NCT02973321

STATISTICAL ANALYSIS PLAN

A 26-week Randomized, Double-blind, Placebo-controlled, Dose-ranging Phase 2 Study to Assess the Safety and Efficacy of SAR425899 in Patients with Type 2 Diabetes Mellitus

SAR425899-DRI13940

STATISTICIAN: [redacted]

DATE OF ISSUE: 18-Sep-2017

Total number of pages: 84

Any and all information presented in this document shall be treated as confidential and shall remain the exclusive property of Sanofi (or any of its affiliated companies). The use of such confidential information must be restricted to the recipient for the agreed purpose and must not be disclosed, published or otherwise communicated to any unauthorized persons, for any reason, in any form whatsoever without the prior written consent of Sanofi (or the concerned affiliated company); 'affiliated company' means any corporation, partnership or other entity which at the date of communication or afterwards (i) controls directly or indirectly Sanofi, (ii) is directly or indirectly controlled by Sanofi, with 'control' meaning direct or indirect ownership of more than 50% of the capital stock or the voting rights in such corporation, partnership or other entity

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

AE:	adverse event
AESI:	adverse event of special interest
ALP:	alkaline phosphatase
ALT:	alanine aminotransferase
ANCOVA:	analysis of covariance
APPADL:	ability to perform physical activities of daily living
AST:	aspartate aminotransferases
ATC:	anatomic category
BMI:	body mass index
BNP:	brain natriuretic peptide
CI:	confidence interval
CMH:	Cochran-Mantel-Haenszel
CSR:	clinical study report
DMC:	Data Monitoring Committee
ECG:	electrocardiogram
eCRF:	electronic case report form
FDA:	Food and Drug Administration
FPG:	fasting plasma glucose
FSH:	follicle-stimulating hormone
GGT:	gamma-glutamyl transpeptidase
HbA_{1c}:	glycosylated hemoglobin
HDL:	high-density lipoprotein
HLGT:	high level group term
HLT:	high level term
HR:	heart rate
hs-CRP:	high-sensitivity C-reactive protein
IMP:	investigational medicinal product
IRT:	interactive response technology
IW-SP:	Impact of Weight on self-perception
KM:	Kaplan-Meier
LDL:	low-density lipoprotein
LLN:	lower limit of normal
LLT:	lower level term
MAR:	Missing at random, Missing at random, Missing at random
MedDRA:	Medical Dictionary for Regulatory Activities
NASH:	nonalcoholic steatohepatitis
NIMP:	non-investigational medicinal product
OC:	observed cases
PCSA:	potentially clinically significant abnormality
PD:	pharmacodynamic
PK:	pharmacokinetic

PPG:	post-prandial plasma glucose
PRO:	patient-reported outcome
PSAC:	Pancreatic Safety Assessment Committee
PT:	preferred term
SAE:	serious adverse events
SAP:	statistical analysis plan
SC:	subcutaneous
SD:	standard deviation
SE:	standard error
SMPG:	self-monitored plasma glucose
SOC:	system organ class
TEAE:	treatment-emergent adverse event
ULN:	upper limit of normal
VAS:	visual analogue scale
WHO-DD:	World Health Organization-Drug Dictionary
WRSM:	weight-related symptoms measure

1 OVERVIEW AND INVESTIGATIONAL PLAN

1.1 STUDY DESIGN AND RANDOMIZATION

This is a multicenter, multinational, randomized, placebo-controlled, 7-arm parallel-group Phase 2 study. The study is double-blind with regard to SAR425899 and placebo treatments. It is open-label with regard to liraglutide comparative treatment. The volume of investigational medicinal product (IMP; ie, dose of SAR425899 or matching placebo) is not blinded.

Patients will be randomized into 1 of the 7 treatment arms:

- 3 SAR425899 treatment arms receiving 3 dose levels (60 patients per arm).
- 3 corresponding SAR425899 placebo treatment arms receiving 3 dose levels (10 patients per arm, 30 patients total).
- 1 active comparator arm receiving 1.8 mg liraglutide (60 patients).

For the biostatistical analyses, data from the 3 SAR425899 placebo arms will be pooled into 1 placebo group.

Patients will be stratified by screening visit HbA_{1c} (<8% versus ≥8%) and Visit 4 (Day 1) body mass index (BMI: <35.0 kg/m^2 versus ≥35.0 kg/m^2).

1.2 OBJECTIVES

1.2.1 Primary objectives

The primary objective of this study is to assess the dose-response relationship of SAR425899 versus placebo in terms of glycemic control as measured by the change in glycosylated hemoglobin (HbA_{1c}) from baseline to Week 26.

1.2.2 Secondary objectives

Main secondary objectives:

- To assess the effect of once daily dosing of SAR425899 on body weight over 26 weeks.
- To assess the safety and immunogenicity profile of SAR425899 when administered as daily subcutaneous (SC) injections over 26 weeks, including assessment of the heart rate (HR) change from baseline to Week 26 by electrocardiogram (ECG) and Holter monitor.

Additional secondary objectives:

- To assess the proportion of patients achieving predefined HbA1c targets of <7% and <6.5% as well as the proportion of patients achieving ≥5% and ≥10% body weight loss at Week 26.

- To assess the effect of once daily dosing of SAR425899 on additional parameters of glycemic control and lipid metabolism.
- To assess the effect of once daily dosing of SAR425899 on additional pharmacodynamic (PD) biomarkers.
- To assess the pharmacokinetic (PK) profile and parameters of SAR425899, inter-individual and inter-occasion variability in PK parameters using a population PK approach.

1.2.3 Exploratory objectives

- To assess and compare the safety and efficacy of SAR425899 versus liraglutide from baseline to Week 26.
- To explore SAR425899 PK/PD relationships for glycemic parameters, body weight loss and heart rate.
- To assess the treatment effects in each group on patient-reported outcomes.
- To assess the treatment effects in each group on nonalcoholic steatohepatitis (NASH) and cardiovascular biomarkers.

1.3 DETERMINATION OF SAMPLE SIZE

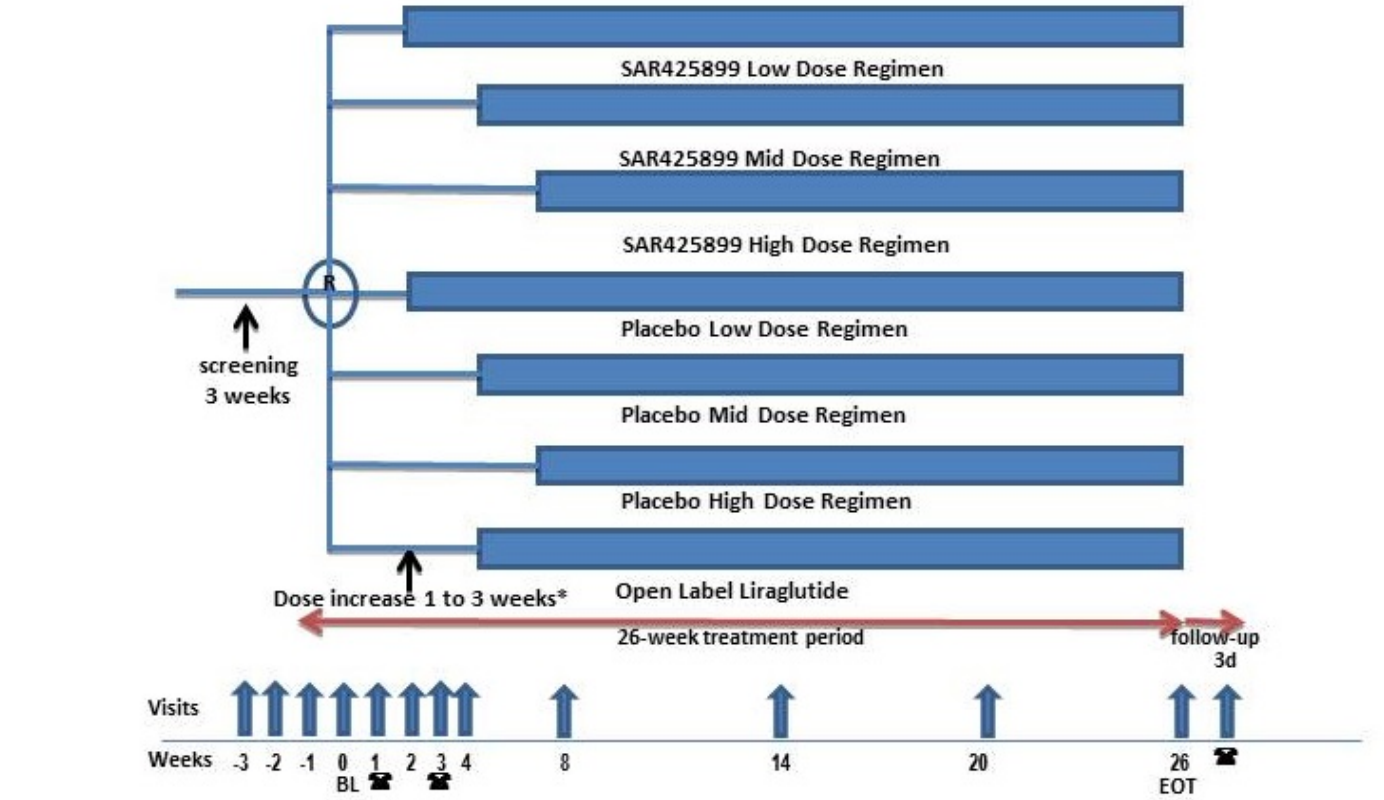
For the primary endpoint of change in HbA_{1c} from baseline to Week 26, a sample size of 60 patients per arm of SAR425899 and 30 patients in the placebo group (pool of the 3 placebo arms) will provide 80% power to detect a difference of 0.7% in the HbA_{1c} change from baseline to Week 26 between a dose group of SAR425899 and placebo (standard deviation (SD) 1.1%; 5% significance level 2-sided).

The total sample size will be 270 patients to be randomized over 7 arms: 3 SAR425899 treatment arms with 3 different dose levels (60 patients per arm), 3 matching SAR425899 placebo treatment arms (10 patients per arm; 30 patients total), and 1 liraglutide arm (active comparator arm with 60 patients).

For the biostatistical analyses, data from the 3 SAR425899 placebo arms will be pooled into 1 placebo group (30 patients total)

Calculations were made using nQuery Advisor® 7.0.

1.4 STUDY PLAN



* depending on the SAR425899/placebo dose regimen
SAR425899 and matching placebo: low dose regimen = 0.06 (1 week) – 0.12 mg (25 weeks)
mid dose regimen = 0.06 (1 week) – 0.12 (1 week) – 0.16 mg (24 weeks)
and high dose regimen = 0.06 (1 week) – 0.12 (1 week) – 0.16 mg (1 week) – 0.20 mg (23 weeks)
Liraglutide: 0.6 mg (1 week) – 1.2 mg (1 week) – 1.8 mg (24 weeks)

1.5 MODIFICATIONS TO THE STATISTICAL SECTION OF THE PROTOCOL

This section summarizes major changes to the protocol statistical section with emphasis on changes after study start (after the first patient was enrolled).

The protocol history table below gives the timing, rationale, and key details of major changes to the protocol statistical section.

The first patient was enrolled on Dec 23rd, 2016 under the original protocol. The amendment 1 was finalized on Jan. 23rd, 2017. The body weight interim analysis is planned on Sept. 2017 and the interim analysis for PK/PD is planned on Dec. 2017. Note that the interim analysis for PK/PD was added on amendment 1.

Table 1 - Protocol amendment statistical changes

Amendment Number	Date Approved	Rationale	Description of statistical changes
1	23-Jan-2017	Increased urinary free-cortisol level was observed in animal studies of SAR425899 but the relevance of this finding to human is unclear. The Food and Drug Administration (FDA) requested to add the measurement of urinary free-cortisol and creatinine from 24h urines at baseline and end of treatment in order to assess the increase in urinary free-cortisol in humans which may suggest a potential effect of SAR425899 on the hypothalamic-pituitaryadrenal axis.	To add urinary free-cortisol and creatinine measured in 24 hours urines at baseline and in Visit 12 (Week 26, End of treatment).
		In phase I studies an increase of ketones bodies was observed for the SAR425899 high dose group (up to 0.18 mg) but not for placebo and the SAR425899 low dose group. In view of the currently requested administration of up to 0.20 mg of SAR425899 and the potential gastrointestinal side effects together with the simultaneous administration of metformin, the BfArM recommended to consider metabolic acidosis as an adverse event of special interest with an immediate notification in order to recognize this state very early in the study and to discontinue the administration of metformin when applicable.	To add metabolic acidosis as a new adverse event of special interest as requested by the BfArM.
		Leptin and microRNA were initially considered as biomarkers of interest in this study. Leptin plasma concentrations are correlated with fat mass. MicroRNA is a surrogate biomarker for nonalcoholic steatohepatitis (NASH). Thus assessment of these two biomarkers was expected to provide information about SAR425899 effect on fat mass and NASH. However the sponsor decided to not measure these parameters anymore during this phase II study. For Leptin, it is not required for any modeling and is not recommended as a surrogate to test for differences in energy expenditure. Nevertheless Leptin and fat mass changes are planned to be assessed in a study dedicated to energy expenditure assessment. For MicroRNA, patients included in this study are not diagnosed for NASH and no liver biopsy is taken which could make the interpretation of microRNA results difficult. Nevertheless microRNA is planned to be assessed in a study dedicated to NASH.	To remove Leptin and microRNA assessments

Amendment Number	Date Approved	Rationale	Description of statistical changes
1	23-Jan-2017	Potentially an interim analysis for population PK and PK/PD will be performed for 60% completers, in order to have data for internal project planning purposes.	To add a potential interim analysis for 60% completers
			To add Amylase/lipase assessment in Visits 9 and 11 as requested by the BfArM.
		To change the reference timepoint for blood sampling during the meal test from the end of the standardized meal to the start of the standardized meal and change the time period for the standardized meal consumption from 15 minutes to 10 minutes.	Correction of a discrepancy between the text in Section 9.2.1.2.2 and Table 2 regarding the reference time for blood sampling during the meal test.
		TSH, T3 and FT4 measurements will be performed at baseline and endpoints visits in order to assess further the thyroid function during the study.	To Add thyroid function assessment at baseline and endpoint visits.

1.6 STATISTICAL MODIFICATIONS MADE IN THE STATISTICAL ANALYSIS PLAN

The statistical analysis plan history table below gives the timing, rationale, and key details for major changes to the statistical analysis features in the statistical analysis plan. Changes also incorporated in a protocol amendment are cross-referenced to Table 1.

Table 2 - Statistical analysis plan statistical changes

SAP version number	Date approved	Rationale	Description of statistical changes
1	This version	Increased urinary free-cortisol level was observed in animal studies of SAR425899 but the relevance of this finding to human is unclear. The FDA requested to add the measurement of urinary free-cortisol and creatinine from 24h urines at baseline and end of treatment in order to assess the increase in urinary free-cortisol in humans which may suggest a potential effect of SAR425899 on the hypothalamic-pituitaryadrenal axis.	Urinary free-cortisol and creatinine measured in 24 hours urines, and the cortisor/creatinine ratio were added to the list of laboratory variables for statistical analyses.
		In phase I studies an increase of ketones bodies was observed for the SAR425899 high dose group (up to 0.18 mg) but not for placebo and the SAR425899 low dose group. In view of the currently requested administration of up to 0.20 mg of SAR425899 and the potential gastrointestinal side effects together with the simultaneous administration of metformin, the BfArM recommended to consider metabolic acidosis as an adverse event of special interest with an immediate notification in order to recognize this state very early in the study and to discontinue the administration of metformin when applicable.	Metabolic acidosis was added to the list of AESI parameters for statistical analyses.
		Leptin and microRNA were initially considered as biomarkers of interest in this study. Leptin plasma concentrations are correlated with fat mass. MicroRNA is a surrogate biomarker for nonalcoholic steatohepatitis (NASH). Thus assessment of these two biomarkers was expected to provide information about SAR425899 effect on fat mass and NASH. However the sponsor decided to not measure these parameters anymore during this phase II study. For Leptin, it is not required for any modeling and is not recommended as a surrogate to test for differences in energy expenditure. Nevertheless Leptin and fat mass changes are planned to be assessed in a study dedicated to energy expenditure assessment. For MicroRNA, patients included in this study are not diagnosed for NASH and no liver biopsy is taken which could make the interpretation of microRNA results difficult. Nevertheless microRNA is planned to be assessed in a study dedicated to NASH.	Leptin and microRNA were removed from the list of pharmacodynamics biomarkers for statistical analyses.
		Potentially an interim analysis for population PK and PK/PD will be performed for 60% completers, in order to have data for internal project planning purposes.	An interim PK/PD analysis will be performed for 60% completers.
		To add Amylase/lipase assessment in Visits 9 and 11 as requested by the BfArM.	Visit 9 and 11 amylase/lipase assessments will be included in the analyses of amylase/lipase.
		To change the reference timepoint for blood sampling during the meal test from the end of the standardized meal to the start of the standardized meal and change the time period for the standardized meal consumption from 15 minutes to 10 minutes.	There will be no change to the planned by timepoint analyses on meal test related efficacy parameters due to the change, however, sensitivity analyses may be performed as necessary.

(electronic 1.0)

SAP version number	Date approved	Rationale	Description of statistical changes
1	This version	To be consistent with company standard.	The definition of treatment compliance was changed, with above-planned dosing defined as any day that the patient took a higher dose than planned, and under-planned dosing defined as any day that the patient took a lower dose than planned.
1	This version	Trend tests on HbA1c will be performed instead pairwise comparison of each dose of SAR425899 versus placebo for primary analysis. Same changes are also made for body weight analyses. The hierarchical testing procedure is also updated to reflect the changes.	Trend tests on HbA1c will be performed instead pairwise comparison of each dose of SAR425899 versus placebo for primary analysis. Same changes are also made for body weight analyses. The hierarchical testing procedure is also updated to reflect the changes. Details of the trend tests and hierarchical testing procedure are presented in Section 2.4.4.
1	This version	Unblinded analyses results on body weight, HbA1c and FPG, and unblinded DMC results (including demographics, disposition, and safety summary) and additional related exploratory analyses will/may be provided to Sanofi internal project planning personnel to support compound development during the 2 interim analyses of the study.	Unblinded analyses results on body weight, HbA1c and FPG, and unblinded DMC results (including demographics, disposition, and safety summary) and additional related exploratory analyses will/may be provided to Sanofi internal project planning personnel to support compound development during the 2 interim analyses of the study. Details are presented in Section 3.
1	This version	To account for the situation that a patient took a SAR425899/placebo dose level that is different from the planned 4 doses levels allow for the study, specify that if a patient is exposed to more than one dose level of SAR425899 since Week 8 visit or later, then the patient will be analyzed in the dose level administered on Week 8 visit or study Day 56 if Week 8 visit is not available.	Patients will be analyzed for safety analyses according to the treatment actually received. If a patient is exposed to more than one dose level of SAR425899 since Week 8 visit or later, then the patient will be analyzed in the dose level administered on Week 8 visit or study Day 56 if Week 8 visit is not available.

2 STATISTICAL AND ANALYTICAL PROCEDURES

2.1 ANALYSIS ENDPOINTS

2.1.1 Demographic and baseline characteristics

The baseline value is defined as the last available value before the first injection of investigational medicinal product (IMP) or the last available value on or before the date of randomization if not treated with IMP.

Baseline safety and efficacy parameters are presented along with the summary statistics for safety and efficacy sections (Section 2.4.5 and Section 2.4.4).

Demographic characteristics

- Demographic characteristics to be summarized are:
- Age (years) derived as: (Date of informed consent - Date of birth)/365.25;
- Age categories (<50, ≥50 to <65, ≥65 to <75, ≥75 years of age);
- Gender (Male, Female);
- Race (American Indian or Alaska Native, Asian/Oriental, Black, Native Hawaiian or Other Pacific Islander, White, Other);
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino, Unknown);
- Screening (V1) HbA1c (%);
- Randomization strata (V1 screening) of HbA1c (<8%, ≥8%);
- Baseline BMI (kg/m^2) derived as: (Weight in kg)/(Height in meters)^2;
- Randomization strata of baseline BMI level (<35 kg/m^2, ≥35 kg/m^2);
- Country.

Disease characteristics at screening or baseline

Disease history includes:

- Duration of diabetes (years) derived as: (Date of informed consent – Date of diagnosis of diabetes + 1)/365.25;
- Categorized duration of diabetes (years) (<10 years, ≥10 years)
- Age at onset of diabetes (years) derived as: (Date of diagnosis of diabetes – Date of birth + 1)/365.25;
- Anti-Hyperglycemic Therapy Categories (Metformin only, Diet/Exercise only, Metformin and diet/exercise)
- Duration of metformin treatment (years) derived as: (Date of informed consent – Date of first dose of metformin + 1)/365.25;

(electronic 1.0)

- Daily dose of metformin at baseline (mg) for patients taking metformin, defined as the last total daily value of metformin intake before the first dose injection of IP for those taking metformin at baseline;
- Categorized daily dose of metformin at baseline (<1500, ≥1500 to <2500, ≥2500 to <3000, ≥3000 mg);
- Baseline diabetic microvascular complications (Yes, No) (ie, diabetic retinopathy, diabetic neuropathy, and diabetic nephropathy);
- History of gestational diabetes (yes or no);
- Baseline urine albumin/creatinine ratio categories (<30 µg/mg [Normal], ≥30 to <300 µg/mg [Microalbuminuria], and ≥300 µg/mg [Macroalbuminuria]);
- Estimated Glomerular Filtration Rate (eGFR) at screening (ml/min);
- eGFR categories at screening (<15 mL/min [End stage renal disease], ≥15 to <30 mL/min [Severe decrease in GFR], ≥30 to <60 mL/min [Moderate decrease in GFR], ≥60 to <90 mL/min [Mild decrease in GFR], and ≥90 mL/min [Normal]).

Medical or surgical history

Medical history and medical findings include:

- Physical examination;
- Medical or surgical history;
- Alcohol habits;
- Tobaco smoking habits.

Medical and surgical history will be coded to a “lower level term (LLT)”, “preferred term (PT)”, “high level term (HLT)”, “high level group term (HLGT)”, and associated primary “system organ class (SOC)” using the version of Medical Dictionary for Regulatory Activities (MedDRA) currently in effect at Sanofi at the time of database lock.

2.1.2 Prior or concomitant medications

All medications taken within 3 months prior to the screening visit or during study are to be reported in the electronic case report form (eCRF) pages.

All medications will be coded using the version of World Health Organization-Drug Dictionary (WHO-DD) currently in effect at Sanofi at the time of database lock.

Medications will be classified into the following 3 groups:

- Prior medications are those the patient took prior to the first injection of study IMP. Prior medications can be discontinued before first IMP injection or can be ongoing during treatment period.
- Concomitant medications are those the patient continued or started on or after the first injection of study IMP up to 3 days after the last injection of study IMP.
- Post-treatment medications are those the patient continued or started on or after 4 days after the last injection of study IMP.

A given medication can be classified in several groups.

Any technical details related to computation, dates, imputation for missing dates are described in Section 2.5.

2.1.2.1 Concomitant diabetes therapy

Patients can be enrolled with a background therapy metformin. Metformin should be administered according to its locally approved label and should be kept at stable dose throughout the study (same dose as used prior to the study) unless there is a specific safety issue related to this medication.

2.1.2.2 Rescue therapy

In case of fasting plasma glucose (FPG) or HbA1c above pre-defined thresholds and if no reasons can be found for insufficient glucose control, it is recommended to add rescue therapy.

Central laboratory alerts on FPG and HbA1c are set up to ensure that glycemic parameters remain under predefined threshold values (see hereinafter).

The threshold values are defined as follows, depending on study period:

From baseline Visit 4 (Week 0) to Visit 9 (Week 8, including value at Visit 9): FPG >270 mg/dL (15.0 mmol/L).

From Visit 9 (Week 8) to Visit 10 (Week 14): FPG > 240 mg/dL (13.3 mmol/L).

From Visit 10 (Week 14) to Visit 12 (Week 26): FPG>200 mg/dL (11.1 mmol/L) or HbA1c >8%.

In case of FPG / HbA1c above the threshold values, the investigator should ensure that no reasonable explanation exists for insufficient glucose control, and in particular that:

- Plasma glucose was actually measured in fasting condition (ie, after at least 8 hrs fast).
- Investigational product is given at planned dose.
- There is no intercurrent disease, which may jeopardize glycemic control (eg, infectious disease).
- Compliance with treatment is appropriate.
- Compliance with diet and lifestyle is appropriate.

If any of the above can reasonably explain the insufficient glycemic control, the investigator should undertake appropriate action, ie,

- Dose increase of investigational product according to protocol (if tolerance allows).
- Set up adequate investigation and treatment of intercurrent disease (to be reported in AE/concomitant medication parts of the eCRF).
- Stress on the absolute need to be compliant with treatment.

- Organize a specific interview with a Registered Dietician or other qualified nutrition professional and stress on the absolute need to be compliant to diet and lifestyle recommendations, and check a FPG / HbA1c assessment at the next visit.

If none from the above-mentioned reasons can be found, or if appropriate action fails to decrease FPG / HbA1c under the threshold values, rescue medication may be introduced.

The choice of rescue therapy is at the investigator's discretion with the exception of using GLP-1 receptor agonists or DPP4 inhibitors.

2.1.2.3 Prohibited concomitant therapy

Use of the following medications will not be permitted during the study:

- Weight control treatment, including any medication with a labelled reference to weight loss or weight gain.
- Systemic glucocorticoid use for more than 7 days.
- Any antidiabetic treatment other than IMP, authorized background antidiabetic therapy and rescue therapy, if necessary.
- GLP-1 receptor agonists or dipeptidyl peptidase 4 (DPP4) inhibitors.
- Drugs that affect gastrointestinal motility (eg, chronic use of anticholinergics, antispasmodics, 5HT3 antagonists, opiates).
- Medications that may cause significant weight gain such as antipsychotic medications, tricyclic antidepressants, and mood stabilizers (eg, imipramine, amitryptiline, mirtazapine, paroxetine, phenelzine, chlorpromazine, thioridazine, clozapine, olanzapine, valproic acid and its derivatives, and lithium), and growth hormone.
- Cannabis, even if prescribed for medical reasons.

Any technical details related to computation, dates, imputation for missing dates are described in Section 2.5.

2.1.3 Efficacy endpoints

All biological efficacy assessments will be performed by a central laboratory. All scheduled measurements collected during the study will be used in the analysis, including those obtained after IMP discontinuation or rescue medication use.

Efficacy variables will be summarized in both standard international units and conventional units when applicable.

The baseline value for efficacy endpoints is the last available value prior to the first injection of IMP or the last available value on or before the date of randomization if not treated with IMP.

HbA1c is assayed at screening (Visit 1); at Visit 4 (Day 1) before first IP injection; at Visit 6 (Week 2); at Visit 8 (Week 4); at Visit 9 (Week 8); at Visit 10 (Week 14); at Visit 11 (Week 20); and at final assessment (Visit 12 at Week 26, primary endpoint assessment visit).

Body weight is measured at screening (Visit 1); at Visit 4 (Day 1) before first IP injection; at Visit 6 (Week 2); at Visit 8 (Week 4); at Visit 9 (Week 8); at Visit 10 (Week 14); at Visit 11 (Week 20); and at final assessment (Visit 12 at Week 26, primary endpoint assessment visit).

FPG is measured at a central laboratory at screening (Visit 1); at Visit 4 (Day 1) before first IP injection; at Visit 6 (Week 2); at Visit 8 (Week 4); at Visit 9 (Week 8); at Visit 10 (Week 14); at Visit 11 (Week 20); and at final assessment (Visit 12 at Week 26, primary endpoint assessment visit).

The self-monitored plasma glucose (SMPG) is measured over a single 24-hr period in the week before Visit 4 (Day 1) before first IP injection, Visit 10 (Week 14), and Visit 12 (Week 26, primary endpoint assessment visit).

Fasting insulin are measured at Visit 4 (Day 1) before first IP injection; at Visit 10 (Week 14); and at final assessment (Visit 12 at Week 26, primary endpoint assessment visit).

Following list of standardized meal test related measurements at premeal, 10 min, 20 min, 0.5 h, 1 h, 1.5 h, 2 h, 3 h, 4 h post meal at Visit 3 (Week -1); and at final assessment (Visit 12 at Week 26, primary endpoint assessment visit) in up to 50% subset of all patients:

- post-prandial plasma glucose (PPG) and associated glucose excursion
- postprandial insulin, proinsulin and C-peptide and associated excursion values

Pre-meal measurements of pro-insulin and C-peptide from the meal test will be used as fasting pro-insulin, and C-peptide values.

The variables HOMA-IR (homeostasis model assessment for insulin resistance) index and

HOMA-IR (homeostasis model assessment for β-cell function) index, derived from fasting

plasma glucose and fasting plasma insulin (FPI), are calculated as:

HOMA-IR index = [FPI (μU/mL) × FPG (mmol/L) / 22.5],

HOMA-β index = ([20 × FPI (μU/mL)] / [FPG (mmol/L) - 3.5]).

HOMA-IR and HOMA-IR will be derived at Visit 4 (Day 1) before first IP injection; at Visit 10 (Week 14); and at final assessment (Visit 12 at Week 26, primary endpoint assessment visit).

Fasting lipid profile (total cholesterol, triglycerides, LDL-C, and HDL-C) and free fatty acids will be assessed at screening, Visit 4 (Day 1) before first IP injection; at Visit 10 (Week 14); and at final assessment (Visit 12 at Week 26, primary endpoint assessment visit).

Ketone bodies is assessed at Visit 4 (Day 1) before first IP injection; at Visit 8 (Week 4); at Visit 9 (Week 8); at Visit 10 (Week 14); at Visit 11 (Week 20); and at final assessment (Visit 12 at Week 26, primary endpoint assessment visit).

2.1.3.1 Primary efficacy endpoint(s)

The primary efficacy endpoint is change in HbA1c (%) from baseline to Week 26.

2.1.3.2 Secondary efficacy endpoint(s)

Continuous secondary efficacy endpoints

The continuous secondary efficacy endpoints are:

- Change in body weight from baseline to Week 26,
- Change in FPG from baseline to Week 26,
- Change in 7-point SMPG profiles from baseline to Week 26 (each time point and average daily value),
- Change in PPG and in blood glucose excursion in response to a standardized meal test in up to 50% subset of all patients from baseline to Week 26 at time points premeal, 10 min, 20 min, 0.5 h, 1 h, 1.5 h, 2 h, 3 h, 4 h post meal,
- Change in fasting insulin, proinsulin, and C-peptide from baseline to Week 26,
- Change in insulin, proinsulin, and C-peptide and associated excursion values in response to a standardized meal test in up to 50% subset of all patients from baseline to Week 26 at time points premeal, 10 min, 20 min, 0.5 h, 1 h, 1.5 h, 2 h, 3 h, 4 h post meal,
- Change in HOMA-β and HOMA-IR from baseline to Week 26,
- Change in lipid profile (total cholesterol, triglycerides, LDL-C, and HDL-C) and free fatty acids, and Ketone bodies from baseline to Week 26,

For 7-point SMPG profiles, the mean daily (ie, average on 7 points) change from baseline to Week 26 will be analyzed. In addition, the change from baseline to Week 26 for each of the 7 points will be evaluated, respectively.

For FPG, only values assessed in fasting condition will be analyzed.

Categorical secondary efficacy endpoints

The categorical secondary efficacy endpoints are:

- Percentage of patients reaching HbA1c <6.5% (49 mmol/mol) at Week 26,
- Percentage of patients reaching HbA1c <7% (53 mmol/mol) at Week 26.
- Percentage of patients achieving ≥5% body weight loss at Week 26,
- Percentage of patients achieving ≥10% body weight loss at Week 26,
- Percentage of patients requiring rescue therapy during the 26-week randomized treatment period.

An algorithm for defining patients requiring rescue therapy is provided in Section 2.5.8.

For further details on missing data handling, see Section 2.5.3.

2.1.4 Safety endpoints

The safety endpoints are assessed by:

- Adverse events (AEs), serious adverse events (SAEs), AE of special interest (AESIs), hypoglycemia (severe, documented, asymptomatic, probable, relative), vital signs, ECG changes, and safety laboratory values,
- Change of heart rate from baseline to Week 26 by ECG,
- Change of heart rate from baseline to Week 26 by Holter monitor in up to 50% subset of all patients,
- Antibody assessments: anti-SAR425899 antibodies samples will be collected in all patients (except patients on Liraglutide) at baseline, Week 2, Week 4, Week 14, and Week 26 before dosing. Patients will be asked to provide sample for anti-SAR425899 antibodies assessments 3 months after the end of the study if needed.

Observation period

The observation period will be divided into 4 epochs:

- The **screening** epoch is defined as the time from the signed informed consent date up to the first administration of the IMP.
- The **treatment** epoch is defined as the time from the first administration of the IMP to the last administration of the IMP
- The **residual treatment** epoch is defined as the time from the last administration of the IMP up to 3 days (1 day for hypoglycemia, 28 days for anti-SAR425899 antibody) after the last administration of IMP.

  The treatment-emergent adverse event (TEAE) period will include both **treatment** and **residual treatment** epochs.
- The **posttreatment** epoch is defined as the period of time starting the day after the end of the treatment-emergent adverse event period up to the end of the study (defined as last protocol-planned visit or the resolution/stabilization of all serious adverse events (SAE) and adverse events with prespecified monitoring as defined in protocol).

The on-study observation period is defined as the time from the start of treatment until the end of study (defined as the last scheduled visit for those who completed the study and the date collected on e-CRF page “Completion of End of Study/Follow-up” for those who did not complete the study, or the resolution/stabilization of all serious adverse events and adverse events with prespecified monitoring as defined in the protocol whichever is the latest).

(electronic 1.0)

2.1.4.1 Hypoglycemia

Hypoglycemia will be identified as events recorded on the dedicated eCRF or eDiary "Hypoglycemic Event Information" page, and will be categorized as follows (see study protocol for further details):

Severe hypoglycemia

Severe hypoglycemia is an event requiring assistance of another person to actively administer carbohydrate, glucagon, or other resuscitative actions. These episodes may be associated with sufficient neuroglycopenia to induce seizure, unconsciousness, 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.

The definition of severe hypoglycemia includes all episodes in which neurological impairment was severe enough to prevent self-treatment and which were thus thought to place patients at risk for injury to themselves or others.

Note that "requires assistance" means that the patient could not help himself or herself. Assisting a patient out of kindness, when assistance is not required, should not be considered a "requires assistance" incident.

Severe hypoglycemia will be qualified as an SAE only if it fulfills SAE criteria. All events of seizure, unconsciousness, or coma must be reported as SAEs.

In the eCRF, severe hypoglycemia is identified based on information captured in the "Symptomatic hypoglycemic event information" page as those

- ticked "Subject was Not Capable of Treating Self and Required Assistance" to the question "Countermeasure Administration" and
- ticked "Yes" to the question "Were Symptoms Present".

Documented symptomatic hypoglycemia

Documented symptomatic hypoglycemia is an event during which typical symptoms of hypoglycemia are accompanied by a measured plasma glucose concentration of ≤70 mg/dL (3.9 mmol/L). Clinical symptoms that are considered to result from a hypoglycemic episode can include (but are not limited to): increased sweating, nervousness, asthenia, tremor, dizziness, increased appetite, palpitations, headache, sleep disorder, confusion, seizures, unconsciousness, and coma.

In the eCRF, documented symptomatic hypoglycemia is identified based on information captured in the "Symptomatic hypoglycemic event information" page as those

- not ticked "Subject was Not Capable of Treating Self and Required Assistance" to the question "Countermeasure Administration" and
- ticked "Yes" to the question "Were Symptoms Present" and
- with a measured plasma glucose value before countermeasure ≤70 mg/dL (3.9 mmol/L).

Asymptomatic hypoglycemia

Asymptomatic hypoglycemia: Asymptomatic hypoglycemia is an event not accompanied by typical symptoms of hypoglycemia but with a measured plasma glucose concentration less than or equal to 70 mg/dL (3.9 mmol/L).

- not ticked "Subject was Not Capable of Treating Self and Required Assistance" to the question "Countermeasure Administration" and
- ticked "No" to the question "Were Symptoms Present" and
- with a measured plasma glucose value ≤70 mg/dL (3.9 mmol/L).

Probable symptomatic hypoglycemia

Probable symptomatic hypoglycemia is an event during which symptoms of hypoglycemia are not accompanied by a plasma glucose determination, but was presumably caused by a plasma glucose concentration less than or equal to 70 mg/dL (3.9 mmol/L); symptoms treated with oral carbohydrate without a test of plasma glucose.

In the eCRF, probable symptomatic hypoglycemia is identified based on information captured in the "Symptomatic hypoglycemic event information" page as those

- not ticked "Subject was Not Capable of Treating Self and Required Assistance" to the question "Countermeasure Administration" and
- ticked "Yes" to the question "Were Symptoms Present" and
- ticked "Yes" to the question "Was Any Countermeasure Given for the Hypoglycemic Event?" and
- ticked "Yes" to the question "Did this countermeasure lead a significant improvement or prompt recovery?" and
- with no plasma glucose value before countermeasure.

Relative hypoglycemia

Relative hypoglycemia: Relative hypoglycemia recently termed "pseudo-hypoglycemia" is an event during which the person with diabetes reports any of the typical symptoms of hypoglycemia, and interprets the symptoms as indicative of hypoglycemia, but with a measured plasma glucose concentration greater than 70 mg/dL (3.9 mmol/L), but approaching this level.

- not ticked "Subject was Not Capable of Treating Self and Required Assistance" to the question "Countermeasure Administration" and
- ticked "Yes" to the question "Were Symptoms Present"
- with a measured plasma glucose value before countermeasure >70 mg/dL (3.9 mmol/L),

Symptomatic hypoglycemia events fulfilling the criteria of a SAE will also be recorded on AE and SAE forms in eCRF.

2.1.4.2 Adverse events variables

Adverse event observation period

- Pre-treatment adverse events are AEs that developed or worsened or became serious during the pretreatment period.
- Treatment-emergent adverse events are AEs that developed or worsened (according to the Investigator's opinion) or became serious during the TEAE period.
- Post-treatment adverse events are AEs that developed or worsened or became serious during the post-treatment period.

All AEs (including SAEs and AEs with prespecified monitoring) will be coded to a lower-level term (LLT), preferred term (PT), high-level term (HLT), high-level group term (HLGT), and associated primary system organ class (SOC) using the version of Medical Dictionary for Regulatory Activities (MedDRA) currently in effect at Sanofi at the time of database lock.

Record the occurrence of AEs (including SAEs and AEs with prespecified monitoring) from the time of signed informed consent until the end of the study.

AESI include:

- Pregnancy
- Symptomatic overdose
- Alanine aminotransferase (ALT) increase >3X upper limit of normal (ULN) (per related decision chart in Appendix C of protocol)
- prolongation of QTc interval ≥500 ms (automatic measurement and confirmed by a manual reading by the Investigator or a physician delegated by the Investigator using the Fridericia formula for correcting QT)
- sinus tachycardia (defined as confirmed sinus tachycardia [heart rate >100 bpm] in supine position reported at more than 2 visits) and associated with a concomitant increase, from baseline, in heart rate of ≥15 beats per minute
- Severe allergic or allergic-like reaction (except local reactions at site of injection)
- Calcitonin ≥20 pg/mL: events recorded on the AE form for increased calcitonin and its associated complementary form.
- Amylase or lipase >2 x ULN
- Metabolic acidosis

Note that all AESI are identified by the corresponding AECAT variable in the AE raw dataset and/or other specific form as listed in the table below.

Table 3 - Summary of AESI reporting instructions

Event category	Reporting timeframe	Specific events in this category	Case Report Form completion		
			AE form	Safety Complementary Form	Other specific forms
Adverse Event of Special Interest	Expedited (within 24 hours)	Pregnancy	Yes	Yes	Yes
		Symptomatic overdose	Yes	Yes	Yes
		Increase in alanine transaminase (ALT)	Yes	Yes	Yes
		Prolongation of QTc interval (automatic measurement) ≥500 ms	Yes	Yes	No
		Persistence of sinus tachycardia per definition in Section 2.1.4.2 and associated with a concomitant increase, from baseline, in heart rate of ≥15 beats per minute	Yes	Yes	No
		Severe allergic or allergic-like reaction (except local reactions at site of injection)	Yes	Yes	Yes
		Calcitonin ≥20 pg/mL	Yes	Yes	Yes
		Amylase/lipase>2xULN	Yes	Yes	Yes
		Metabolic acidosis	Yes	Yes	No

Other AEs to be separately analyzed include:

- Major cardiovascular events: events adjudicated as major cardiovascular events by CAC:
  - Death
  - MI/UA requiring hospitalization
  - Congestive heart failure requiring hospitalization
  - Coronary Revasularization
  - Arrhythmia
  - Stroke

Adverse events sent for adjudication for major cardiovascular events will be identified by the corresponding AECAT variable in the AE raw dataset and/or other specific form and/or adjudication form.

- Pancreatic events:
  - Pancreatitis adjudicated by the Pancreatic Safety Assessment Committee (PSAC) as: 1) acute pancreatitis, 2) acute on chronic pancreatitis, 3) chronic pancreatitis, 4) unknown pancreatitis.
  - Pancreatic neoplasms adjudicated by the PSAC

Adverse events sent for adjudication for pancretitis and/or Pancreatic neoplasms will be identified by the corresponding AECAT variable in the AE raw dataset and/or other specific form and/or adjudication form.

- Local tolerability at injection site: identified by searching the term "injection site" in the PTs coded from the investigator reported terms.

2.1.4.3 Deaths

The deaths observation period are per the observation periods defined above.

- Death on-study: deaths occurring during the on-study observation period
- Death on-treatment: deaths occurring during the TEAE period
- Death post-study: deaths occurring after the end of the study

2.1.4.4 Laboratory safety variables

Clinical laboratory data consists of blood analysis (including hematology and clinical chemistry, amylase, lipase, calcitonin) and urinalysis (albumin/creatinine ratio). Clinical laboratory values will be summarized in both standard international units and conventional units.

The laboratory data will be collected at designated visits (see study flowchart in Appendix C. The following laboratory data will be measured at a central laboratory and used as safety endpoints:

- Hematology
  - Red blood cells, platelets: hemoglobin, hematocrit, red blood cells (erythrocytes) count, platelets count.
  - White blood cells: white blood cells count, and differential counts (neutrophils, lymphocytes, monocytes, basophils, and eosinophils).
- Clinical chemistry
  - Electrolytes: sodium, potassium, calcium, phosphorus.
  - Renal function: creatinine, calculated eGFR, cystatin C, uric acid.
  - Liver function: alanine aminotransferases (ALT), aspartate aminotransferases (AST), alkaline phosphatase (ALP), gamma-glutamyl transpeptidase (GGT), total bilirubin (and, in case of values above the normal range, differentiation in conjugated and non-conjugated bilirubin).
- Thyroid panel: Thyroid-stimulating hormone (TSH), Total Triiodothyronine (T3) and Free Thyroxine (FT4).
- Pancreatic enzymes: lipase, amylase.
- Serum calcitonin.
- Urine analysis: albumin/creatinine ratio.
- Urinary free-cortisol and creatinine based on 24-hour urine sample in some patients, 24-hr free-cortisol/creatinine ratio will be derived for these patients.

Lipid parameters (fasting): total cholesterol, triglycerides, LDL-C, and HDL-C are presented as efficacy parameters in Section 2.1.3.2.

Other laboratory data not mentioned above (serology tests (hepatitis B antigen, hepatitis C antibodies), serum pregnancy test, serum follicle-stimulating hormone [FSH], and estradiol) will not be defined as safety endpoints (collected as per study flowchart in Appendix C).

Technical formulas are described in Section 2.5.1.

2.1.4.5 Vital signs variables

Vital signs include supine heart rate (bpm) and sitting systolic and diastolic blood pressures (mmHg).

2.1.4.6 Physical examination

A exam will be performed at Visit 1 (Screening), Visit 4 (baseline), Visit 9 (Week 8), Visit 10 (Week 14), and at final assessment (Visit 12 at Week 26, primary endpoint assessment visit). "Normal", "Abnormal" or "Not done" as determined by the Investigator will be reported in the eCRF. Any abnormality related to physical examination will be reported in the eCRF "Adverse Event" page or "Medical History" page.

2.1.4.7 Electrocardiogram variables

Standard 12-lead ECGs are recorded at Week -3, 0 (Day 1), 4, Visit 10 (Week 14), and on Visit 12 (Week 26, primary endpoint assessment visit). A single 12-lead ECG printout with heart rate, PR, QRS, QT, QTc automatic correction evaluation (by the ECG device) will be generated. ECG status of "normal" or "abnormal" will be reported in the eCRF as determined by the Investigator.

2.1.4.8 Holter Monitoring

Holter ECG will be collected on Visit 3 (Week -1), on Visit 10 (Week 14), and on Visit 12 (Week 26, primary endpoint assessment visit). The Holter monitor will be proposed to first up to 50% of patient who agree. Twenty–four hour average and night time average heart rate will be derived based on the holter monitoring. Twenty-four hour interval is defined to be from the time of dosing until 24 hours post dose. Night time is defined to be the interval from 12 hours post dose until 24 hours post dose. For Visit 3, time of dosing is expected time of dosing on Day 1.

2.1.4.9 Anti-SAR425899 antibody assessment

Anti-SAR425899 antibody status (Positive, Negative), antibody levels (titer or conecentration) will be collected on the safety population and only in patients from SAR treatment groups.

The TEAE period for antibody variables is defined in Section 2.1.4.

2.1.5 Pharmacokinetic variables

Pharmacokinetics variables include plasma concentrations of patients in the SAR425899 groups.

2.1.6 Pharmacodynamic biomarkers, NASH and cardiovascular biomarkers

Pharmacodynamic biomarkers are:

- Change waist circumference, hip circumference, waist to hip ratio from baseline to Week 26,
- Change in visual analogue scale (VAS) for appetite/satiety score in response to a standardized meal test in up to 50% subset of all patients from baseline to Week 26 at time points premeal, 4 hours post meal,
- Change in fasting level of adiponectin from baseline to Week 26.

Waist circumference, hip circumference, waist to hip ratio are assessed at Visit 4 (Day 1) before first IP injection; at Visit 10 (Week 14); and at final assessment (Visit 12 at Week 26, primary endpoint assessment visit).

VAS for appetite/satiety score will be measured at premeal and 4 hours post meal at Visit 3 (Week -1); and at final assessment (Visit 12 at Week 26, primary endpoint assessment visit) in up to 50% subset of all patients

Fasting level of adiponectin will be assessed at Visit 4 (Day 1) before first IP injection; at Visit 10 (Week 14); and at final assessment (Visit 12 at Week 26, primary endpoint assessment visit).

2.1.7 NASH and cardiovascular biomarkers

- Change in NASH biomarkers from baseline to Week 26: α2-macroglobulin, apolipoprotein A1, haptoglobin, CK18, GGT, ALT and AST.
- Change in Cardiovascular biomarkers from baseline to Week 26: high-sensitivity C-reactive protein (hs-CRP), brain natriuretic peptide (BNP).

2.1.8 Patient reported outcomes

The patient-reported outcome (PRO) questionnaires in this study are the Weight-Related Symptoms Measure (WRSM), the Ability to Perform Physical Activities of Daily Living (APPADL, initially called the Impact of Weight on Activities of Daily Living) and the Impact of Weight on self-perception (IW-SP). The questionnaires are attached in the Appendix E.

The 3 PRO questionnaires will be administered at baseline and Week 26. The patients will be requested to complete the questionnaires by themselves during selected clinical visits in specific booklets, independently from investigator, site staff and any help from friends or relatives. For validity purposes, patients will be asked to answer all the questions of the questionnaires at the start of the visit in a quiet place, before any procedures/tests and before discussing with the investigator, and while on site to return the completed questionnaires to the investigator or his/her designee on the same day.

(electronic 1.0)

In case of early IMP discontinuation, the questionnaires will be administered at the visit planned for the last dosing day with IMP and afterwards as normally planned.

The WRSM is a validated PRO questionnaire to measure the presence and bothersomeness of obesity symptoms (1). It was developed with and for obese and overweight adults. It includes 20 items on shortness of breath, tiredness, sleep problems, sensitivity to cold, increased thirst, increased irritability, back pain, frequent urination, pain in the joints, water retention, foot problems, sensitivity to heat, snoring, increased appetite, leakage of urine, lightheadedness, increased sweating, loss of sexual desire, decreased physical stamina, and skin irritation. The presence of each symptom is first asked to patient (Yes/No) and then the bothersomeness of symptom is evaluated on a 7-point Likert scale (from 0 [not at all] to 6 [a very great deal bothersomeness]). A total score is calculated by summing the bothersomeness for each symptom, ranging from 0 to 120, with lower scores corresponding to less bothersomeness. The completion time of the WRSM is less than 5 minutes. It demonstrated robust psychometric properties in obese and overweight people including its responsiveness to weight decrease.

APPADL:

The APPADL, previously named the Impact of Weight on Activities of Daily Living (IW-ADL) is a validated PRO questionnaire to measure the ability of moderately obese individuals with T2DM to perform daily physical activities (2). It was developed with T2DM individuals with BMI between 25 and 40 kg/m². It includes 7 items on flexibility, mobility, and activity level, evaluated on a 5-point Likert scale (from 1 [unable to do] to 5 [not at all difficult]). A total score is calculated by summing the 7 items and dividing by the number of items, it ranges from 1 to 5, with higher scores corresponding to greater ability to do physical daily activities. The completion time of the APPADL is estimated to be less than 5 minutes. It demonstrated robust psychometric properties in patients with T2DM and BMI >30 kg/m² and was able to discriminate patients who achieved at least 5% weight loss from others who did not (3).

A raw total APPADL score is calculated by summing the 7 item scores and dividing by the number of items (7). The raw APPADL total score ranges from 1 to 5. To transform the APPADL total score to a scale ranging from 0 to 100, the following transformation formula is used:

[(actual raw score – lowest possible score)/raw score range (highest possible score – lowest possiblescore)] x 100.

For example, using the transformation formula, a raw APPADL total score of “3.0” becomes a transformed total score of “50.0” on a scale of 0 to 100:

$$[(3.0-1.0)/(5-1)] *100 = 2/4 *100 = 50.0$$

Higher raw APPADL raw total scores and higher transformed APPADL total scores correspond to better self-reported ability to perform physical activities of daily living.

Missing data imputation for APPADL can be found in Section 2.5.9.

IW-SP:

The IW-SP is a validated questionnaire to assess an individual's self-perception related to his or her weight. It was developed with moderately obese individuals with T2DM (BMI 25 to 40 kg/m²) (4). It includes 3 items on unhappiness with appearance, self-consciousness in social situations, and overall self-perception, evaluated on a 5-point Likert scale (from 1 [always] to 5 [never]). A total score is calculated by summing the 3 items and dividing by the number of items, it ranges from 1 to 5, with higher scores corresponding to better self-perception. The completion time of the IW-SP is estimated to be less than 2 minutes. It demonstrated robust psychometric properties in patients with T2DM and BMI >30 kg/m² and was able to discriminate patients who achieved at least 5% weight loss from others who did not.

A raw total IW-SP score is calculated by summing the 3 item scores and dividing by the number of items (3). The raw IW-SP total score ranges from 1 to 5. To transform the IW-SP total score to a scale ranging from 0 to 100, the same transformation formula as APPADL is used. For example, using the formula, a raw IW-SP total score of "2.0" becomes a transformed total score of "25.0" on a scale of 0 to 100:

$$[(2.0-1.0)/5-1] *100 = 1/4 *100 = 25.0$$

Higher raw IW-SP total scores and higher transformed IW-SP total scores correspond to better self-perception.

Missing data imputation for IW-SP can be found in Section 2.5.9.

Patient qualitative assessment of the treatment

In addition of the 3 PRO questionnaires, patients will be asked to answer questions at the end of the treatment period (Week 26). This patient qualitative assessment of treatment aims to better understand patients' views on benefits and disadvantages of the treatment they experienced during the trial. The two first questions will ask patients to describe in free text the benefits and disadvantages of the drug they experienced during the trial. A third question will ask patients if they would be willing to pursue the treatment they had during the trial (yes/no) and to describe the reasons in free text. Finally, patients will be asked to give their perception of the drug benefit-risk on 7-point Likert scale ranging from -3 (disadvantages significantly outweigh the benefits) to 3 (Benefits significantly outweigh the disadvantages) and where 0 corresponds to 'Equal benefits and disadvantages'. This patient qualitative assessment should take between 10 and 20 minutes. Patients' answers to the 3 first questions will be analyzed qualitatively. Due to the availability of this questionnaire only in English, it will be limited to English speakers in United States and Canada only.

2.2 DISPOSITION OF PATIENTS

This section describes patient disposition for both patient study status and the patient analysis populations.

Screened patients are defined as all patients who have signed the informed consent.

Randomized patients consist of all patients who have signed informed consent, with a treatment kit number allocated and recorded in the interactive response technology (IRT) database, regardless of whether the treatment kit was actually used or not.

For patient study status, the total number of patients in each of the following categories will be presented in the clinical study report (CSR), using a flowchart diagram or summary table:

- Screened patients
- Screen failure patients and reasons for screen failure
- Nonrandomized but treated patients
- Randomized patients
- Randomized but not treated patients
- Randomized and treated patients
- Patients included in each of the substudy (meal test and holter monitor)
- Patients who have completed the 26-week treatment period as per protocol
- Patients who permanently discontinued the IMP, the main reason for treatment discontinuation, and subject's request for treatment discontinuation
- Patients who complete the study as per protocol
- Patients who did not complete the study as per protocol and the reasons for study discontinuation
- Patients' end of study status (completed, not completed) by end of treatment status (completed, not completed)
- Status at last study contact.

For the screened, screen failure, and non-randomized but treated patients, percentages will be calculated using the number of screened patients as the denominator. All other categories of patients will be presented by treatment group, and percentages will be calculated using the number of randomized patients within each treatment group as the denominator. Reason for treatment/study discontinuation will be provided in tables giving numbers and percentages by treatment group.

A summary of the distribution of patients by country and center will also be provided (overall number of patients screened, screened but not randomized, randomized, and randomized but not treated, and number of patients randomized and discontinued from study treatment for each treatment group).

Patients treated but not randomized, randomized but not treated, patients randomized but not treated as randomized will be identified and described in separate listings. Patients of the second category (randomized and not treated) will be part of efficacy analyses. Patients of the third

category (randomized and not treated as randomized) will be part of efficacy and safety analyses (see Section 2.3). Safety data of patients treated but not randomized will be reported separately.

The randomization strata [HbA1c at screening (<8%, ≥8%) and Visit 4 (Day 1) BMI (<35.0 kg/m^2, ≥35.0 kg/m^2)] assigned by IRT will be summarized. The percentages will be calculated using the number of randomized patients as the denominator. The discrepancy between the strata assigned by IRT and the information reported on eCRF will be listed for all randomized patients.

Kaplan-Meier (KM) plots of the cumulative incidence of IMP discontinuation due to any reason or due to AE will be provided for the 26-week treatment period. Additionally, a listing of these patients, along with the reason for treatment discontinuation, study completion status and the reason for study discontinuation, will be provided. Time to treatment discontinuation will be defined as the number of days from the first injection of IMP until the day of treatment discontinuation. All completers will be considered as censored observations. The censoring time will be the number of days from the first injection of IMP until the last injection date during 26-week treatment period.

All major or critical deviations including deviations related to randomization procedure will be listed and summarized in tables giving numbers and percentages of deviations by randomized treatment group.

Additionally, the analysis populations for efficacy, safety, PK, and each of the substudy defined in Section 2.3 will be summarized in a table by number of patients on the randomized population:

- Efficacy population: modified intent-to-treat (ITT) population,
- Safety population.
- PK population.
- Substudies

2.2.1 Major deviations related to randomization procedure

Major deviations related to randomization procedure occur whenever:

1. A randomization is not in accordance with the protocol-defined randomization method, such as a) an ineligible patient is randomized, b) a patient is randomized based on an incorrect stratum, c) a patient is randomized twice, or d) in a dynamic randomization scheme the treatment assignment is, in fact, not random, due to a computer program error.

   OR

2. A patient is dispensed an IMP kit not allocated by the protocol-defined randomization, such as a) a patient at any time in the study is dispensed a different treatment kit than as randomized (which may or may not contain the correct-as-randomized IMP), or b) a nonrandomized patient is treated with IMP reserved for randomized patients.

Major deviations related to randomization procedure will be monitored throughout the study and reviewed on an ongoing basis.

All major deviations related to randomization procedure will be documented in the CSR. The deviations will be categorized and summarized among randomized patients (number and percentages). Nonrandomized, treated patients will be described separately.

2.3 ANALYSIS POPULATIONS

Patients treated without being randomized (patients who received IMP without calling the IRT or before calling the IRT, or patients who did not give their informed consent) will not be considered as randomized. They will be excluded from any population for analysis, including efficacy and safety. However, if these patients experience any safety event, they will be presented and listed separately in the appendix of the CSR.

The randomized population includes any patient who has been allocated to a randomized treatment regardless of whether the treatment kit was used.

For any patient randomized more than once, only the data associated with the first randomization will be used in any analysis population. The safety experience associated with any later randomization will be assessed separately.

2.3.1 Efficacy populations

Efficacy analyses will be based on the ITT population, defined as all randomized patients, irrespective of compliance with the study protocol and procedures. Efficacy analyses will be based on the treatment group allocated by the IRT according to the randomization schedule at randomization visit (as randomized), irrespective of the treatment actually received.

2.3.2 Safety population

Safety analyses will be based on the safety population, defined as all randomized patients who receive at least 1 dose of IMP (regardless of the amount of treatment administered). Patients will be analyzed for safety analyses according to the treatment actually received. For a patient randomized to a SAR425899 treatment arm, the patient will be analyzed in the treatment he/she receives starting Week 8 visit (inclusive) when no adjustment of dose level is allowed. However, for a patient randomized to a SAR425899 treatment arm and discontinues the IMP before Week 8 visit, the patient will be analyzed in the SAR425899 treatment arm the patient is randomized to. If a patient is exposed to more than one dose level of SAR425899 since Week 8 visit or later, then the patient will be analyzed in the dose level administered on Week 8 visit or study day 56 if Week 8 visit is not available. If a patient is treated with both placebo and SAR425899 any time during the study, the patient will be analyzed in the SAR425899 treatment arm. Liraglutide patients will be included in the Liraglutide treatment arm regardless of dose amount patients receive.

In addition:

- Nonrandomized but treated patients will not be part of the safety population; however, their safety data will be presented separately;

(electronic 1.0)

- Randomized patients for whom it is unclear whether they took the IMP will be included in the safety population as randomized;
- Patients will be excluded from the safety population only if there is documented evidence (ie, all study dates recorded as no medication taken) that patients have not taken the study medication.

2.3.3 Pharmacokinetic population

For PK analyses, the PK population is defined as all randomized and treated patients who contribute with at least 1 valid plasma concentration of SAR425899. The PK data will be analyzed according to the treatment actually received (see Section 2.3.2).

2.3.4 Pharmacodynamic population

The analysis of pharmacodynamics biomarkers will be conducted on the ITT population. The data will be analyzed according to the treatment as randomized (see Section 2.3.1).

2.3.5 NASH and cardiovascular biomarkers population

The analysis of NASH and cardiovascular biomarkers will be conducted on the ITT population. The data will be analyzed according to the treatment as randomized (see Section 2.3.1).

2.3.6 Patient-reported outcome population

The analysis of PROs will be conducted on the ITT population. The PROs data will be analyzed according to the treatment as randomized (see Section 2.3.1).

2.3.7 Substudy population

Patients included in a substudy (meal test, or holter monitor substudy) are defined as patients who signed the informed consent for the substudy, and have at least one post baseline measurement for any of the substudy variables.

Meal test related variables are:

- PPG and in blood glucose excursion in response to a standardized meal test in up to 50% subset of all patients from baseline to Week 26 at time points premeal, 10min, 20min, 0.5h, 1h, 1.5h, 2h, 3h, 4h post meal,
- Insulin, proinsulin, and C-peptide and associated excursion values in response to a standardized meal test in up to 50% subset of all patients from baseline to Week 26 at time points premeal, 10min, 20min, 0.5h, 1h, 1.5h, 2h, 3h, 4h post meal,
- VAS for appetite/satiety score in response to a standardized meal test in up to 50% subset of all patients from baseline to Week 26 at time points premeal, 4 hours post meal.

Holter monitor related variables are

- Twenty–four hour average heart rate,
- Night time average heart rate.

2.4 STATISTICAL METHODS

2.4.1 Demographics and baseline characteristics

Continuous data will be summarized by treatment group using the number of available observations (N), mean, standard deviation (SD), minimum, median, and maximum. Categorical and ordinal data will be summarized using counts and percentages in each treatment group.

Parameters will be summarized based on the randomized population analyzed in the treatment group to which they were randomized. Analyses for the safety population will be included in the appendices if the size of the safety population is different (>10%) from the size of that in the primary analysis population (ie, randomized patients) for any treatment group.

Parameters described in Section 2.1.1 will be summarized by treatment group and overall (pooled across treatment groups) using descriptive statistics.

P-values on the treatment difference for the demographic and baseline characteristic data will not be calculated.

In general, no specific description of the safety parameters will be provided at baseline. If relevant, the baseline values will be described along with each safety analysis.

In general, no specific description of the efficacy parameters will be provided at baseline. If relevant, the baseline values will be described along with each efficacy analysis.

2.4.2 Prior or concomitant medications

The prior, concomitant and posttreatment medications will be presented in the randomized population for each treatment group (and overall for the summary of prior medications), using counts and percentages. No statistical test for the between-group difference will be performed.

Medications will be summarized by treatment group according to the WHO-DD dictionary, considering the first digit of the anatomic category (ATC) class (anatomic category) and the first 3 digits of the ATC class (therapeutic category). A given medication may be classified in more than 1 ATC class. All ATC codes corresponding to a medication will be summarized, and a patient will be counted once in each ATC category (anatomic or therapeutic) linked to the medication. Therefore a patient may be counted several times for the same medication.

Prior medications will be presented by anatomic and therapeutic categories and sorted by decreasing frequency of ATC based on the overall incidence across treatment groups. In case of equal frequency regarding ATCs (anatomic or therapeutic categories), alphabetical order will be used.

Concomitant and posttreatment medications will be presented by anatomic and therapeutic categories and sorted by decreasing frequency of ATC based on the incidence in the SAR425899 0.20 mg dose group. In case of equal frequency regarding ATCs (anatomic or therapeutic categories), alphabetical order will be used.

(electronic 1.0)

2.4.3 Extent of investigational medicinal product exposure and compliance

The extent of IMP exposure and compliance will be assessed and summarized by actual treatment received within the safety population (Section 2.3.2).

2.4.3.1 Extent of investigational medicinal product exposure

The extent of IMP exposure will be assessed by the duration of IMP exposure.

The duration of IMP exposure is defined as last dose date of double-blind IMP – first dose date of double-blind IMP + 1 day, regardless of unplanned intermittent discontinuations (see Section 2.5.3 for calculation in case of missing or incomplete data).

Duration of IMP exposure (days) will be summarized descriptively as a quantitative variable (number of patients exposed, mean, SD, minimum, median, and maximum). In addition, duration of treatment exposure will also be summarized categorically by numbers and percentages for each of the following categories and cumulatively according to these categories:

- 1 to 14 days,
- 15 to 28 days,
- 29 to 56 days,
- 57 to 98 days,
- 99 to 182 days,
- >182 days.

Additionally, the cumulative duration of treatment exposure will be provided, defined as the sum of the duration of treatment exposure for all patients, and will be expressed in patient years.

2.4.3.2 Compliance

A given administration will be considered noncompliant if the patient did not take the planned dose of treatment as required by the protocol. No imputation will be made for patients with missing or incomplete data.

Percentage of compliance for a patient will be defined as the number of days that the patient was compliant divided by the total number of days that the patient was planned to take during the treatment epoch defined in Section 2.1.4 (ie, from the first date to the last date of double-blind IMP administration).

Above-planned dosing percentage for a patient will be defined as the number of days that the patient took a higher dose than planned divided by the total number of days that the patient was planned to take during the treatment epoch.

Under-planned dosing percentage for a patient will be defined as the number of days that the patient took a lower dose than planned divided by the total number of days that the patient was planned to take during the treatment epoch.

Treatment compliance, above-planned, and under-planned dosing percentages will be summarized descriptively as quantitative variables (number, mean, SD, median, minimum, and maximum). The percentage of patients whose compliance is <80% will be summarized. In addition, numbers and percentages of patients with at least 1 above-planned dose administration will be provided, as well as numbers and percentages of patients with 0, (0, 20%], and >20% under-planned dose administrations.

The planned dose is defined to be the intended dose as collected in the eCRF.

Cases of overdose (see study protocol for further details) will constitute AEs and be analyzed as such. More generally, dosing irregularities related to randomization procedure will be listed in Section 2.2.1.

2.4.4 Analyses of efficacy endpoints

Efficacy analyses will be performed on the ITT population using efficacy assessment collected during the study, including those obtained after IMP discontinuation or introduction of rescue therapy, unless otherwise specified.

Unscheduled post-baseline measurements will be allocated to a post-baseline visit according to the criteria described in Section 2.5.4

2.4.4.1 Analysis of primary efficacy endpoint(s)

The statistical test will be 2-sided tests at a nominal 5% significance level.

Primary analysis

The primary efficacy endpoint of change in HbA1c from baseline to Week 26 will be analyzed with missing post-baseline values imputed by placebo control-based multiple imputation (MI) method under the missing not at random framework (5, 6, 7).

- For placebo group patients, missing data will be imputed based on the placebo group data.
- For patients in the SAR425899 dose groups and liraglutide group, missing data will be imputed as if the patients were on placebo group throughout the study.

For missing HbA1c value at baseline, missing data will be imputed using MI under the missing at random (MAR) assumption. Missing data at baseline will be imputed using regression method that includes randomization stratum of screening HbA1c (≤8.0%, >8.0%), randomization stratum of Visit 4 (Day 1) BMI (<35.0 kg/m^2, ≥35.0 kg/m^2), and baseline value in the imputation model (see sample code Part 1c in Appendix D).

The change from baseline at Week 26 values will be derived from observed or imputed Week 26 HbA1c values (see sample code Part 1a in Appendix D). Each of the complete dataset will be analyzed by the Analysis of Covariance (ANCOVA) model with treatment groups (SAR425899 0.12 mg, SAR425899 0.16 mg, SAR425899 0.20 mg, placebo and liraglutide), randomization strata of screening HbA1c value (<8, ≥8%), randomization strata of Visit 4 (Day 1) BMI

(<35.0 kg/m^2, ≥35.0 kg/m^2), and country as fixed effects, and baseline HbA1c value as a covariate(see sample code Part 2 in Appendix D).

Results from each complete dataset will be combined using Rubin's rule to provide the following. Three trend tests are planned on HbA1c. First trend test will be performed to test the hypothesis that HbA1c reduction from baseline at Week 26 increases with increase in dose. The test will be based on a contrast under the ANCOVA framework and with coefficients of +3, +1, -1, -3 and 0 for SAR425899 0.20 mg, SAR425899 0.16 mg, SAR425899 0.12 mg, placebo and liraglutide, respectively. The second trend test (SAR425899 0.16 mg versus placebo) is performed with coefficients of 0, +1, 0, -1 and 0 for SAR425899 0.20 mg, SAR425899 0.16 mg, SAR425899 0.12 mg, placebo and liraglutide, respectively. The third trend test (SAR425899 0.12 mg versus placebo) is performed with coefficients of 0, 0, 1, -1 and 0 for SAR425899 0.20 mg, SAR425899 0.16 mg, SAR425899 0.12 mg, placebo and liraglutide, respectively. Adjusted mean change in HbA1c from baseline to Week 26 for each treatment group, as well as the between-group difference (comparing each SAR425899 dose group versus placebo) and the 95% confidence interval (CI) for the between-group difference will also be provided. The overall Type 1 error for multiple comparisons of the HbA1c and body weight will be controlled by a Hierarchical testing procedure as detailed in Section 2.4.4.3.

The ANCOVA model will be implemented using statistical analysis system (SAS®) (Version 9.4 or higher) MIXED procedure (PROC MIXED).

Sensitivity analyses

Tipping point analysis based on the control-based MI method will be performed to examine the robustness of the results from the primary analysis (7). Patients who were randomized to SAR425899 dosing groups and liraglutide group and had no value at Week 26 will be given a penalty of 0.2%, ie, they are assumed to have received a treatment inferior to placebo by 0.2%. The penalty will be gradually increased by 0.2% at a time to evaluate at which level the conclusion of the analyses in terms of statistical significance is changed for the first trend test. The tipping point is the penalty level, at which the magnitude of efficacy reduction in patients without HbA1c value at Week 26 creates a shift in the treatment effect of each of SAR425899 dose group so that the trend test result becomes insignificant. The steps to perform the tipping point analysis are as follows:

1. Missing data will be imputed using control-based MI (see sample code Part 1b in Appendix D),
2. All imputed Week 26 HbA1c values in the SAR425899 dose groups and Liraglutide group will be penalized by adding 0.2% in each complete dataset (see sample code Part 1b in Appendix D),
3. Change from baseline at Week 26 in HbA1c will be analyzed using the same ANCOVA model as specified in the primary analysis in each complete dataset (see sample code Part 2 in Appendix D),
4. Results will be combined across complete datasets using Rubin's rule (see sample code Part 3 in Appendix D),
5. Steps 2 to 5 will be repeated with penalty increased by 0.2% until the p-value for the first trend test is >0.05.

Summary statistics at scheduled visits

Summary statistics (for screening value, baseline value, observed values, and observed changes from baseline) at scheduled visits (using OC) will be provided for each treatment group. The summary will include the number of observations, mean, SD, standard error (SE), minimum, median, and maximum. Graphical presentations will also be used to examine trends over time using mean values (±SE) and mean changes from baseline (±SE) at each of the scheduled visits. These analyses will be performed using efficacy assessments obtained during the 26-week treatment period, including those obtained after IMP discontinuation or introduction of rescue therapy.

2.4.4.2 Analyses of secondary efficacy endpoints

Descriptive statistics (number, mean, standard deviation, median, minimum, and maximum) will be provided by treatment for all continuous secondary variables at the scheduled visits.

All continuous secondary efficacy endpoints at Week 26 defined in Section 2.1.3.2 will be analyzed using the same ANCOVA model with missing post baseline values imputed by placebo control-based MI method as described in Section 2.4.4.1 to compare each SAR425899 dose group with placebo. For missing value at baseline, missing data will be imputed using the same MI under the MAR assumption as described in Section 2.4.4.1. The model will include fixed effect terms of treatment groups (SAR425899 0.12 mg, SAR425899 0.16 mg, SAR425899 0.20 mg, placebo and liraglutide), randomization strata of screening HbA1c value (<8, ≥8%), randomization strata of Visit 4 (Day 1) BMI (<35.0 kg/m^2, ≥35.0 kg/m^2), and country, and the covariate of baseline value. The three trend tests described in Section 2.4.4.1 are planned on body weight only. For all continuous secondary efficacy endpoints, means and adjusted means of each treatment group will be provided, as well as adjusted mean and associated two-sided 95% CI of the differences between each SAR425899 dose group and placebo. The analyses include values at baseline and Week 26 (observed or imputed) regardless of treatment discontinuation or initiation of rescue therapy.

All categorical secondary efficacy endpoints defined in Section 2.1.3.2 will be analyzed using a Cochran-Mantel-Haenszel (CMH) method stratified on randomization strata of screening HbA1c value (<8, ≥8%) and randomization strata of Visit 4 (Day 1) BMI(<35.0 kg/m^2, ≥35.0 kg/m^2). The proportion in each treatment group will be provided, as well as the difference of proportions between each SAR425899 dose group and placebo with associated 2-sided 95% CI. For HbA1c (<6.5%, <7% respectively at Week 26) and body weight (≥5%, ≥10% body weight loss respectively from baseline at Week 26) responders, all values at Week 26 will be used to determine whether a patient is a responder or not, even if they are measured after IMP discontinuation or rescue medication use. Patients who have no measurement at Week 26 will be treated as non-responders.

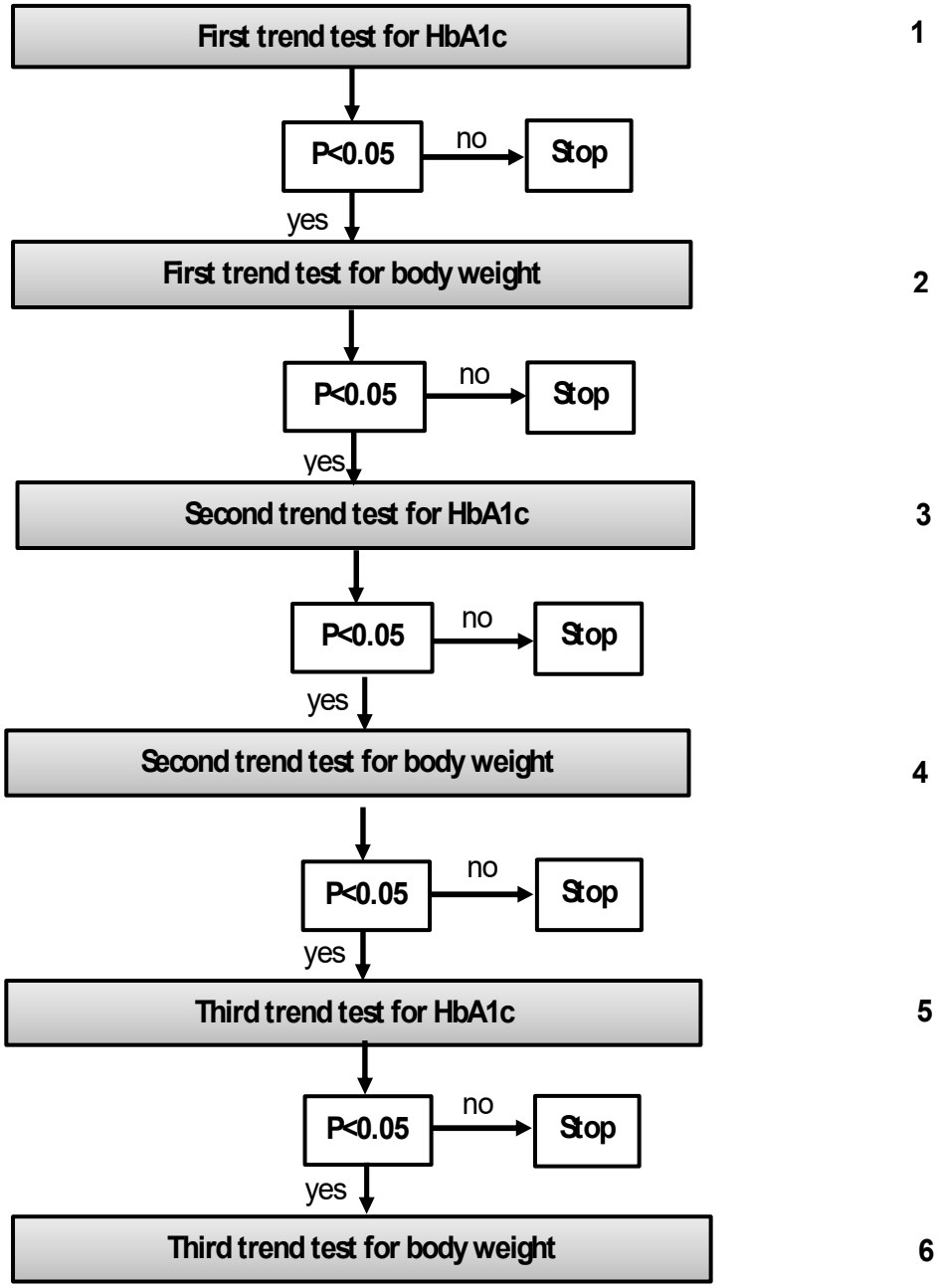
For percentage of patients requiring rescue therapy, patients who have received rescue therapy during the 26-week treatment period will be summarized.

For HbA1c and body weight responders, a sensitivity analysis will be performed treating patients as non-responders if they receive rescue therapy, or have no measurement at Week 26. The sensitivity analysis includes measurement at Week 26, regardless of treatment discontinuation.

2.4.4.3 Multiplicity issues

To control the type I error, a hierarchical testing procedure will be applied. The testing procedure will be performed to test the primary and secondary efficacy variables (HbA1c and body weight) by the following prioritized order. The test stop as soon as an endpoint is found not statistically significant at α=0.05 (2-sided).

Figure 1 - The hierarchical testing procedure



No multiplicity adjustment will be made on other secondary efficacy variables other than body weight.

(electronic 1.0)

2.4.5 Analyses of safety data

The summary of safety results will be presented by treatment group.

The "observation period" defined in Section 2.1.4 is applicable in all safety analyses for the classification of AEs, potentially clinically significant abnormality (PCSA) values for the laboratory, vital sign, ECG and Holtor monitoring.

General common rules

All safety analyses will be performed on the safety population as defined in Section 2.3.2, unless otherwise specified, using the following common rules:

- Safety data in patients who do not belong to the safety population (eg, exposed but not randomized) will be listed separately;
- The baseline value is defined as the last available value before the first injection of IMP.
- The PCSA values are defined as abnormal values considered medically important by the Sponsor according to predefined criteria/thresholds based on literature review and defined by the Sponsor for clinical laboratory tests, vital signs, and ECG (PCSA version dated May 2014 [Appendix A]).
- PCSA criteria will determine which patients had at least 1 PCSA during the TEAE period, taking into account all evaluations performed during the TEAE period, including unscheduled or repeated evaluations. The number of all such patients will be the numerator for the treatment-emergent PCSA percentage;
- The treatment-emergent PCSA denominator by group for a given parameter will be based on the number of patients assessed for that given parameter during the TEAE period by treatment group in the safety population;
- For laboratory parameters cited in the protocol as efficacy endpoints (including HbA1c and plasma glucose), PCSA summaries will not be provided. These parameters will be summarized in Section 2.4.4;
- For quantitative safety parameters based on central laboratory/reading measurements, descriptive statistics will be used to summarize results and change from baseline values by visit and treatment group. Summaries will include the last on-treatment value (see Section 2.4.5.4);
- The analysis of the safety variables will be essentially descriptive and no systematic testing is planned. Relative risks versus placebo and their 95% CIs may be provided, if relevant;
- Selected safety analyses will be summarized by age group (<50, ≥50 to <65, ≥65 to <75, ≥75 years of age), gender, race subgroups, and any pertinent subgroups as appropriate(see details in Section 2.4.5.1 and Section 2.4.5.2).

2.4.5.1 Analyses of hypoglycemia

Analyses of hypoglycemia will be performed on the TEAE period, as defined in Section 2.1.4. Hypoglycemia will be classified as severe hypoglycemia, documented symptomatic hypoglycemia, asymptomatic hypoglycemia, probable symptomatic hypoglycemia or relative hypoglycemia (see Section 2.1.4.1).

The number (%) of patients with any hypoglycemia, severe hypoglycemia and documented symptomatic hypoglycemia will be summarized respectively by treatment group for the TEAE period, as well as the incidence rate in patient years. Two types of incidence rates will be presented: the number of patients with at least 1 event per patient years (calculated as the number of patients with at least 1 event / total exposure in patient years), and the number of events per patient years (calculated as the number of events / total exposure in patient years). Note: here exposure is duration of TEAE period, ie, duration of IMP treatment in days +1 (see Section 2.1.4).

The summary of frequency and incidence rate in patient years for severe hypoglycemia or documented symptomatic hypoglycemia will be provided as appropriate by gender (Male, Female), age group (<50, ≥50 to <65, ≥65 to <75, ≥75 years of age), race (American Indian or Alaska Native, Asian/Oriental, Black, Native Hawaiian or Other Pacific Islander, White, Other).

The summary of frequency and incidence rate in patient years for severe hypoglycemia or documented symptomatic hypoglycemia will be provided as appropriate by anti-SAR425899 antibody status (Positive, Negative) during TEAE period. A patient is defined as anti-SAR425899 antibody positive during the TEAE period if the patient is anti- SAR425899 antibody positive at any time during TEAE period, ie, duration of IMP treatment in days +28 (see Section 2.1.4).

A KM curve will also be provided by treatment group for the time to first onset of severe or documented symptomatic hypoglycemia during the TEAE period.

A listing of patients for all hypoglycemic events reported on the dedicated eDiary or eCRF "Hypoglycemic Event Information" form will be provided with each category flagged (ie, severe hypoglycemia, documented symptomatic hypoglycemia, asymptomatic hypoglycemia, probable symptomatic hypoglycemia and relative hypoglycemia).

2.4.5.2 Analyses of adverse events

Generalities

The primary focus of AE reporting will be on treatment-emergent adverse events (TEAEs). Pre- and post-treatment AEs will be described separately.

If an AE date/time of onset (occurrence, worsening, or becoming serious) is incomplete, an imputation algorithm will be used to classify the AE as pre-treatment, treatment-emergent, or post-treatment. The algorithm for imputing date/time of onset will be conservative and will classify an AE as treatment-emergent unless there is definitive information to determine it is pre- or post-treatment. Details on classification of AEs with missing or partial onset dates are provided in Section 2.5.3.

Adverse event incidence tables will be presented by SOC, HLGT, HLT, and PT, sorted by the internationally agreed order for SOCs and alphabetic order for HLGT, HLT and PT within a SOC, the number (n) and percentage (%) of patients experiencing an adverse event. Multiple occurrences of the same event in the same patient will be counted only once in the tables within a treatment phase. The denominator for computation of percentages is the safety population within each treatment group.

Sorting within tables ensures the same presentation for the set of all adverse events within the observation period (pretreatment, treatment-emergent, and posttreatment). For that purpose, the table of all treatment-emergent adverse events presented by primary SOC and PT (sorted by the internationally agreed SOC order and decreasing frequency of PTs within SOCs in the SAR425899 0.20 mg group) will define the presentation order for all other similar tables unless otherwise specified. In case of equal frequency regarding PTs, alphabetical order will be used.

Analysis of all treatment-emergent adverse events

The following TEAE summaries will be generated for the safety population.

- Overview of TEAEs, summarizing number (%) of patients with any:
  - TEAE,
  - Serious TEAE,
  - TEAE leading to death,
  - TEAE leading to permanent treatment discontinuation;
- All TEAEs by primary SOC, showing number (%) of patients with at least 1 TEAE, sorted by internationally agreed order of primary SOCs;
- All TEAEs by PT, showing number (%) of patients with at least 1 TEAE, sorted by decreasing incidence of PT in the SAR425899 0.20 mg group;
- All TEAEs by primary SOC, HLGT, HLT and PT, showing number (%) of patients with at least 1 TEAE, sorted by internationally agreed SOC order. The other levels (HLGT, HLT, PT) will be presented in alphabetic order;
- All TEAEs by primary SOC and PT, showing number (%) of patients with at least 1 TEAE, sorted by internationally agreed SOC order and decreasing incidence of PTs within a SOC in the SAR425899 0.20 mg group. This sorting order will be applied to all other similar tables, unless otherwise specified;
- Common TEAEs (PTs with an incidence ≥3% in any treatment group) by primary SOC, HLGT, HLT, and PT, sorted by internationally agreed order of SOCs. The other levels (HLGT, HLT, PT) will be presented in alphabetic order;
- All TEAEs regardless of relationship and related to IMP by primary SOC, HLGT, HLT, and PT, showing number (%) of patients with at least 1 TEAE, sorted by internationally agreed order of SOCs. The other levels (HLGT, HLT, PT) will be presented in alphabetic order;
- All TEAEs by maximal severity, presented by primary SOC and PT, showing number (%) of patients with at least 1 TEAE by severity (ie, mild, moderate, or severe), sorted by sorting order defined above;
- Kaplan-Meier curves will be provided, when appropriate, for the time to first onset of the following PTs: nausea and vomiting;

(electronic 1.0)

- Summaries of common TEAEs (PTs with an incidence ≥3% in any treatment group) will be provided as appropriate by primary SOC, HLGT, HLT, and PT and by demographic factors including gender (Male, Female), age group (<50, ≥50 to <65, ≥65 to <75, ≥75 years of age), and race (American Indian or Alaska Native, Asian/Oriental, Black, Native Hawaiian or Other Pacific Islander, White, Other). SOC will be sorted by internationally agreed order and the other levels (HLGT, HLT, PT) in alphabetic order.
- Summaries of common TEAEs (PTs with an incidence ≥3% in any treatment group) will be provided as appropriate for SAR425899 treatment groups by primary SOC, HLGT, HLT, and PT and by anti- SAR425899 antibody status (Positive, Negative) during the TEAE period as necessary. SOC will be sorted by internationally agreed order and the other levels (HLGT, HLT, PT) in alphabetic order.

Analysis of all treatment emergent serious adverse event(s)

- All treatment-emergent SAEs by primary SOC, HLGT, HLT, and PT, showing the number (%) of patients with at least 1 treatment-emergent SAE, sorted by the internationally agreed SOC order. The other levels (HLGT, HLT, PT) will be presented in alphabetical order.
- All treatment-emergent SAEs regardless of relationship and related to IMP, by primary SOC, HLGT, HLT, and PT, showing the number (%) of patients with at least 1 treatment-emergent SAEs, sorted by the internationally agreed SOC order. The other levels (HLGT, HLT, PT) will be presented in alphabetical order.

Analysis of all treatment-emergent adverse event(s) leading to treatment discontinuation

- All TEAEs leading to treatment discontinuation, by primary SOC, HLGT, HLT, and PT, showing the number (%) of patients with at least 1 TEAE leading to treatment discontinuation, sorted by the internationally agreed SOC order. The other levels (HLGT, HLT, PT) will be presented in alphabetical order.

Analysis of AESI

The number (%) of patients with specific AESIs will be summarized by category (as identified in Section 2.1.4.2) and PT, and sorted by decreasing incidence in the SAR425899 0.20 mg group.

All AESI events reported by the Investigators will be listed.

Analysis of pancreatic events

- Pancreatitis
  - The number (%) of patients with events positively adjudicated as pancreatitis by the PSAC will be summarized by type: 1) acute pancreatitis, 2) acute on chronic pancreatitis, 3) chronic pancreatitis, 4) unknown pancreatitis.
  - All events sent to PSAC for adjudication for pancreatitis will be listed along with the adjudication outcome.
- Pancreatic neoplasm
  - All the events sent to PSAC for adjudication for pancreatic neoplasm will be listed along with the adjudication outcome including PSAC diagnosis, type of neoplasm (malignant, benign), cancer stage etc.

(electronic 1.0)

Analysis of major cardiovascular events

The number (%) of patients with events adjudicated as major cardiovascular events by CAC will be summarized by treatment group. All events reported by the Investigators on the AE forms for cardiovascular events and the associated complementary forms (confirmed or not confirmed by CAC) will be listed along with the adjudication outcome.

Local tolerability at injection site

Adverse events related to local intolerability at the injection site will be identified by searching the term "injection site" in the PTs coded from the Investigator reported terms. The number (%) of patients with related events will be summarized by treatment group, and sorted by decreasing incidence in the SAR425899 0.20 mg group.

All adverse events related to local intolerability at the injection site will be listed.

Analysis of pre-treatment and post-treatment adverse events

- All pre-treatment AEs by primary SOC and PT, showing the number (%) of patients with at least 1 pre-treatment AE, sorted by the internationally agreed SOC order and decreasing incidence of PTs within each SOC in the SAR425899 0.20 mg group.
- All post-treatment AEs by primary SOC and PT, showing the number (%) of patients with at least 1 post-treatment AE, sorted by the internationally agreed SOC order and decreasing incidence of PTs within each SOC in the SAR425899 0.20 mg group.

Listings

Supportive AE listings will be provided as appropriate for all AEs, SAEs and AEs leading to treatment discontinuation, TEAE leading to death, AESI, pancreatic events, major cardiovascular events, and adverse events related to local intolerability at the injection site. These listings will include at least the following information, sorted by treatment, patient identification, and onset date: treatment, patient identification, country, age, gender, race, BMI, primary SOC, PT, reported term, onset date, study day (relative day to the start date of treatment), AE duration, duration of exposure, intensity, corrective treatment, action taken with IMP, date of treatment discontinuation (if relevant), relationship to IMP/NIMP/study procedures, outcome, date of death (if any), seriousness, seriousness criteria, and AE status ("Pre" for a pre-treatment AE; "T" for a TEAE; and "Post" for a post-treatment AE).

2.4.5.3 Deaths

The following summaries of deaths will be generated for the safety population.

- Number (%) of patients who died by study period (on-study, on-treatment, post-study)
- Deaths in nonrandomized patients or randomized but not treated patients
- TEAEs leading to death (death as an outcome on the eCRF AE page as reported by the investigator) by primary SOC, HLGT, HLT, and PT, showing number (%) of patients, sorted by internationally agreed order of SOC and alphabetic order of HLGT, HLT, and PT.

2.4.5.4 Analyses of laboratory variables

Laboratory parameters will be grouped and summarized by biological function as described in Section 2.1.4.4.

The summary statistics (including number, mean, median, Q1, Q3, standard deviation, minimum and maximum) of all laboratory variables (central laboratory values and changes from baseline) will be calculated for each applicable visit or study assessment (screening, baseline, postbaseline time points, last on-treatment value) by treatment group.

The incidence of PCSAs (list provided in Appendix A) at any time during the treatment-emergent adverse event period will be summarized by biological function and treatment group whatever the baseline level and/or according to the following baseline status categories:

- Normal/Missing,
- Abnormal according to PCSA criterion or criteria.

For calcitonin and the albumin/creatinine ratio, no PCSA criterion is defined. Similar summaries using the pre-defined categories will be provided for the TEAE period. The pre-defined categories are, for calcitonin ≤ULN, >ULN - 20 ng/L, ≥20 - <50 ng/L, and ≥50 ng/L (Note that ng/L is the standard international unit and is equivalent to pg/mL); and for albumin/creatinine ratio <30 µg/mg creatinine [Normal], ≥30 to <300 µg/mg creatinine [Microalbuminuria], and ≥300 µg/mg creatinine [Macroalbuminuria].

For parameters for which no PCSA criteria are defined, similar table(s) using the normal range will be provided.

All measurements collected during the TEAE period, including values from unscheduled visits, will be considered for the PCSA summaries. These summaries will include patients in the safety population who have at least 1 assessment performed during the TEAE period. When a PCSA definition involves a change from baseline value, patients must also have a baseline value to be included in the summaries, and when required by the definition of the abnormality, patients must also have available laboratory normal ranges.

A listing of patients with at least 1 post-baseline PCSA (or out of normal range when no PCSA criterion is defined) will be provided and will display the entire patients' profile across time for all parameters belonging to the corresponding biological function. Individual data listings will include flags when applicable:

- Baseline values will be flagged "B".
- Normal laboratory ranges, available for most laboratory parameters, will be identified as ULN and LLN. Baseline, last on-treatment value, and individual data will be flagged "L" if the value is below the LLN and will be flagged "H" if it is above the ULN.
- Laboratory PCSA criteria will be used for the corresponding laboratory parameters. Values reaching a PCSA limit will be flagged (+, ++, -, or -- depending upon the direction and level of the abnormality). Flags for WBC and differential counts will be determined using data expressed in international units.

For parameters whose PCSA criteria are multiples of the ULN, the parameter's value will also be expressed as a multiple of the ULN in the individual data provided.

Drug-induced liver injury

The liver function tests, namely AST, ALT, ALP, and total bilirubin, are used to assess possible drug-induced liver toxicity. The proportion of patients with PCSA values at any post-baseline visit by baseline status will be displayed by treatment group for each parameter. The proportion of patients with PCSA values at any post-baseline visit will also be displayed by duration of exposure for each treatment group (only if a tabulation summary is necessary).

Listing of possible Hy's law cases identified by treatment group (eg, patients with any elevated ALT>3 x ULN, and associated with an increase in bilirubin >2 x ULN) with ALT, AST, alkaline phosphatase, total bilirubin, and the following complementary parameters (if available): conjugated bilirubin and prothrombin time/international normalized ratio, creatine phosphokinase, serum creatinine, complete blood count, anti-HAV IgM, anti-HBc IgM, anti-HCV and HCV RNA, anti-CMV IgM and anti-HEV IgM antibodies, auto-antibodies: anti-nuclear, anti-DNA, anti-smooth muscle, Epstein-Barr virus, herpes viruses, and anti-LKM.

2.4.5.5 Analyses of vital sign variables

The summary statistics (including number, mean, median, Q1, Q3, standard deviation, minimum and maximum) of all vital signs variables (HR, SBP, DBP and changes from baseline) will be calculated for each visit or study assessment (baseline, each postbaseline time point, last on-treatment and/or worst on-treatment value and/or other specific assessment) by treatment group. For parameters HR, SBP, and DBP, mean changes from baseline with the corresponding standard error will be plotted over time (at same time points) in each treatment group.

The incidence of PCSAs at any time during the treatment-emergent adverse event period will be summarized by treatment group irrespective of the baseline level and/or according to the following baseline status categories:

- Normal/missing
- Abnormal according to PCSA criterion or criteria

The summaries will include patients in the safety population who have at least 1 assessment performed during the TEAE period. When a PCSA definition involves a change from baseline value, patients must also have a baseline value to be included in the summaries.

A listing of patients with at least 1 post-baseline PCSA will be provided and will display the patient's profile over time of all vital sign parameters. Individual data listings will include the following flags:

- Baseline values will be flagged "B",
- Parameter values reaching a PCSA limit will be flagged (+, or - depending of the direction).

2.4.5.6 Analyses of 12 lead electrocardiogram variables

The summary statistics (including number, mean, median, Q1, Q3, standard deviation, minimum and maximum) of all 12 lead ECG variables (HR, PR, QRS, QT, QTc and changes from baseline) will be calculated for each applicable visit or study assessment (screening, baseline, postbaseline time points, last on-treatment value) by treatment group. For parameters HR, PR, QRS, QT, QTc mean changes from baseline with the corresponding standard error will be plotted over time (at same time points) in each treatment group.

The incidence of PCSAs at any time during the treatment-emergent adverse event period will be summarized by treatment group irrespective of the baseline level and/or according to the following baseline status categories:

- Normal/missing
- Abnormal according to PCSA criterion or criteria

The summaries will include patients in the safety population who have at least 1 assessment performed during the TEAE period. When a PCSA definition involves a change from baseline value, patients must also have a baseline value to be included in the summaries.

A listing of patients with at least 1 post-baseline PCSA will be provided and will display the patient's profile over time of all vital sign parameters. Individual data listings will include the following flags:

- Baseline values will be flagged "B",
- Parameter values reaching a PCSA limit will be flagged (+, ++, -, or -- depending of the direction and level of the abnormality).

A listing of subjects with at least one abnormality in qualitative assessment (ie, abnormal ECG) during the TEAE will be also provided.

2.4.5.7 24-hours Holter ECG

The summary statistics (including number, mean, median, Q1, Q3, standard deviation, minimum and maximum) of all Holter ECG variables (24-hour average and night time average heart rate, and changes from baseline) will be calculated for each applicable visit or study assessment (screening, baseline, postbaseline time points, last on-treatment value) by treatment group. For parameters 24-hour average and night time average heart rate, mean changes from baseline with the corresponding standard error will be plotted over time (at same time points) in each treatment group.

2.4.5.8 Analyses of anti-drug antibody variables

Analyses of antibody variables will be performed on the safety population (only in patients from SAR treatment groups).

The number and percentage of patients by antibody status will be listed and summarized by visit, as well as the percentage of conversion from negative to positive status from baseline to Week 26.

Antibody levels (titer or concentration) will be listed and summarized by visit using descriptive statistics by N, geometric mean, coefficient of variation, median, minimum and maximum.

2.4.6 Analyses of pharmacokinetic variables

SAR425899 plasma concentrations of patients in the SAR425899 treatment groups will be listed and summarized by visit and time window and by antibody status in the PK population, using descriptive statistics by N, geometric mean, coefficient of variation, median, minimum and maximum.

Population PK modeling and simulation results would be provided in a separate report.

2.4.7 Analyses of pharmacodynamics biomarkers

Descriptive statistics (number, mean, standard deviation, median, minimum, and maximum) will be provided by treatment for pharmacodynamics biomarkers at the scheduled visits.

2.4.8 Analyses of NASH and cadiovascular biomarkers

Descriptive statistics (number, mean, standard deviation, median, minimum, and maximum) will be provided by treatment for NASH and cadiovascular biomarkers at the scheduled visits. GGT, ALT, and AST will be analyzed according to Section 2.4.5.4, no additional analyses will be provided.

2.4.9 Analyses of patient reported outcomes (health-related quality of life/health economics variables)

2.4.9.1 Analysis of PRO questionnaires (WRSM, APPADL and IW-SP)

Descriptive statistics (number, mean, standard deviation, median, minimum, and maximum) will be provided by treatment for PRO questionnaires (WRSM, APPDL and IW-SP) at the scheduled visits.

PRO questionnaires (WRSM, APPDL and IW-SP) at Week 26 will be analyzed using the same ANCOVA model with missing post baseline values imputed by placebo control-based MI method as described in Section 2.4.4.1 to compare each SAR425899 dose group with placebo. For missing value at baseline, missing data will be imputed using the same MI under the MAR assumption as described in Section 2.4.4.1. The model will include fixed effect terms of treatment groups (SAR425899 0.12 mg, SAR425899 0.16 mg, SAR425899 0.20 mg, placebo and liraglutide), randomization strata of screening HbA1c value (<8, ≥8%), randomization strata of Visit 4 (Day 1) BMI (<35.0 kg/m^2, ≥35.0 kg/m^2), and country, and the covariate of baseline value. Means and adjusted means of each treatment group will be provided, as well as adjusted mean and associated two-sided 95% CI of the differences between each SAR425899 dose group and placebo. The analyses include values at baseline and Week 26 (observed or imputed) regardless of treatment discontinuation or initiation of rescue therapy.

2.4.9.2 Analysis of the patient qualitative assessment of the treatment

Patients' answers to the 3 open-ended questions will be analyzed qualitatively using qualitative data analysis software. This qualitative thematic analysis of patients answers will be based on grounded theory methods and involve creating codes corresponding to concepts and sub-concepts identified in the text. A global synthesis will establish the most relevant concepts reported by the patients (benefits, disadvantages and reasons to continue or not the drug). Additional descriptive statistics will be provided on the closed-ended questions. The analyses method for this exploratory analysis will be provided in a separate SAP and the analyses results will be documented in a separate report.

2.5 DATA HANDLING CONVENTIONS

2.5.1 General conventions

The following formulas will be used for computation of parameters.

Renal function formulas

The estimated glomerular filtration rate (GFR) will be calculated by 4 variable MDRD formula using the serum creatinine, race, age, and gender of the patient:

Standard unit: eGFR (mL/min/1.73 m^2) = 175 x [Serum Creatinine (µmol/L)/88.4] $^{-1.154}$ x Age (year) $^{-0.203}$ x 1.212 (if Black) x 0.742 (if female)

Conventional unit: GFR (mL/min/1.73 m^2) = 175 x serum creatinine (mg/dL) $^{-1.154}$ x age (yr) $^{-0.203}$ x 1.212 [if black] x0.742 [if female]

Urine albumin/creatinine ratio:

Standard unit: Urine albumin/creatinine ratio (mg/g) = Urine Albumin (mg/dL) / [Urine Creatinine (mmol/L) x 11.31] x 1000

Conventional unit: Urine albumin/creatinine ratio (mg/g) = Urine Albumin (mg/dL) / Urine Creatinine (mg/dL) x 1000

Calculation of LDL-C

When triglycerides is lower than or equal to 4.52 mmol/L (400 mg/dL), LDL-C is calculated using the Friedewald equation as:

Standard unit (mmol/L): total cholesterol – HDL-C – triglycerides/2.17;

Conventional unit (mg/dL): total cholesterol – HDL-C – triglyceride /5.

When triglycerides is greater than 4.52 mmol/L (400 mg/dL), the Friedewald equation can not be used to derive LDL-C value, and the value of LDL-C will be considered as missing.

Linked adverse events that worsened or became serious

An AE that worsened or became serious will have a separate record in the data from the original event record with an AE identification number that links the new record to the original record. An AE that worsened or became serious will be considered a new recurring AE in the summary of recurrent events or in the summary of events by time intervals.

2.5.2 Data handling conventions for secondary efficacy variables

All continuous secondary efficacy endpoints at Week 26 will be analyzed using an ANCOVA model as described in Section 2.4.4.2. The analyses include all observed values at baseline and Week 26 regardless of treatment discontinuation or initiation of rescue therapy. Missing values will be imputed by control-based multiple imputation method under the missing not at random frame work.

For the categorical secondary efficacy endpoints, data handling conventions are described in Section 2.4.4.2.

2.5.3 Missing data

For categorical variables, patients with missing data are not included in calculations of percentages unless otherwise specified. When relevant, the number of patients with missing data is presented.

Derived variables will be considered missing if any of the original variables required to calculate them are missing. For example, if either a baseline assessment or an endpoint assessment is missing for a particular patient, then change from baseline at endpoint will be missing. Depending upon the assessment, analyses may not include all patients in the analysis population, because certain patients in the intended population may have missing data.

Incomplete date of first injection

Date/time of first injection is the first non-missing start date/time of IMP completed in the eCRF "Dosing(SAR425899/Placebo)" or "Dosing(Liraglutide)" module .

For patients who are randomized and dispensed a treatment kit but who are lost to follow-up just after Visit 4 (only the treatment kit number is reported in the eCRF dosing module without any dose information), the date of first injection will be imputed using the Visit 4 date. When a patient is randomized but not dispensed a treatment kit, the dosing module should be blank. In this case, the patient will be considered as randomized but not treated, "Not Taken" should be ticked in the e-CRF dosing module.

Handling of computation of treatment duration if IMP end of treatment date is missing

For the calculation of the treatment duration, the date of the last injection of IMP is equal to the date of last administration reported on the eCRF "Treatment Status" page. If this date is missing, the exposure duration should be kept as missing.

The last dose intake should be clearly identified in the eCRF and should not be approximated by the last returned package date.

Handling of medication missing/partial dates

No imputation of medication start/end dates or times will be performed. If a medication date or time is missing or partially missing and it cannot be determined whether it was taken prior or concomitantly, it will be considered a prior, concomitant, and post-treatment medication.

Handling of missing data for categorical secondary efficacy endpoints

Please see Section 2.4.4.2.

Handling of missing data for continuous efficacy endpoints

Please see Section 2.4.4.1 and Section 2.4.4.2.

Handling of missing data in the calculation of average 7-point SMPG

In order to account for the situation that a patient could have SMPG at bedtime after 12 midnight, a day interval is defined for the study as 2:01 am-2:00 am the next day. All the SMPG measurements taken during the interval will be considered as measurements of the day. If there are more than one measurement for a time-point of a day, the average value will be used instead.

Only glucose measurements taken after the previous visit with scheduled 7-point SMPG measurements, and within a week of the next visit with scheduled 7-point SMPG measurements will be used to choose a SMPG profile for a patient and a visit. A patient has to have value for at least 4 time-points during a day interval for the SMPG profile to be considered as a complete profile. The day with a complete profile and the most glucose measurements (in terms of time-points) will be chosen as the SMPG profile for the patient and the visit. If there are more than one days with the most glucose measurements, the profile from the day closer to the next visit with scheduled 7-point SMPG measurements will be chosen.

For SMPG by time-point analysis, a similar principal will be applied in selecting the data for analysis. A glucose measurement for the time-point taken after previous visit with scheduled 7-point SMPG measurements, and closest to the next visit with scheduled 7-point SMPG measurements will be selected for the by time-point analysis.

Handling of adverse events/hypoglycemia with missing or partial date/time of onset

Missing or partial missing AE/hypoglycemia onset dates and times will be imputed so that if the partial onset date/time information does not indicate that the AE/hypoglycemia started prior to treatment or after the treatment-emergent period, the AE/hypoglycemia will be classified as treatment-emergent. No imputation of AE end dates/times will be performed. These data imputations are for categorization purpose only and will not be used in listings. No imputation is planned for date/time of AE resolution.

Handling of adverse events/hypoglycemia when date and time of first IMP administration is missing

When the date and time of the first double-blind IMP administration is missing, the day of randomization should be considered as the start date of TEAE period**.** The exposure duration should be kept as missing.

Handling of adverse events/hypoglycemia when IMP end of treatment date is missing

For the purpose of defining TEAE period, the date of the last administration of IMP is equal to the date of the last administration reported on the eCRF "Treatment Status" page.

If the date of last administration reported on the eCRF "Treatment Status" page is

- Partially missing, it will be imputed with a date as late as possible before or on the date of last available information on eCRF "Completion of End of Study/Follow-up".
- Completely missing, it will be imputed with the date of last available information on eCRF "Completion of End of Study/Follow-up" page.

If the date of last available information on eCRF "Completion of End of Study/Follow-up" page is

- Partially missing, it will be imputed with a date as late as possible.
- Completely missing, all adverse events occurred on or after the first administration of double-blind IMP will be considered as treatment emergent adverse events.

Handling of missing assessment of relationship of adverse events to IMP/NIMP/study procedures

If the assessment of the relationship to IMP/NIMP/study procedures is missing, the relationship to IMP/NIMP/study procedures has to be assumed and the AE considered as such in the frequency tables of possibly related AEs, but no imputation should be done at the data level.

Handling of missing severity/grades of adverse events

For a patient and a preferred term, if the severity is missing for one of the treatment-emergent occurrences of an AE, the maximal severity on the remaining occurrences for the patient and the preferred term will be considered. If the severity is missing for all the occurrences, a "missing" category will be added in summary table.

Handling of potentially clinically significant abnormalities

If a patient has a missing baseline he will be grouped in the category "normal/missing at baseline."

For PCSAs with 2 conditions, one based on a change from baseline value or a normal range and the other on a threshold value, with the first condition being missing, the PCSA will be based only on the second condition.

For a PCSA defined on a threshold and/or a normal range, this PCSA will be derived using this threshold if the normal range is missing; eg, for eosinophils the PCSA is >0.5 GIGA/L or >ULN if ULN ≥0.5 GIGA/L. When ULN is missing, the value 0.5 should be used.

Measurements flagged as invalid by the laboratory will not be summarized or taken into account in the computation of PCSA values.

2.5.4 Windows for time points / Measurements for analyses

For statistical analyses on efficacy, PRO questionnaires (WRSM, APPADL and IW-SP), a scheduled measurement will be used if it is available, regardless of treatment discontinuation or initiation of rescue therapy. An unscheduled measurement will be allocated to a post-baseline visit based on the study day that the unscheduled measurement is taken. An allocation is performed only if there is no scheduled measurement for the visit to be allocated to. However, if the unscheduled measurement (central laboratory measurement only for laboratory data) is taken on the same date of the scheduled visit, the value will be used in the analyses, regardless whether or not there is a scheduled measurement for the visit.

Table 4 - Allocating an unscheduled measurement to a post-baseline visit

Unscheduled measurement taken in study days[a]	Visit to be allocated to	Targeted study day for the visit
2 - 20	Visit 6 (Week2)	14
21 - 41	Visit 8 (Week 4)	28
42 - 76	Visit 9 (Week 8)	56
77 - 118	Visit 10 (Week 14)	98
119 - 160	Visit 11 (Week 20)	140
≥161	Visit 12 (Week 26)	182

a Study days are calculated from the day of first IMP injection; the day of first IMP injection is Day 1.

If more than one measurements are allocated to a visit, the one closest to the targeted study day (as shown in table above) will be used. Furthermore, if there are more than one measurement that are equally close to the targeted study day but measured on two different dates, measurements from the later date will be used. If there are more than one measurement taken on the same date, the average will be used.

The premature treatment discontinuation visit is entered in the eCRF if a patient prematurely discontinues treatment. These measurements will be allocated to a scheduled visit using the same criteria described above.

Pre-rescue therapy visit is entered in the eCRF if a patient receives rescue therapy. These measurements will also be allocated to a scheduled visit using the same criteria described above.

Furthermore, end of treatment visit measurements from patients prematurely discontinue the study ("Subject Status" in "Completion of End of Study/Follow-up" page is not "Completed") with will also be allocated to a scheduled visit using the same criteria described above.

Similar criteria of allocating an unscheduled measurement to a scheduled visit will be used for safety variables (laboratory variables, vital sign, ECGs, holter monitor variables), including measurements from premature treatment discontinuation visit, pre-rescue therapy visit, and end of treatment visit for patients prematurely discontinue the study if any. However, descriptive statistics, PCSA analyses, and shift summaries for safety variables are based only on data from TEAE period, a detailed definition of the TEAE period is provided in Section 2.1.4.

After allocation, if there is still no measurement for a given parameter at a scheduled visit, multiple imputation would be applied efficacy and PRO questionnaires as described in Section 2.4.4 and Section 2.4.9.

Baseline definition for efficacy/safety data

For the safety analyses, the baseline for a given parameter is defined as the last available measurement, including unscheduled assessments, assessed prior to the first administration of IMP. For the efficacy analyses, the baseline for a given parameter is defined as the last available measurement, including unscheduled assessments, assessed prior to the first administration of IMP or the last available value on or before the date of randomization if not treated with IMP.

Summary statistics by visit for continuous efficacy endpoints and PRO questionnaires (WRSM, APPADL and IW-SP)

Summary statistics (N, mean, SD, SE, minimum, median, maximum) of continuous efficacy endpoints (observed data and change from baseline) will be provided at scheduled visits as per protocol. Summaries showing data by visit will be presented according to the visit number (or re-allocated visit number) and labeled with the targeted week. The analyses will be based on data after applying the allocation of measurements from unscheduled visit, premature treatment discontinuation visit, pre-rescue therapy visit, and end of treatment visit as described above in this section.

Last on-treatment value for safety variables (laboratory variables, vital sign, ECGs, holter monitor variables)

The last on-treatment value is the final measurement assessed during treatment epoch for safety variables, regardless of the introduction of rescue therapy, including measurements at unscheduled visits. The definition of treatment epoch is provided in Section 2.1.4, also see details in Section 2.4.5.

Display of safety data by visit (laboratory variables, vital signs ECGs, holter monitor variables)

Descriptive statistics (N, mean, SD, minimum, median, maximum) of quantitative laboratory variables and vital signs (observed data and change from baseline) during the TEAE period for all scheduled visits as per protocol will be provided (ie, only including patients having non-missing assessments at a nominal visit). The analyses will be based on data after applying the allocation of measurements from unscheduled visit, premature treatment discontinuation visit, pre-rescue therapy visit, and end of treatment visit as described above in this section. In addition, these summaries will also include a row for the ‘last value on-treatment’ to describe the last available on-treatment value (see above). Summaries showing data by visit will be presented according to the visit number (or re-allocated visit number) and labeled with the targeted week.

As specified in the study protocol, laboratory data from scheduled visits are reported by central laboratories. The local results will not be used in the efficacy analyses or in the definition of baseline for both safety and efficacy analyses. In the safety analyses, local results will only be used in the PCSA summary, and only if they are accompanied by a local laboratory normal range.

When a patient has more than 1 measurement from the central laboratory for the same laboratory parameter on the same date, the average of the measurements will be used. For the same laboratory parameter, if a patient has more than 1 measurement on different dates for the same scheduled visit, the value closest to the date of the visit will be used for the scheduled visit. When the values for the same scheduled visit are equidistant to the visit date, the last value should be used for the scheduled visit. Similar rules will be applicable to other safety variables (vital signs ECGs, holter monitor variables).

2.5.5 Unscheduled visits

Unscheduled visit measurements of laboratory data, vital signs, ECG, and holter monitor variables will be used for computation of baseline, the last on-treatment value, PCSAs, and the shift summaries for safety or efficacy. They will be included in the by-visit summaries if they are re-allocated to scheduled visits (see Section 2.5.4) for efficacy variables. For safety variables, they also need to be measured during the TEAE period to be included in the by-visit summaries.

2.5.6 Pooling of centers for statistical analyses

Center will not be included in the statistical models for efficacy analyses. However, all centers within a country will be pooled, and country will be included as fixed effect in the ANCOVA model for primary and secondary efficacy endpoints. Countries with fewer than 5 randomized patients will be grouped with the country with the lowest number of patients that is 5 or more.

2.5.7 Statistical technical issues

None.

2.5.8 Rescued patient data

Patients who met the following condition are considered “rescued”:

- The answer to “Was a rescue therapy started since previous visit?” is “Yes” in the eCRF “Rescue Therapy” page.
- The answer to “Reason for treatment” is “Rescue therapy” in the eCRF “Medications” page.

2.5.9 PRO missing data imputation

APPADL and IW-SP

If >50% of item scores is available, impute mean of that data for missing values. For the APPADL, this means that there must be at least 4 items completed and for the IW-SP, there must be at least 2 items completed for imputation. The raw total score will be derived by summing the nonmissing item scores and dividing by the number of nonmissing items.

If the data for any respondent does not meet the missing data rule (>50% completed), then the data is considered a missing data point and the procedure for missing data points (placebo control-based MI method) is used to analyze the data.

Missing data imputation for WRSM

For WRSM, no more than 4 items can be missing out of 20 items.

If the data for any respondent does not meet the missing data rule (≥80% completed), then the data is considered a missing data point and the procedure for missing data points (placebo control-based MI method) is used to analyze the data.

3 INTERIM ANALYSIS

Body weight interim analysis:

An interim analysis for body weight will be conducted when 108 patients (about 24 patients per dose arm of SAR425899 and liraglutide and 12 patients from the SAR425899 placebo arm) have been randomized and completed 14 weeks of treatment.

Descriptive statistics will be used to summarize the body weight results and the changes from baseline by visit for each treatment group. Number of observations, mean, SD, SE, minimum, median, and maximum will be used to summarize the results (using OC). Other exploratory analyses on body weight will also be performed including using the ANCOVA model with missing post baseline values imputed by placebo control-based MI method as described in Section 2.4.4.1. Modeling and simulation will be performed on body weight data.

In addition, unblinded DMC results (including demographics, disposition, and safety summary) and possibly related additional exploratory analyses will also be provided to support compound development.

PK/PD interim analysis:

An interim analysis for PK/PD will be conducted when 60% of patients have completed the end of treatment visit. Potentially 4 analyses are planned :

- PK
- PK/PD with FPG, HbA1c, and body weight
- PK/safety: with heart rate
- PK/safety: with GI AEs (nausea and vomiting)

In addition, Descriptive statistics (Number of observations, mean, SD, SE, minimum, median, and maximum) will be used to summarize the body weight, HbA1c and FPG results and the changes from baseline by visit for each treatment group based on the efficacy data collected when 60% of patients have completed the end of treatment visit.

Other exploratory analyses on body weight, HbA1c and FPG may also be performed including using the ANCOVA model with missing post baseline values imputed by placebo control-based MI method as described in Section 2.4.4.1.

In addition, unblinded DMC results (including demographics, disposition, and safety summary) and possibly related additional exploratory analyses may also be provided to support compound development.

Unblinding for interim analyses:

The descriptive statistics and other exploratory analyses on body weight, HbA1c and FPG, as well as unblinded DMC report will be performed by an internal sponsor team, independent of the study

team. Body weight modeling and simulation will be performed by a sponsor internal modeling and simulation team providing body weight extrapolation to 6-months and 12-months for obese T2DM and obese patients as compared to liraglutide in obese patients supporting project planning. The unblinded study data will be kept on a restricted access folder, only independent programmers who will perform analyses on body weight, HbA1c and FPG, and provide unblined DMC results and related exploratory analyses, or perform data preparation for body weight modeling and simulation will have access to the folder. Only those necessary for the conducting the analyses and those responsible for internal project planning/overall portfolio planning needs (eg, to aid in the planning of future studies) will have the access to the interim analyses results before study completion. A list of these individuals will be maintained.

All other PK/PD analyses (besides body weight modeling and simulation) will be performed by a sponsor internal modeling and simulation team, independent of the study team. Only those necessary for conducting the analyses/building the model(s) will have the access to the unblinding information and analyses results before study completion and database lock. The purpose of the PK/PD analyses is model establishment to accelerate the final PK/PD analyses after study completion and database lock. These PK/PD analyses results will not be shared with personnel outside of the modeling and simulation team before study completion and database lock.

All sponsor internal personnel with access to unblinding information will be asked to sign a study confidentiality agreement before having access to unblinding information. Study team (including PKDM lead) and investigational sites will not have access to interim study results, and continue to be blinded to individual randomization codes except for the open-label liraglutide until after study completion and database lock. Therefore the two interim analyses will not lead to changes in the conduct of the protocol.

DMC:

In addition, an independent Data Monitoring Committee (DMC) will be used to monitor and assess the safety of patients from this trial through periodic review of the accumulated safety data provided by an independent statistical group. Related details are provided in separate documents (DMC charter and DMC SAP).

4 DATABASE LOCK

The database is planned to be locked approximately 4 weeks after the last patient last visit.

5 SOFTWARE DOCUMENTATION

All summaries and statistical analyses will be generated using SAS version 9.4 or higher.

6 REFERENCES

1. Patrick DL, Bushnell DM, Rothman M. Performance of Two Self-Report Measures for Evaluating Obesity and Weight Loss. Obes Res. 2004;12(1):48-57.

2. Hayes RP, Nelson DR, Meldahl ML, Curtis BH. Ability to perform daily physical activities in individuals with type 2 diabetes and moderate obesity: a preliminary validation of the Impact of Weight on Activities of Daily Living Questionnaire. Diabetes Technol Ther. 2011;13:705-12.

3. Hayes RP, Schultz EM, Naegeli AN, Curtis BH. Test-retest, responsiveness, and minimal important change of the ability to perform physical activities of daily living questionnaire in individuals with type 2 diabetes and obesity. Diabetes Technol Ther. 2012;14:1118-25

4. Hayes RP, DeLozier AM. Reliability, validity, and responsiveness of the Impact of Weight on Self-Perceptions Questionnaire (IW-SP) in individuals with type 2 diabetes and obesity. Diabetes Technol Ther. 2015;17:210-4.

5. Tang YQ. An efficient multiple imputation algorithm for control-based and delta-adjusted patternmixture models using SAS.

6. Ratitch B, O'Kelly M. Implementation of Pattern-Mixture Models Using Standard SAS/STAT Procedures. PharmaSUG2011 - Paper SP04.

7. Yuan Y. Sensitivity Analysis in Multiple Imputation for Missing Data. SAS270-2014.