	mg/dL	Age 17-65: ≤ 1.8mg/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 Creatinine Ratio)	mg/g		> 1800 mg/g

Standard International Units

Clinical Laboratory variables	Units	Marked Abnormality Criteria	
		Low	High
Hematology			
HCT males/females	vol	< 0.20	> 0.55
HCT males/females	vol		> 0.60
Hemoglobin males/females	g/L	< 60 g/L	> 180 g/L
Hemoglobin males/females	g/L		> 200 g/L
Blood Chemistry			
Albumin	g/L	≤ 20 g/L	> 60 g/dL
Total protein	g/L		> 100 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	µmol/L		> 1.5X ULN if PreRx ≤ ULN; > 2X ULN if PreRx > ULN
Glucose, Plasma Unspecified	mmol/L	< 3 mmol/L	> 19.4 mmol/L
Na (Sodium)	mmol/L	< 130 mmol/L	> 150 mmol/L

Clinical Laboratory variables	Units	Marked Abnormality Criteria	
		Low	High
Na (Sodium)	mmol/L	< 120 mmol/L	
K (Potassium)	mmol/L	≤ 2.5 mmol/L	≥ 6.0 mmol/L
HCO3 (Bicarbonate)	mmol/L	≤ 13 mmol/L	
BUN	mmol/L		≥ 21.4 mmol/L
Creatinine	µmol/L		≥ 1.5X PreRx CREAT
Creatinine	µmol/L		≥ 221 µmol/L
CK (Creatine Kinase)	U/L		> 5X ULN
CK (Creatine Kinase)	U/L		> 10X ULN
Calcium	mmol/L	< 1.88 mmol/L	≥ 0.25 mmol/Lfrom ULN **and** ≥ 0.13 mmol/L from PreRx CA
Magnesium	mmol/L	< 0.5 mmol/L	> 2 mmol/L
Inorganic phosphorus	mmol/L	Age 17-65: ≤ 0.58 mmol/L Age ≥ 66: ≤ 0.68 mmol/L	Age 17-65: ≥ 1.81 mmol/L Age ≥ 66: ≥ 1.65 mmol/L
Urine			
UACR (Urinary Albumin to Creatinine Ratio)	mg/mmol		> 203.62 mg/mmol

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 > 3XULN) and (Bilirubin > 1.5XULN within 14 days on or after AT elevation)
- (AST or ALT > 3XULN) and (Bilirubin > 2XULN within 14 days on or after AT elevation)
- (AST or ALT > 3XULN) and {(Bilirubin > 2XULN and no ALP >= 2XULN) within 14 days on or after AT elevation)}

9.4 Summary of updates

Following is a brief summary of updates compared to MB102 T1D Core Statistical Analysis Plan:

1) The SAP is authorized based on AstraZeneca template and a SAP for study MB102230 only.

2) The method for type I error control has been switched from Dunnett's to a Dunnett and Tamhane step-up procedure. This method applies to primary and secondary endpoints.

3) The randomized subjects data set has been replaced with the Full Analysis Set. The full analysis set will consist of all randomized subjects who took at least one dose of double-blind study medication during the short-term double-blind period.

4) The proportion of subjects with a Pre-defined Characteristic at Week t will be analysed using a logistic regression procedure, adjusting for baseline values and randomization strata, instead of methodology of Zhang, Tsiatis and Davidian and Tsiatis, Davidian, Zhang and Lu.

5) The Section 4.2.3.2 has updated to include the analysis of number of non-severe hypoglycemic events per subject in those achieving a HbA1c reduction of ≥ 0.5% from baseline to the week 24 visit (LOCF), an exploratory endpoint.

6) The definition of insulin up-titration more than 25% than baseline has been updated to include allowable missing total insulin dose days in a 14 consecutive days window (not more than 50% of days and less than 5 consecutive days in a 14 consecutive days window).

7) It has been clarified that primary estimand for primary endpoint is *treatment difference at Week 24 if subjects did not discontinue randomized treatment* and the secondary primary estimand is ITT estimand, *treatment difference at Week 24 regardless of treatment discontinuation.* In both, the values after insulin up-titration will be included in the analysis.

8) The measurements obtained after the first instance the 25% up-titration criteria is met will be excluded in the sensitivity analysis for primary endpoint's primary estimand only. This clause is removed from all of secondary and exploratory efficacy endpoints.

9) Self-monitored blood ketones that are not associated with an event sent for DKA adjudication will be analysed.

10) The sections on PK and PD analysis sets, and PK and PD summaries are removed.

11) Malignant and unspecified tumors is removed from Adverse Events of Special Interest. It is not defined in the protocol.

12) A de facto ITT estimand, i.e. treatment difference at Week 24 regardless of whether treatment was actually taken, has been added.