STATISTICAL ANALYSIS PLAN

A randomized, open-label, parallel group real world pragmatic trial to assess the clinical and health outcomes of Toujeo® compared to commercially available basal insulins for initiation of therapy in insulin naïve patients with uncontrolled type 2 diabetes mellitus

HOE901-LPS14347

STATISTICIAN: ████████████ / ████████████

Statistical Project Leader: ██████████ / ████████████

DATE OF ISSUE: 11-Nov-2016

NCT Number: NCT02451137

Total number of pages: 66

(electronic 1.0)

TABLE OF CONTENTS

(electronic 1.0)

LIST OF TABLES

(electronic 1.0)

LIST OF FIGURES

(electronic 1.0)

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

ADA: American Diabetes Association
AE: adverse event
ALP: alkaline phosphatase
ALT: alanine transaminase
ANCOVA: analysis of covariance
AST: aspartate aminotransferase
ATC: anatomic or therapeutic category
BG: blood glucose
BMI: body mass index
CI: confidence interval
CLcr: creatinine clearance
CRF: case report form
DMC: data monitoring committee
DTSQc: diabetes treatment satisfaction questionnaire change version
DTSQs: diabetes treatment satisfaction questionnaire status version
ECG: electrocardiogram
e-CRF: electronic case report form
EOT: end of treatment
ES: effect size
FPG: fasting plasma glucose
GES: global effectiveness scale
GLP-1:   glucagon like peptide-1
HbA1c: hemoglobin A1c
HEDIS: healthcare effectiveness data and information set
HLGT: high level group term
HLT: high level term
IA: interim analysis
IFCC: international federation of clinical chemistry
IMP: investigational medicinal product
IRT: interactive response technology
ITT: intent-to-treat
IVRS: interactive voice response system
IWRS: interactive web response system
KM: Kaplan-Meier
LBGE: low blood glucose event
LFT: liver function test
LLT: lower level term
LS means: least square means
MAR: missing at random
MCID: minimal clinically important difference
MCMC: Markov Chain Monte Carlo

MedDRA:	medical dictionary for regulatory activities
MMRM:	mixed-effect model with repeated measure
MPR:	medication possession ratio
NIMP:	non-investigational medical product
OAD:	oral anti-diabetic agent
PAM:	patient activation measure
PRO:	patient reported outcome
PSP:	patient support program
PT:	preferred term
RA:	receptor agonist
SAE:	serious adverse event
SD:	standard deviation
SE:	standard error
SIDES:	subgroup identification based on difference effect search
SMPG:	self-measured plasma glucose
SMQ:	standardized MedDRA query
SOC:	system organ class
SU:	sulfonylurea
T2DM:	type 2 diabetes mellitus
TEAE:	treatment-emergent adverse event
ULN:	upper limit of normal
WHO-DD:	World Health Organization-Drug Dictionary

1   OVERVIEW AND INVESTIGATIONAL PLAN

1.1   STUDY DESIGN AND RANDOMIZATION

This is a multi-center open label randomized active controlled 2 arm parallel group comparative "real world" or "pragmatic" trial in patients with type 2 diabetes (T2DM) who are insulin naïve and uncontrolled on two or more oral anti-diabetic agents (OADs) and/or glucagon like peptide-1 (GLP-1) receptor agonist (RA). The goal of the study is to demonstrate that Toujeo HOE901 U300 insulin glargine in combination with its patient support program (PSP) are superior to other commercially available basal insulin, specifically insulin glargine (Lantus) and insulin detemir (Levemir), in attainment of individualized Healthcare Effectiveness Data and Information Set (HEDIS) glycosylated hemoglobin (HbA1c) targets at 6 months without documented symptomatic (blood glucose [BG] ≤70 mg/dL) or severe hypoglycemia at any time of day from baseline to 6 months. The protocol (1) allows for nearly all OADs and GLP-1 RA as background medication as allowed for use with insulin in the label.

The trial will consist of:

- One week screening period at site by clinician to screen the identified potential subjects,

- Twenty-six week treatment period for primary efficacy endpoint,

- Twenty-six week extension period for collecting persistence data and additional efficacy and safety data.

Table 1 - Visit schedule

Visit	Comment	Time	Timeframe
V1	Screening visit	Week -1	
V2	Baseline Visit	Week 0, Day 1	±7 days
V3	Primary efficacy endpoint visit	Day 180	Should occur within ±30 days
V4	End of treatment visit (EOT)	Day 360	May occur within ±30 days

ABBREVIATION: EOT = end of treatment.

Randomization will be performed centrally on V2 (Day 1) by interactive voice response system/interactive web response system (IVRS/IWRS), and will be stratified by individualized HbA1c target (<8%/<7%), GLP-1 RA use (y/n), sulfonylurea (SU) use (y/n) and HbA1c (<9%/≥9%). Patients receive either Sanofi's Toujeo as the test drug or Lantus or Levemir as the comparator/control drugs with or without a PSP.

The maximum study duration is 53 weeks per patient. For patients who prematurely discontinue the trial, the EOT visit assessments are performed at the time of discontinuation. In addition, an early end of trial case report form (CRF) page will be completed which includes the reason for stopping the treatment.

Toujeo SoloSTAR pen and the control drugs, Lantus SoloSTAR and Levemir FlexTouch are distinguishable and so this study is an open-label design. Despite the open label administration of the study insulin, the assessments of outcome, the HbA1c and fasting plasma glucose (FPG) are determined in central laboratories blinded to the treatment received. The study team will review the data for the efficacy parameters, adverse events (AE) and hypoglycemia without treatment assignments.

1.2   OBJECTIVES

1.2.1 Primary objectives

The primary objective of this study is to demonstrate clinical benefit of Toujeo in achieving individualized HEDIS HbA1c targets as defined below at 6 months without documented symptomatic (BG ≤70 mg/dL) or severe hypoglycemia at any time of day from baseline to 6 months in uncontrolled insulin naive patients with type 2 diabetes initiating basal insulin therapy in a real world setting.

The individualized HEDIS HbA1c targets for this study are:
- HbA1c <8% if the patient is ≥65 years,
- HbA1c <8% if the patient is <65 years, and with evidence of any of these diseases/conditions:
    - Coronary artery bypass surgery or Percutaneous coronary intervention,
    - Ischemic vascular disease,
    - Thoracic aortic aneurysm,
    - Chronic heart failure,
    - Prior myocardial infarction,
    - Chronic renal failure/End stage renal disease,
    - Dementia,
    - Blindness,
    - Lower extremity amputation,
- HbA1c <7% if the patient is <65 years, and without evidence of any of these diseases/conditions listed above.

1.2.2 Secondary objectives

Secondary objectives are to compare Toujeo to other commercially available basal insulins at 6 and 12 months after initiating basal insulin therapy in a real world setting in terms of:
- Patient persistence with assigned basal insulin therapy,
- Risk of hypoglycemia including the incidence and rate of documented symptomatic and severe hypoglycemia,
- Changes in HbA1c, FPG, body weight,

- Differences in patient and provider reported outcomes (including Diabetes Treatment Satisfaction Questionnaire Status and Change Versions [DTSQs and DTSQc], Hypoglycemia Patient Questionnaire, and patient and provider reported Global Effectiveness Scales [GES]),

- Healthcare resource utilization including hospitalizations and emergency department or other provider visits and healthcare costs.

1.3   DETERMINATION OF SAMPLE SIZE

1.3.1   Original sample size

At time of initial protocol, calculations to determine the sample size of 1635 patients per treatment group are based on superiority testing of the primary efficacy variable of the proportion of patients reaching individualized HbA1c target at Month 6 without a documented symptomatic (BG ≤70 mg/dL) or severe hypoglycemic event at any time of day from baseline to Month 6, with the following assumptions:

- Individualized HbA1c target attainment per HEDIS criteria of <8% if older or presence of medical comorbidities, or otherwise <7%,

- An absolute difference of 5.5% in favor of Toujeo in patients assigned a target HbA1c <7%, assuming a 30.1% rate in the Toujeo arm, based on data from the EDITION 3 study,

- An absolute difference of 5.4% in favor of Toujeo in patients assigned a target HbA1c <8%, assuming a 55.8% rate in the Toujeo arm, based on data from the EDITION 3 study,

- An absolute overall (weighted average) difference of 5.5% in favor of Toujeo for all patients combined, assuming that 70% of the study patients would have a target HbA1c <7%, and 30% would have a target HbA1c <8%, based on data from a commercially available database,

- A maximum overall response rate of 33% in the comparator basal insulins arm,

- A 2-sided chi-squared test at the 5% significance level, with 90% power,

- All patients will be eligible for the primary endpoint analysis.

The assumptions are based on data from the EDITION program.

As a logistic regression model will be used for the primary analysis of the composite endpoint of individualized HbA1c target attainment without hypoglycemia, it is anticipated that the power will be higher due to reduced estimate variability compared with the chi-square test. Calculations were made using nQuery Advisor 6.01.

1.3.2 Sample size for Interim Analysis

One interim analysis (IA) was added per Clinical Trial Protocol Amendment No. 01, for the purpose of efficacy to provide an early evaluation of Toujeo effectiveness, when 1800 randomized patients (54.2% information fraction of targeted sample size) have completed their 6-month (Day 180) visit.

Due to the addition of the IA the sample size will be increased to 1662 patients per treatment group (total of 3324 patients). The calculation takes into account one IA using a group sequential approach with an efficacy boundary based on a gamma (-3) alpha spending function, and an overall two-sided alpha-level of 0.05, the two-sided nominal significance level is 0.010 at IA, and 0.046 at final analysis.

The total sample size calculation of 3324 is based on the assumption of a maximum of a 33% overall (weighted average) response rate with Lantus, and at least a 5.5% increase in response rate with Toujeo, and assumes an overall 2-sided test with alpha-level 0.05 with an IA performed on 1800 patients followed 6 months.

Overwhelming efficacy will be shown at IA if the efficacy boundary is met, ie, if the two-sided p-value for comparison of primary composite endpoint incidence is below 0.010, with a difference in favor of Toujeo between percentages of patients reaching the primary composite endpoint. The corresponding power to show overwhelming efficacy at IA under the alternative hypothesis is of 45%.

1.4 STUDY PLAN

The study will be a randomized, controlled, open label parallel group real world trial.

It will include a 6 month treatment period ending in assessment of the primary and secondary endpoints followed by a 6 month extension period for assessment of certain additional endpoints.

There will be three database locks:
- One at interim 6-month analysis,
- One at 6 months, once the last randomized patient has had an opportunity to complete 6 months of treatment,
- And one at 12 months.

Figure 1 - Graphical study design
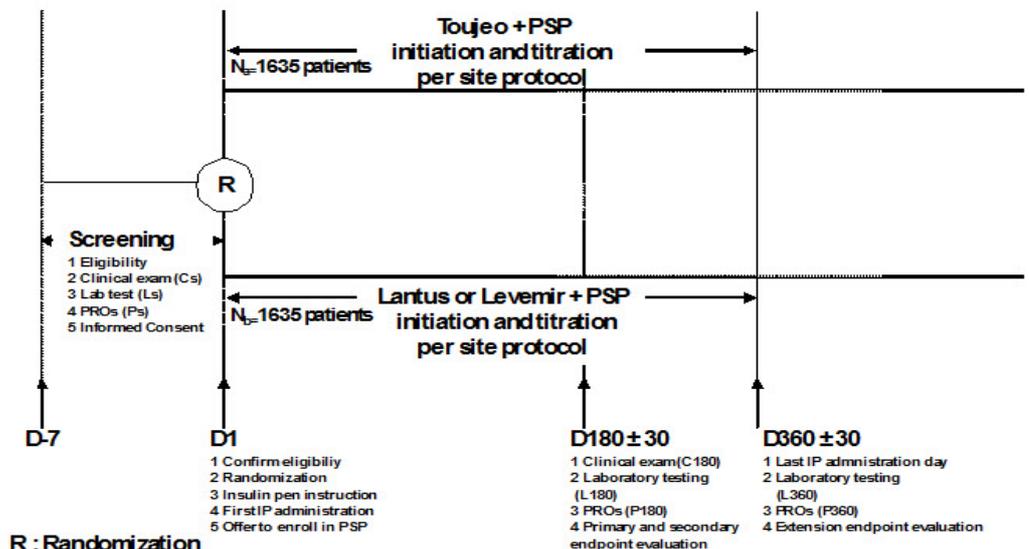
R : Randomization

Cs and C180 : Clinical exam including vital signs: weight, height, systolic and diastolic blood pressure, heart rate as well as comprehensive physical exam

Ls: Hemoglobin A1c, fasting plasma glucose, serum creatinine, serum liver function tests including AST, ALT, total bilirubin, and alkaline phosphatase, serum pregnancy test for women of childbearing potential

L180: Hemoglobin A1c, fasting plasma glucose, serum pregnancy test for women of childbearing potential

L360: Hemoglobin A1c, fasting plasma glucose, serum pregnancy test for women of childbearing potential

Ps: DTSQ, Hypoglycemia Patient Questionnaire

P180 and P360: DTSQ, Hypoglycemia Patient Questionnaire, Patient GES, Provider GES

Table 2 - Study flow chart

	Screening	Treatment Period			
VISIT	**1**	**2**		**3**	**4**
WEEK	**-1**	**0**		**26 ±4**	**52 ±4**
DAY	**-7**	**1 ±7**	**45**	**90**	**180 ±30**	**360 ±30**
Informed Consent	X				
Inclusion/Exclusion Criteria	X				
Patient Demography	X				
Medical/Surgical History	X				
Prior Medication History	X				
Physical Examination	X			X	X
Vital Signs (including body weight)	X			X	X
DTSQs	X	X		X	X
DTSQc					X
Hypoglycemia Patient Questionnaire	X	X		X	X
Patient activation measure (PAM)	X				
Patient: GES				X	X
Provider: GES				X	X
Randomization		X			
IVRS call	X	X		X	X

		Screening	Treatment Period			
	VISIT	**1**	**2**		**3**	**4**
	WEEK	**-1**	**0**		**26 ±4**	**52 ±4**
	DAY	**-7**	**1 ±7**	**45**	**90**	**180 ±30**	**360 ±30**
Treatment:	Dispensation of investigational medicinal product (IMP)		X				
	Concomitant Medication	X	X			X	X
	Compliance					X	X
	Insulin Pen Instruction		X				
	BG meter dispensing and training		X				
	E-diary dispensing and training		X				
	Offer to Enroll in PSP		X				
	Insulin dose					X	X
Safety:	AE/serious adverse event (SAE) recording	X	X			X	X
	Diary Reviews at visits		X			X	X
	Hypoglycemia review of electronic source [a]			X	X		
Laboratory Testing:	HbA1c	X				X	X
	FPG	X				X	X
	Serum creatinine	X					
	Serum liver function tests (LFT) (aspartate aminotransferase [AST], alanine aminotransferase [ALT], total bilirubin, alkaline phosphatase [ALP])	X					
	Pregnancy test (for women of childbearing potential) [b]	X				X	X

a   Approximately Day 45 and approximately Day 90, review of electronic source without on site or telephone visit,

b   Pregnancy status should be checked by serum pregnancy testing prior to exposure to the investigational product, urine pregnancy tests may be done at subsequent visits.

ABBREVIATION: DTSQs = diabetes treatment satisfaction questionnaire status version; DTSQc = diabetes treatment satisfaction questionnaire change version; PAM = patient activation measure; GES = global effectiveness scale; IVRS = interactive voice response system; IMP = investigational medicinal product; BG = blood glucose; PSP = patient support program; SAE = serious adverse event; FPG = fasting plasma glucose; LFT = liver function test; AST = aspartate aminotransferase; ALT = alanine aminotransferase; ALP = alkaline phosphatase; HbA1c = hemoglobin A1c

1.5   MODIFICATIONS TO THE STATISTICAL SECTION OF THE PROTOCOL

In the protocol (1) Section 11.4.2, Analyses of efficacy endpoints, defines the analysis of HbA1c endpoints (change from baseline to 6 months) and self-reported satisfaction score using a mixed-effect model with repeated measures (MMRM) approach under the missing at random (MAR) framework, using an adequate contrast at Month 6. Since at the time of the 6 month database lock, only baseline and 6 month HbA1c data are available, an analysis of covariance (ANCOVA) will be performed at time of 6 month database lock and a MMRM at time of 12 month database lock.

In the protocol (1) Section 11.4.2, Analysis of primary efficacy endpoint(s), exploratory analysis for primary endpoint was modified. The Predictive Enrichment method is replaced by Subgroup Identification Based on Difference Effect Search (SIDES). The rationale is that the predictive enrichment procedure does not adjust for multiplicity. Additionally, the definition of subgroups is based on a score, defined as a function of multiple covariates, which makes difficult the clinical interpretation. On the other side, SIDES method uses fewer covariates (ie, 2 or 3 covariates), that is easier to interpret from a clinical point of view. Additionally, SIDES corrects for multiplicity.

As mentioned in Clinical Trial Protocol Amendment No. 01, an IA is planned when 1800 patients have completed their 6 month visit. The calculation takes into account one IA using a group sequential approach with an efficacy boundary based on a gamma (-3) alpha spending function.

The cut-off date for the IA is the date when 1800 patients have completed their 6 month visit.

1.6   STATISTICAL MODIFICATIONS MADE IN THE STATISTICAL ANALYSIS PLAN

Not applicable.

2 STATISTICAL AND ANALYTICAL PROCEDURES

2.1 ANALYSIS ENDPOINTS

2.1.1 Demographic and baseline characteristics

The efficacy baseline value is defined as the last available value obtained before and up to the date of randomization.

Baseline HbA1c will be obtained at screening visit within one week of study initiation. For all safety measures, the baseline value is defined as the last available value prior to the first injection of IMP.

Demographic characteristics

Demographic variables are,
- Age in years,
- Age groups (<65, ≥65 and <75, and ≥75 years),
- Gender (Male, Female),
- Race (Caucasian/white, Black, Asian/Oriental, American Indian or Alaska Native, Native Hawaiian or other Pacific Island, other), and,
- Ethnicity (Hispanic, non-Hispanic).

Medical or surgical history

Medical or surgical history include medical history or surgical history related to diabetes and other medical or surgical history collected at Screening.

Medical history or surgical history related to diabetes are pre-specified medical or surgical history related to diabetes.

Other medical or surgical history will be coded to "Lower Level Term (LLT)", "Preferred Term (PT)", "High Level term (HLT)", "High Level Group Term (HLGT)" and associated primary "System Organ Class (SOC)" by the version of Medical Dictionary for Regulatory Activities (MedDRA) currently in effect at the sponsor at the time of database lock.

Disease characteristics at baseline

Specific disease history includes history of diabetes collected at Screening and HEDIS components.

Diabetes history includes,
- Duration of type 2 diabetes (date of diagnosis),
- Age at diagnosis of diabetes,
- Diabetes duration (years),

- Diabetic complications, (diabetic retinopathy, diabetic sensory or motor neuropathy, diabetic autonomic neuropathy, and diabetic nephropathy),
- Anti-diabetic non-insulinic drugs,
- Previous participation to a PSP, and,
- PAM.

Note: PAM as an assessment of patient engagement. PAM score is detailed in Appendix F of Achieve Control protocol (1).

Vital signs

Vital signs at baseline are,
- Weight in kilograms,
- Weight group (<50, ≥50 and <100, and ≥100 kg),
- Height in centimeter,
- Body mass index (BMI) in kg/m^2,
- Category of BMI (<25, [25-30], [30-40] and ≥40 kg/m²),
- Blood pressures, and,
- Heart rate.

Other baseline characteristics

Other baseline characteristics include,
- Randomization strata Individualized HbA1c target (<8%/<7%) from IVRS and electronic CRF (e-CRF) data,
- Randomization strata SU use (yes/no) from IVRS and e-CRF data,
- Randomization strata GLP-1 RA use (yes/no) from IVRS and e-CRF data,
- Payer or non-payer patients, defined as patients who have (or not) provided consent to release of insurance data,
- Geographical region of the sites, and,
- Physician type (endocrinologist, internal medicine or primary care physician).

Baseline efficacy data
- HbA1c (% and mmol/mol),
- Baseline HbA1c category (<9%/≥9%), and,
- FPG (mmol/L and mg/dL).

Baseline safety laboratory data

Clinical chemistry data for centralized clinical laboratories will be described after conversion into standard international units.

Renal function
- Baseline Creatinine (µmol/L),
- Baseline Creatinine Clearance : CLcr (mL/min),
- CLcr category (mL/min): ≥90, [60–90], [30–60], <30.

Pancreatic function
- Lipase (U/L and ULN ratio),
- Amylase (U/L and ULN ratio).

Alcohol and smoking habit
- Tobacco habit (never smoke, quit smoking, currently smokes) including the average number of cigarettes per day for currently smokers,
- Alcohol habit (never, occasionally, at least monthly, at least weekly, at least daily) including the number of daily standard drink (1 or 2, >2).

2.1.2 Prior and Concomitant medications

Prior medications are those the patient used prior to the first IMP intake. Prior medications can be discontinued before first administration or can be ongoing during treatment phase.

Concomitant medication is any treatment received by the patient at the same time as the IMP treatment, from 1st IMP injection up to last IMP injection + 1 day (0 day for anti-diabetic therapy) whichever comes earlier.

All medications will be coded using the most current version of World Health Organization-Drug Dictionary (WHO-DD) at the time of each database lock.

Prohibited concomitant medication is limited to drugs contraindicated with use of basal insulin as per drug labels.

2.1.3 Efficacy endpoints

HbA1c and FPG will be measured by central laboratory. Blood sample taken used for analyses are those collected at baseline, Month 6, Month 12 or premature EOT. Analysis time point will be either the nominal visit or the reallocated visit (reallocation process is described in Section 2.5.4).

For patients randomized and not treated, the baseline value is defined as the last available value obtained up to the date and time of randomization.

Observation period of efficacy variables
- The **12-month randomized period** is defined as the time from randomization up to Day 360 (Visit 4). It includes the extension period (after Day 180 [Visit 3] to Day 360 [Visit 4]) to allow for collection of additional endpoint data such as health care utilization[2].

- The **6-month randomized period** is defined as the time from randomization up to Day 180 (Visit 3) or date of end of study whichever come earlier[1].

- The **6-month on-treatment period** is defined as the time from 1st IMP intake up to Day 180 (Visit 3) or date of last IMP + 1 day (7 days for HbA1c), whichever comes earlier[1].

- The **6 to 12-month randomized period** named "from start of Month 7 to Month 12" is defined as the time from the day following the 6-month randomized treatment period up to the end of the 12-month randomized period[2].

1. Used at time of 6 month database lock,
2. Used at time of 12 month database lock.

2.1.3.1 Primary efficacy endpoint(s)

The primary efficacy endpoint is the proportion of patients with individualized HbA1c target attainment per HEDIS criteria assessed at randomization without documented symptomatic (BG ≤70 mg/dL) or severe hypoglycemia at any time of day from baseline to 6 months.

The individual components of individualized HbA1c target attainment and absence of any documented symptomatic (BG ≤70 mg/dL) or severe hypoglycemia will also be evaluated as supportive analyses to evaluate consistency:

- Proportion of patients with individualized HbA1c target attainment at 6 months, and,
- Proportion of patients without documented symptomatic (BG ≤70 mg/dL) or severe hypoglycemia.

2.1.3.2 Secondary efficacy endpoint(s)

- Change in HbA1c from baseline (obtained at screening visit within one week of study initiation) to 6 months,
- Proportion of patients who remain on assigned basal insulin therapy (persistent with assigned therapy) at 6 and 12 months,
- Proportion of patients with individualized HbA1c target attainment per HEDIS criteria at 6 months without documented symptomatic (BG <54 mg/dL) or severe hypoglycemia at any time of day from baseline to 6 months,
- Proportion of patients with individualized HbA1c target attainment per HEDIS criteria at 12 months without documented symptomatic (BG <54 mg/dL) or severe hypoglycemia at any time of day from baseline to 12 months, and,
- Proportion of patients with individualized HbA1c target attainment (per HEDIS criteria) at 12 months without documented symptomatic (BG ≤70 mg/dL) or severe hypoglycemia at any time of day from baseline to 12 months.

2.1.3.3 Other efficacy endpoint(s)

- Change in HbA1c from baseline to 12 months,
- Change in FPG from baseline to 6 months and 12 months,
- Change in body weight from baseline to 6 months and 12 months,

- Basal insulin dose collected at the study visit on the e-CRF at 6 and 12 months,
- Proportion of patients with HbA1c <8% at 6 months without documented symptomatic (BG ≤70 mg/dL) or severe hypoglycemia,
- Proportion of patients with HbA1c <7% at 6 months without documented symptomatic (BG ≤70 mg/dL) or severe hypoglycemia,
- Proportion of patients with HbA1c <8% at 6 months, and,
- Proportion of patients with HbA1c <7% at 6 months.

2.1.4 Safety endpoints

The following safety parameters will be analyzed in the clinical study report:

- Hypoglycemia (according to American Diabetes Association [ADA] Workgroup on Hypoglycemia) (2),
- AEs, SAEs,
- Injection site reactions,
- Hypersensitivity reactions,
- Vital signs, and,
- Healthcare resource utilization including hospitalizations and emergency department or other provider visits and healthcare costs.

Observation period for safety endpoints

- The **pre-treatment period** is defined as the time from informed consent up to the time of the first injection of IMP.
- The **12-month on-treatment period** is defined as the time from the 1st IMP intake up to 1 day after the trial is completed[2].
- The **6-month on-treatment period** named "Month 1 to Month 6" is defined as the time from 1st IMP intake up to Day 180 (Visit 3) or 1 day after date of last insulin (either IMP or commercial basal insulin prescribed following a switch) or 6-month visit, whichever comes earlier[3].
- The **6 to 12-month on-treatment period** named "from start of Month 7 to Month 12" is defined as the time from the day following the 6-month on-treatment period up to the end of the 12-month on-treatment period[2].
- The **post-treatment period** is defined as the time starting 2 days after last injection of IMP (after the on-treatment period).
- The **on-study observation period** is defined as the time from randomization until the end of the study (defined as last protocol planned visit or the resolution/stabilization of all SAEs and AE with pre-specified monitoring).

2. Used at time of 12 month database lock,
3. Used for 6 month and 12 month database lock.

2.1.4.1 Hypoglycemia variables

Hypoglycemia observation periods
- Pre-treatment hypoglycemia events are events that occur during the pre-treatment period,
- Treatment-emergent hypoglycemia events are events that occur during the 12-month on-treatment period,
  - The 6-month treatment-emergent hypoglycemia events are events that occur during the 6-month on-treatment period,
- Post-treatment hypoglycemia events are events that occur during the post-treatment period.

Hypoglycemia endpoints will be classified in categories after events have been characterized using ADA Working group (2). More specifically, the following variables will be analyzed:
- Incidence and rate of documented symptomatic (BG ≤70 mg/dL) or severe hypoglycemia (24 hour and nocturnal) from baseline to 6 and 12 months,
- Incidence and rate of documented symptomatic (BG <54 mg/dL) or severe hypoglycemia (24 hour and nocturnal) from baseline to 6 and 12 months, and,
- Incidence and rate of severe hypoglycemia (24 hour, per ADA definition) from baseline to 6 and 12 months.

Nocturnal hypoglycemia is defined as any hypoglycemia of the above categories that occurs between $00:00^{AM}$ and $05:59^{AM}$ hours, regardless whether patient was awake or woke up because of the event. Daytime hypoglycemia is defined as any hypoglycemia of the above categories that occurs between $6:00^{AM}$ to 23:59.

For each class of hypoglycemia as defined in Appendix A, documented hypoglycemia will be analyzed in addition to the threshold of plasma glucose of less than or equal to 70 mg/dL (3.9 mmol/L), by using a threshold of plasma glucose of <54 mg/dL (3.0 mmol/L).

Patients are dispensed:
- A Bluetooth enabled study glucometer,
- E-diary,
- And Low Blood Glucose Event (LBGE) paper back-up page.

For all self-measured plasma glucose (SMPG) equal to or less than 70 mg/dL (or equivalent) LBGE symptoms, patients are instructed to complete:
- A hypoglycemia page in the e-diary,
- Or paper back-up page.

Hypoglycemic events will be transferred from the glucometer to a web based portal that serves as an electronic source document. Some hypoglycemic events may be recovered from paper source such as the hypoglycemia back up report or the patient paper diary. In all cases, the source will be reviewed by the site and manually entered into the CRF.

2.1.4.2 Adverse events variables

All reported AEs and SAEs is collected from the time of signed informed consent until the end of the study (defined as patient´s last visit planned per protocol or the resolution/stabilization of all SAE and AEs with pre-specified monitoring). All events will be managed and reported in compliance with all applicable regulations, and included in the final clinical study report.

All AEs and SAEs will be coded to a "LLT", "PT", "HLT", "HLGT" and associated primary "SOC" using MedDRA. MedDRA terms for hypersensitivity and injection site reactions will be included. The occurrence of AEs and SAEs is recorded from the time of signed informed consent until the end of the study.

If an AE/SAE is ongoing and not resolved by the end of the study observation period and leads to subsequent death of the patient, it will have to be reported.

The AE observations will be classified per the observation periods of safety data as defined above into:

- **Pre-treatment AEs** are AEs that developed or worsened or became serious during the pre-treatment period.
- **Treatment-emergent AEs (TEAE)** are AEs that developed or worsened or became serious during the 12-month on-treatment period.
  - The **6-month TEAEs** are AEs that developed or worsened or became serious during the 6 month on-treatment period.
    to be used at time of 6-month database lock only.
- **Post-treatment AEs** are AEs that developed or worsened or became serious during the post-treatment period.

Symptomatic overdose

Symptomatic overdose with IMP/anti-hyperglycemic non-investigational medicinal product (NIMP) will be identified through a specific question on the AE page in the e-CRF.

Pregnancy

Pregnancy occurring, either in female subject or in a female partner of a male subject, entered in a study will be identified through a dedicated AE page in the e-CRF.

2.1.4.3 Injection site reactions, hypersensitivity reactions and immunogenicity

No immunogenicity assessment is planned as part of the protocol (1).

Injection site and hypersensitivity reactions are AEs recorded on AE page and identified for the statistical analysis using the following MedDRA codes:

- Injection site reactions will be identified using the following MedDRA searches: HLGT "Administration site reactions" and HLTs "Administration site reactions NEC", "Infusion site reactions", "Injection site reactions" and "Application and instillation site reactions" and excluding HLTs "Implant and catheter site reactions" and "Vaccination site reactions".

- Hypersensitivity reactions will be identified using the following MedDRA searches: Angioedema standardized MedDRA query (SMQ) (Narrow), Severe cutaneous adverse reactions SMQ (Broad), Hypersensitivity SMQ (Broad and Narrow) and excluding PTs related to administration, application, injection and infusion sites. High level term "Anaphylactic Responses" are included in those SMQs.

2.1.4.4 Deaths

The death observations are per the observation periods defined above. In addition, after the post-treatment period, death related to IMP as well as death resulting from TEAE must be reported to the Sponsor.

- Death on-study: deaths occurring during the on-study observation period (defined as the time from the signed informed consent date up to the end of the study, ie, last protocol planned visit or the resolution/stabilization of all SAE),

- Death pre-treatment: deaths occurring before the on-study observation period,

- Death on-treatment: deaths occurring during the on-treatment period,

- Death post study: death occurring after the end of study.

2.1.4.5 Laboratory safety variables

The clinical laboratory data consist of blood analysis (including, LFTs namely AST, ALT, ALP, total bilirubin, ALP, and creatinine) at Screening. They are not considered as part of safety endpoints to evaluate the study objectives during the treatment period.

2.1.4.6 Vital signs variables

Vital signs include: observed and change from baseline of heart rate (bpm), systolic and diastolic blood pressures (mmHg), and weight (kg) at baseline, 6 months and 12 months.

2.1.4.7 Electrocardiogram variables

Electrocardiogram (ECG) is not done routinely in the protocol (1). It may be done at the discretion of the investigator when clinically indicated. This data is not routinely recorded in the e-CRF and if abnormal, should be recorded as an AE.

2.1.5 Pharmacokinetic variables

Not applicable.

2.1.6 Pharmacodynamic/genomics endpoints

Not applicable.

2.1.7 Quality-of-life endpoints and patient reported outcomes

Patients will complete self-reported questionnaires at designated time points in the study.

2.1.7.1 Diabetes Treatment Satisfaction Questionnaire Status Version

The DTSQs is a validated questionnaire to assess patient satisfaction with treatment and patient perception of BG control (3). The DTSQs will be used to evaluate patient satisfaction with treatment and patient perception of BG control over a several week period. The DTSQs is comprised of 8 questions which are answered on a Likert scale from 0 to 6. Responses to these questions would be summarized into the three domain scores of Total Treatment Satisfaction (Questions 1, 4-8), Hyperglycemia Perception (Question 2) and Hypoglycemia Perception (Questions 3) such that a higher score would be indicative of better satisfaction. Mean differences in scores between groups would be evaluated to understand the impact of treatment on patient satisfaction. DTSQs will be evaluated at Screening, Day 1, 6 months and 12 months in this study.

The total treatment satisfaction score and the 2 other scores are determined by computing the mean of item responses:
- Total treatment satisfaction score: sum of Items 1, 4, 5, 6, 7 and 8, ranging from 0 (no satisfaction) to 36 (high satisfaction),
- Perceived frequency of hyperglycemia score: Item 2, ranging from 0 (none of the time) to 6 (most of time),
- Perceived frequency of hypoglycemia score: Item 3, ranging from 0 (none of the time) to 6 (most of time).

The DTSQs endpoints are:
- Average scores on total treatment satisfaction from DTSQs at 6 months and 12 months,
- Average scores on hyperglycemia perception from DTSQs at 6 months and 12 months; and,
- Average scores on hypoglycemia perception from DTSQs at 6 months and 12 months.

Additionally, on these 3 scores (total treatment satisfaction, hyperglycemia perception and hypoglycemia perception), evaluation will be performed on:
- Change from baseline to 6 months,
- Change from baseline to 12 months,
- Change between 6 and 12 months.

Additional endpoints are:
- Percentage of patient responders on total treatment satisfaction score, defined as patients presenting a change from baseline to respectively Month 6 and Month 12 of at least the distribution-based minimal clinically important difference (MCID) value (Section 2.5.1).
- Percentage of patient responders on perceived frequency of hyperglycemia score, defined as patients presenting a change from baseline to respectively Month 6 and Month 12 of at least the standard MCID value (Section 2.5.1).

- Percentage of patient responders on perceived frequency of hypoglycemia score, presenting a change from baseline to respectively Month 6 and Month 12 of at least the standard MCID (Section 2.5.1).

2.1.7.2 Diabetes Treatment Satisfaction Questionnaire Change Version

The DTSQc was developed from the original DTSQ to evaluate the change in treatment satisfaction at a specific time point (4).

The DTSQc will be used to produce a measure of relative change in satisfaction rather than measure of absolute satisfaction, which is subject to ceiling and floor effects.

A score for DTSQc treatment satisfaction items (Questions 1, 4-8) range from −3 to +3, and the sum of the treatment satisfaction scores range from −18 to +18.

Positive scores are indicative of increases in treatment satisfaction relative to one year ago (start of the study).

Perceived hyperglycemia (Question 2) and hypoglycemia (Question 3) items have scores that range from −3 to +3 in the DTSQc, with higher scores indicating more frequent perceived hyperglycemia or hypoglycemia by the end of the study.

To control for baseline scores, DTSQs will be used as a measure of treatment satisfaction at baseline. DTSQc will be evaluated at 12 months in this study.

The total treatment satisfaction score and the 2 other scores are determined by computing the mean of item responses:
- Total treatment satisfaction score: sum of Items 1, 4, 5, 6, 7 and 8, ranging from -18 (deterioration in treatment satisfaction) to +18 (improvement in treatment satisfaction),
- Perceived frequency of hyperglycemia score: Item 2, ranging from -3 (fewer problems than before with hyperglycemia) to + 3 (more problems),
- Perceived frequency of hypoglycemia score : Item 3, ranging from -3 (fewer problems than before with hypoglycemia) to +3 (more problems).

The DTSQc endpoints are:
- Average scores on change in total treatment satisfaction from DTSQc at 12 months,
- Average scores on change in hyperglycemia perception from DTSQc at 12 months,
- Average scores on change in hypoglycemia perception from DTSQc at 12 months.

2.1.7.3 Hypoglycemia Patient Questionnaire

Patients will also complete the Hypoglycemia Patient Questionnaire, detailed in Appendix H of Achieve Control protocol (1).

This questionnaire will require them to report frequency of hypoglycemia experienced during the course of the study. Responses to Questions 4 and 5 on the Hypoglycemia Patient Questionnaire will be used to assess the endpoints. a patient report of one or more episodes of either moderate or severe hypoglycemia since the last visit will be indicative of patient experience of hypoglycemia.

The Hypoglycemia Patient Questionnaire endpoints are:
* Proportion of patients who report moderate or severe hypoglycemia since last visit on the Hypoglycemia Patient Questionnaire (Questions 4 and 5) at baseline, at 6 months and at 12 months.

2.1.7.4 Global Effectiveness Scale

Patient and Physician reported GES: This measure, originally developed to assess the impact of treatment on asthma control, will be adapted for diabetes (GES). The GES assesses impact of treatment on a 5-point scale:
* Complete control of diabetes,
* Marked improvement of diabetes,
* Discernible, but limited improvement in diabetes,
* No appreciable change in diabetes,
* Worsening of diabetes.

These assessments will be completed twice by patients at 6 and 12 months. Physicians will complete the GES twice for each patient they manage at 6 and 12 months.

The GES endpoints are:
* Proportions of patient and provider reported "Complete control of diabetes" or "Marked improvement of diabetes" responses to GES question at 6 months and 12 months.

2.1.8 Health economic endpoints

The healthcare utilization endpoints are:
* Mean number and proportion of patients with hospitalizations over a 6 month and 12 month period,
* Mean number and proportion of patients with emergency room visits over a 6 month and 12 month period,
* Mean number of provider office visits and proportion of patients with specialty visits over a 6 month and 12 month period,
* Mean overall and diabetes-related costs at 6 months and 12 months.

2.2   DISPOSITION OF PATIENTS

Screened patients are defined as any patient who met the inclusion criteria and signed the informed consent.

All randomized patients will be included in the Intent-to-Treat (ITT) Population.

Patients will not be randomized more than once.

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

- Screened: all patients who originally met inclusion criteria and signed the informed consent[3],
- Screen failure patients and reason for screen failure[3],
- Non-randomized but treated patients[3],
- Randomized: all screened patients with a treatment arm allocated and recorded in the IRT database, regardless of whether IMP was used or not and regardless of use of a PSP program[3],
- Randomized but not treated patients[3],
- Randomized and treated patients[3],
- The **12-month treatment completer population** (presented as randomized): patients who complete the 12-month treatment period (have performed Visit 4 [D360]) without permanently discontinue treatment[2],
- The **6-month-treatment completer population** (presented as randomized): patients who patients who complete the 6-month treatment period (have performed Visit 3 [D180]) without permanently discontinuing treatment[1],
- Patients who permanently discontinued treatment during the 12-month study period and reason for permanent treatment discontinuation presented as randomized[2],
- Patients who permanently discontinued treatment during the 6-month treatment period (before Month 6) and reason for permanent treatment discontinuation presented as randomized[1],
- Patients who permanently discontinued treatment at or after the 6-month treatment period and reason for permanent treatment discontinuation presented as randomized[2],

1.   Used at time of 6 month database lock,
2.   Used at time of 12 month database lock,
3.   Used for 6 month and 12 month database lock.

For all categories of patients (except for the screened and non-randomized categories) percentages will be calculated using the number of randomized patients as the denominator. Reasons for treatment discontinuation will be supplied in tables giving numbers and percentages by treatment group. This summary will be provided by treatment group within each center and will also be further sub-grouped by geographic region.

Kaplan-Meier (KM) plots and estimates of the cumulative incidence of IMP discontinuation based on CRF collected data will be performed.

All critical or major deviations potentially impacting efficacy analyses, randomization, and drug-dispensing irregularities, and other major or critical deviations will be summarized in tables giving numbers and percentages of deviations by treatment group.

Additionally, the analysis populations for safety and efficacy will be summarized in a table by number of patients on the randomized population (Section 2.3).

2.2.1 Randomization and drug dispensing irregularities

On V2 (Day 1), the investigator or designee contacts the IVRS/IWRS and provides the patient number provided by IVRS/IWRS at screening visit, date of birth, HbA1c, individualized HbA1c target (<8%/<7%), GLP-1 RA use (y/n), SU use (y/n) and HbA1c (<9%/≥9%). Afterwards the IVRS/IWRS is called again each time for each new visit but does not include "standard of care" visits, which are not considered trial visits.

Treatment kits are not dispensed since the drugs are provided through a pharmacy. The investigator provides the patient with a prescription for the insulin and a card to provide to the pharmacy to obtain the insulin, a "smart card" as a method of reimbursement.

Randomization will be performed centrally by IVRS/IWRS, and will be stratified by individualized HbA1c target (<8%/<7%), GLP-1 RA use (y/n), SU use (y/n) and HbA1c (<9%/≥9%). a patient cannot be randomized more than once in the study. Patients who need to be moved to another site will have a new randomization number assigned at the new site.

Randomization and drug-dispensing irregularities 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,
2. A patient is dispensed/prescribed an IMP not allocated by the protocol-defined randomization, such as,
   a) A patient at any time in the study is dispensed/prescribed different IMP than as randomized (which may or may not contain the correct-as-randomized IMP), or,
   b) A non-randomized patient is treated with IMP reserved for randomized patients.

Randomization and drug-dispensing irregularities will be monitored throughout the study and reviewed on an ongoing basis.

All randomization and drug-dispensing/prescription irregularities will be documented in the clinical study report. If the number of irregularities is large enough to make a tabular summary useful, the irregularities will be categorized and summarized among randomized patients (number and percentages). Non-randomized, treated patients will be described separately.

Patients with the following deviations will be identified and described in separate listings:

- Treated but not randomized,

- Randomized but not treated,

- Randomized but not following the treatment arm as randomized.

Randomization and drug-dispensing irregularities to be prospectively identified include but are not limited to:

Table 3 - Randomization and drug allocation irregularities

Randomization and drug allocation irregularities
Treatment dispensation without IRT transaction
Erroneous treatment delivery
Treatment not available
Randomization by error
Patient randomized twice
Forced randomization
Stratification error
Patient switched to another site

2.3   ANALYSIS POPULATIONS

Patients treated without being randomized will not be considered randomized and will not be included in any safety or efficacy population.

Patients who are dispensed/prescribed study drug without calling the IRT or before calling the IRT are considered non-randomized patients. They are excluded from any population for analysis, including efficacy and safety. However, if these patients experienced any significant safety event, they should be documented separately in the clinical study report.

The randomized population includes any patient who has signed her/his informed consent and has been allocated to a randomized treatment arm regardless of whether the treatment was delivered or used.

The safety and efficacy populations will be defined at time of 6-month data base-lock based on data collected during the 6 month-randomized period.

2.3.1 Efficacy populations

The efficacy population is the ITT population, defined as all randomized population, irrespective of the treatment actually received, analyzed according to the treatment group allocated by randomization.

2.3.2 Safety population

The safety population is defined as: all randomized patients who actually received at least one dose or part of a dose of IMP, analyzed according to the treatment actually received.

In addition:

- Non-randomized but treated patients will not be part of the safety population. This is a post marketing real world trial with commercially available insulin so that non-randomized but treated patients are not considered part of the trial and safety data are listed separately.

- Randomized patients for whom it is unclear whether they took the study medication will be included in the safety population as randomized.

- For any patient receiving more than one study treatment (Toujeo or "standard of care" basal insulin) during the trial, the patient will be analyzed in the treatment arm in which he/she was treated longer by either IMP (allocated by IRT or prescribed/dispensed) or Commercial basal insulin prescribed following a switch during the 6-month on-treatment period. This actual treatment arm, defined at time of 6-month database lock on the 6-month on-treatment period, will be applied at time of 12-month database lock (ie, to the 12-month on-treatment period).

If a patient is randomized more than once, only the data associated with the first randomization are used in the analysis population. The subject will be analyzed according to the first randomization treatment arm. All data collected after the second randomization will be excluded from the primary efficacy analysis. The safety experience associated with any later randomization is described separately.

2.4 STATISTICAL METHODS

Continuous data will be summarized using the number of available data, mean, standard deviation (SD), median, minimum, and maximum for each treatment group. Categorical and ordinal data will be summarized using the number and percentage of patients in each treatment group.

2.4.1 Demographics and baseline characteristics

Parameters will be summarized on the ITT 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 for any treatment group.

Parameters described in Section 2.1.1 will be summarized by treatment group and overall using descriptive statistics.

Other medical or surgical history not related to diabetes will be summarized by SOC and HLT for each treatment group. Table will be sorted in descending order of the Toujeo treatment arm.

P-values on demographic and baseline characteristic data will not be calculated.

2.4.2 Prior and Concomitant medications

The prior and concomitant medications will be presented for the ITT population, separately for anti-diabetic and non-anti-diabetic medication.

Prior medication will be presented by treatment arm and overall, whereas concomitant and post-treatment medication will be presented by treatment arm.

2.4.2.1 Non anti-diabetic medication

Non anti-diabetic medications will be summarized by treatment arm according to the WHO-DD dictionary, considering the first digit of the anatomic category (ATC) class and the first 3 digits of the ATC class (therapeutic category). All ATC codes corresponding to a medication will be summarized, and patients will be counted once in each ATC category (anatomic or therapeutic) linked to the medication. Therefore patients may be counted several times for the same medication.

The tables for non-anti-diabetic prior medications will be sorted by decreasing frequency of anatomic class followed by therapeutic class based on the overall incidence across treatment arms. In case of equal frequency regarding ATCs, alphabetical order will be used.

The tables for non-anti-diabetic concomitant and post-treatment medications will be sorted by decreasing frequency of anatomic class followed by therapeutic classes based on the incidence in Toujeo treatment arm. In case of equal frequency regarding ATCs, alphabetical order will be used.

An individual listing sorted by treatment arm and subject number will be populated instead of summary table only if few non-anti-diabetic either prior or post-treatment medications are observed.

2.4.2.2 Anti-diabetic medication

Anti-diabetic medications will be summarized by treatment arm according to the WHO-DD dictionary considering the pharmacological class (ATC3), chemical class (ATC4) and standardized medication name.

The table for anti-diabetic prior medications will be sorted by decreasing frequency of pharmacological class followed by chemical class and standardized medication name based on the overall incidence across treatment arm. In case of equal frequency, alphabetical order will be used.

The tables for anti-diabetic concomitant and post-treatment medications will be sorted by decreasing frequency of pharmacological class followed by chemical class and standardized medication name based on the incidence in Toujeo treatment arm. In case of equal frequency, alphabetical order will be used.

Furthermore,

- Prior and post-treatment anti-diabetic insulin therapy will also be summarized by predefined classification (basal insulin, short-acting insulin, pre-mixed insulin) and standardized medication name.
- Prior, concomitant and post-treatment anti-diabetic non-insulin therapy will also be summarized by predefined classification:
  - Biguanides,
  - SU,
  - Glinides,
  - Thiazolidinedione,
  - DPP-4 inhibitors,
  - SGLT-2 inhibitors,
  - GLP1-RA,
  - Alpha-glucosidase inhibitors,
  - Other.

  and standardized medication name.
- Medication will be summarized by pre-defined classification (insulin and non-insulin therapy) then insulin therapy will be split by pre-defined classification (basal insulin, short acting insulin, premixed insulin) and standardized medication name whereas non-insulin therapy will be split by predefined classification (as described just above) and standardized medication name.

2.4.2.3 Prohibited medication

Previous and concomitant prohibited medications will be presented separately by prohibited medication category as defined in deviation and standardized medication name.

2.4.3 Extent of study treatment exposure

2.4.3.1 Exposure and extent of investigational medicinal product exposure

Extent of exposure

The extent (or duration) of exposure during the study will be the total number of days of administration of IMP, ignoring temporary drug discontinuation, as follows:

Extent (or duration) of exposure to the open-label IMP during the study is defined as:

(Date of the last IMP administration – date of the first IMP administration) + 1.

Duration of IMP exposure will be summarized descriptively as a quantitative variable (number, mean, SD, median, minimum, maximum, Q1 and Q3).

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:

Categories

At time of 6 month database lock,

- Up to 4 weeks,
- >4 to 13 weeks,
- >13 to 25 weeks,
- >25 to 26 weeks,
- >26 to 30 weeks,
- >30 weeks.

At time of 12 month database lock,

- Up to 4 weeks,
- >4 to 13 weeks,
- >13 to 26 weeks,
- >26 to 39 weeks,
- >39 to 50 weeks,
- >50 to 52 weeks,
- >52 to 60 weeks,
- >60 weeks.

2.4.3.2 Persistence on study medication

Treatment persistence will be determined based on vendor claims database that would be responsible for managing and administration of the study drugs (Toujeo, Lantus, Levemir) from the vendor database administering the study drugs, information on prescription fill date, quantity dispensed, days of supply, number of units dispensed, prescription refill dates will be obtained in estimation of treatment persistence and discontinuation.

A medication possession ratio (MPR) will be assessed based on number of days that patients had possession of study drug based on prescription date and days of supply. a MPR of 80% would be indicative of patients being persistent on study medication (5). Those not persistent will be considered discontinued on their study medication. Patients identified by the provider and where there is not data available for calculation of MPR will be summarized by observance of the actual treatment received from the e-CRF.

Treatment persistence will be summarized descriptively (N, Mean, SD, Median, Minimum, and Maximum):

- Mean proportion of patients who have MPR, and,
- The proportion of MPR <80%,

- The incidence of patients completing the 6 month and the 12 month periods based on study visits will be described by treatment arm,

- The proportion of patients who withdraw from study treatment before Month 6 and before Month 12 will be analyzed by treatment group using a logistic regression methodology with the same covariates used in the primary efficacy analysis (Section 2.4.4.1),

- Cumulative incidence of treatment discontinuation based on the e-CRF collected treatment discontinuation dates will be plotted/estimated by treatment arm using a KM by treatment arm.

Patients are not responsible for returning insulin pens or dosing records to their clinic visits to assess compliance. Medication use will be assessed by MPR and persistence measures based on data collected by the Smart card vendor (date of fill or refill and quantity of medication dispensed for 30 day supply). The site does not complete a treatment log form and no compliance information is entered into the e-CRF.

2.4.3.3 Basal insulin dose

The baseline for basal insulin dose is the total IMP dose taken the 1st day of administration (ie, 1st IMP dose for subject following a once daily regimen and the two 1st IMP dose for patients following a twice daily injection regimen).

Observed and change and relative change from baseline value of total daily basal insulin dose (U and U/kg) will be described by visit (at baseline, 6 months and 12 months) per treatment arm.

Mean ($\pm\sqrt{2}$ SE) insulin dose per group at baseline, 6-month and 12-month will be plotted.

2.4.4 Analyses of efficacy endpoints

The baseline value for laboratory efficacy endpoints (HbA1c and FPG), is the last available value prior and up to randomization.

Unless otherwise specified the efficacy analyses will be performed on the ITT population based on the 6-month randomized period at time of 6 month database lock and based on the 12-month randomized period at time of 12-month database lock.

2.4.4.1 Analysis of primary efficacy endpoint

The primary efficacy analysis will be performed on the ITT population, using all post-baseline HbA1c data available and hypoglycemia events collected on the 6-month randomized period, regardless of IMP discontinuation.

The primary efficacy endpoint: proportion of patients reaching individualized HbA1c target at 6 months without documented symptomatic (BG $\leq$70 mg/dL) hypoglycemia will be analyzed using a logistic regression model, adjusting for HbA1c target and other stratification factors used at randomization: SU use (y/n), GLP-1 RA use (y/n) and baseline HbA1c (as continuous). a superiority test will be performed to compare this proportion between treatment groups at IA

(2-sided, alpha = 0.010) and at final 6-month analysis (2-sided, alpha = 0.046). Treatment difference will be assessed by odds ratios and CIs. For the primary efficacy analysis, the randomization strata from IVRS will be used.

Patients with missing HbA1c data at 6 months will be regarded as not having reached HbA1c target. In order to examine the consistency of this analysis, the individual components of HbA1c target attainment and absence of any documented symptomatic (BG ≤70 mg/dL) or severe hypoglycemia will also be evaluated. Each individual component (HbA1c target attainment at 6 month, or absence of any documented symptomatic [BG ≤70 mg/dL] or severe hypoglycemia), will be compared between treatment groups, respectively, using same logistic regression model as the primary efficacy analysis.

Sensitivity analyses

The sensitivity analyses of primary efficacy endpoint will be performed at time of 6-month database lock.

- **On-treatment sensitivity analysis.**

  A sensitivity analysis for the primary efficacy endpoint will be performed on the 6-month on-treatment period, as defined in Section 2.1.3.

- **Multivariate analysis.**

  As a sensitivity analysis, a multivariate analysis will be performed, adding the following baseline covariates in the model: BMI, duration of diabetes, baseline PAM score, hypoglycemic history and age. a backward selection procedure at 0.10 level will be used get the final model with these exploratory factors. Treatment effect will be re-evaluated based on the final model.

  Country (USA/Canada) and US region effect systematic group difference will be explored.

- **Sensitivity analysis of primary efficacy endpoint with randomization strata from e-CRF.**

  The information collected in the e-CRF on SU use, GLP-1 RA use and comorbidity at baseline slightly differs from the definition of randomization stratification factors as per IVRS, as detailed below in Table 4.

Table 4 - Stratification factors discrepancies

	SU use	GLP-1 RA use	Comorbidity (used for HEDIS HbA1c target definition)
IVRS	"Used previously" (anytime in patient history)	"Used previously" (anytime in patient history)	Did the patient exhibit comorbid conditions as outlined per the IVRS worksheets? Refer to Protocol (1) Appendix G (anytime in patient history)
e-CRF	Currently being treated (at study entry)	Within past 6 months	Within past 12 months

IVRS = interactive voice response system; e-CRF = electronic case report form; GLP-1 = glucagon like peptide-1; HEDIS = healthcare effectiveness data and information set; RA = receptor agonist; HbA1c = glycosylated hemoglobin.

To assess the impact of these differences between IVRS randomization strata and the strata which can be derived based on e-CRF data, two sensitivity analyses will be performed.

1. SU use (y/n), GLP-1 RA use (y/n) as per e-CRF and the HEDIS target from IVRS will be used in the logistic regression model.

2. SU use (y/n), GLP-1 RA use (y/n) from IVRS included in the logistic regression model and HEDIS target strata will be dropped from the logistic model.

- **Sensitivity analysis of primary efficacy endpoint with MAR multiple imputation.**

As a sensitivity analysis to assess the impact of missing data, missing HbA1c data and documented (BG ≤70 mg/dL) or severe hypoglycemia will be handled using the multiple imputation method.

Months 6 and 12 data will be imputed separately. For 6-months data, randomization strata, and treatment will be included in the imputation model. For 12-month data, 6-month data, randomization strata, and treatment will be included in the imputation model.

Documented hypoglycemia (yes/no/missing) and HbA1c values by visit will be presented by for each treatment group, using descriptive statistics.

For documented hypoglycemia, if a patient drops out from the study prior to 6 months and no documented hypoglycemia is entered into the CRF during the first 6-month treatment period, the patient status with regard to documented hypoglycemia (presence or absence of hypoglycemia) will be considered as missing. Similarly, if a patient drops out from the study prior to 12 months and no documented hypoglycemia is entered during the 12-month treatment period, the patient's documented hypoglycemia status will be considered missing. If a patient has a documented hypoglycemia event within the time frame, then this subject will be considered as a subject with documented hypoglycemia, no matter it is a completer or early withdrawal.

Missing data will be imputed 100 times to generate 100 complete data sets with the MI with Markov Chain Monte Carlo (MCMC) method in SAS procedure in two steps.

- Step 1: Six-month HbA1c value will be imputed using the regression method with stratification factors and treatment. For each simulation leading to imputed HbA1c values outside of 3.5% to 16%, another value will be redrawn for imputation.

- Step 2: If the 6-month documented (BG ≤70 mg/dL) hypoglycemia status is missing, after Step 1, HbA1c is imputed, 6-month documented (BG ≤70 mg/dL) hypoglycemia status will be subsequently imputed using a logistic regression method with 6-month HbA1c, stratification factors, and treatment.

For each of the 100 imputed datasets, the composite endpoint will be constructed and compared between treatment groups using the same logistic regression model as specified above in the primary efficacy endpoint analysis. The estimates from the 100 fitted models will be combined to provide an overall estimate with corresponding confidence intervals (CI). The difference between treatments and the corresponding CI will be presented. The imputed values of HbA1c will also be used in the analysis of HbA1c change as continuous variable (secondary efficacy endpoint).

- **Sensitivity analysis of primary efficacy endpoint with tipping point analysis.**

  Robustness of the primary analysis results to departure from the MAR assumption will be explored in the ITT population using tipping-point analysis based on the pattern mixture model approach.

  Estimations will be performed using the same multiple imputation approach as described previously. In each treatment arm, a penalty δ will be added to the imputed values ($\delta = 0$ corresponds to the MAR assumption) in the HbA1c multiple imputation step.

  The documented hypoglycemia status will be imputed as described previously following the HbA1c imputation with various δ.

  To investigate how the conclusions for the composite endpoint depend on the adopted values δ, the testing will be repeated over a range of plausible values for the pairs (δ_{Toujeo}, $\delta_{\text{std\_care\_basal\_insulin}}$). The δ values range from -1 to 1 by 0.1, in each treatment group. Results will then be summarized using graphs.

  The smallest value of δ_{Toujeo} for a value of $\delta_{\text{std\_care\_basal\_insulin}}$ equal to 0 which will statistically reverse the conclusion for the composite primary endpoint will be assessed and interpreted from its clinical plausibility.

- **Sensitivity analysis excluding patient data from sites at which research activities were terminated due to non-compliance.**

  A sensitivity analysis using the same method as for the primary efficacy analysis will be performed excluding patient data from sites at which research activities were terminated due to ongoing non-compliance with the clinical protocol and violations of GCP.

  Sites ███████████████████ and site ███████████████████ are the two concerned. They will be excluded for a sensitivity analysis at 6 months.

Subgroup Identification based on Differential Effect Search

If superiority of Toujeo with respect to standard-of-care basal insulin is not shown in the primary efficacy analysis, an exploratory analysis will be conducted to identify potential subgroups of patients with enhanced treatment effect of Toujeo versus standard of care (Lantus or Levemir) basal insulin.

A procedure based on Subgroup Identification based on Differential Effect Search (6) will be employed to identify such subgroups of potential responders.

The outcome variable used in the SIDES method is the primary endpoint of the study which is the proportion of patients with individualized HbA1c target at 6 months, without any documented symptomatic or severe hypoglycemia.

Patient characteristics on which potential subgroups will be defined, include randomization strata which are used in the primary efficacy analysis and other variables collected at baseline: randomization strata; SU use (y/n), GLP-1 RA use (y/n), baseline HbA1c (as continuous), baseline age, race/ethnicity, gender, baseline fasting glucose level, baseline BMI, duration of diabetes, diabetic disease history (diabetic retinopathy, diabetic sensory or motor

neuropathy, diabetic autonomic neuropathy, diabetic nephropathy), comorbidity (coronary artery bypass graft, percutaneous coronary intervention, ischemic vascular disease, thoracic aortic aneurysm, chronic heart failure, prior myocardial infarction, chronic renal failure or end stage of renal disease, dementia, blindness, and leg amputation). Since HbA1 target as per HEDIS depends on age and comorbidity, both already included, it will not be added to the list of covariates for subgroup search.

The global procedure includes the following steps:

- Balanced allocation procedure:

  The data is split into one training data set on which the subgroup identification procedure will be applied, and one or several validation data set. This allocation procedure is based on an imbalanced score so that the distribution of patients' treatment and covariates is balanced between sets.

- SIDES recursive partitioning algorithm:

  Subgroups are constructed from the training data set by recursively partitioning from covariates the data set into two subgroups at each parent group so that the treatment effect in one group is maximized compared to the other.

- Validation of identified candidates subgroups:

  One validation set is used to confirm or reject candidate subgroups identified in the previous step based on-treatment effect p-value estimated on the validation set.

Note:

As specified in the article (6), the identification of candidate subgroups step can be applied to the entire data set, and confirmation performed in a subsequent independent study.

This will be the preferred option: the dataset will not be divided in training and test datasets, and only Step 2 (SIDES recursive partitioning algorithm) will be applied.

The SIDES recursive partitioning algorithm can be described as follows:

Consider the current parent subgroup. At the beginning, no covariate is used to define subgroup and the parent subgroup is the entire training set.

- Among covariates not already used to define parent subgroup,

- For each covariate, consider all possible splits of the covariate into two child subgroups (the number of possible splits depends on the type of the covariate),

- If the subgroup with the larger treatment effect of one split has a sample size lower than a pre-defined threshold, then the split is excluded,

- For each split among all splits of the previous step, calculate the splitting criterion that aims at maximizing the differential effect between the two child subgroups,

- $\rho = 2\left[1 - \phi\left(\dfrac{Z_1 - Z_2}{\sqrt{2}}\right)\right]$,

- Where Z_1 and Z_2 denote the Z-statistics for efficacy of the subgroups pair composing one split, and ϕ represents the standard normal distribution function,

- Apply multiplicity adjustment to splitting criterion, based on modified Sidak correction, for covariates with more than two levels.

- Order all splits in terms of adjusted splitting criterion, and select the best M splits, where M is a pre-defined value.

- From each pair composing one split, select the subgroup with the larger treatment effect. These subgroups are called "promising child subgroups".

- If a promising child subgroup meets the continuation criterion, that is if $\rho_c \leq \rho_p$ where ρ_c, respectively ρ_p, are the p-values of treatment effect for child, respectively parent, subgroups; then it is added to the set of parent groups.

- In addition of meeting the continuation criterion, if a promising child subgroup meets the selection criterion, it is added to the set of candidate subgroups return at the end of the recursive search algorithm. Treatment effect p-values of promising child subgroups are adjusted using a standard resampling-based method to have a weak control of type I error at a level α. Then, the selection criterion is met if the adjusted p-value in the subgroup is significant at a pre-defined level α.

- Call recursively the algorithm on each updated parent subgroups. The algorithm stops when a maximum number of L covariates are used to define parents, where L is pre-defined, or when no promising child subgroup meets the continuation criterion.

2.4.4.2   Analyses of secondary efficacy endpoints

Change in HbA1c

The change in HbA1c from baseline to 6 months will be analyzed using an analysis covariance (ANCOVA) model. The model will include a fixed categorical effect of treatment arm, HbA1c target and other randomization strata; SU use (y/n), GLP-1 RA use (y/n), as well as baseline BMI (as continuous) and baseline HbA1c (as continuous).

The change in HbA1c from baseline to 12 months will be analyzed a MMRM approach under the MAR framework, using an adequate contrast at Month 12. The model will include fixed categorical effects of treatment arm, visit, treatment-by-visit interaction, HbA1c target and other randomization strata; SU use (y/n), GLP-1 RA use (y/n), and baseline A1c (as continuous), as well as baseline HbA1c-by-visit interaction. Both 6 months and 12 months HbA1c data will be included in the MMRM model.

The analysis of HbA1c is based on patients in the ITT population. For those with missing HbA1c value at 6 months, the change in HbA1c will be missing. Therefore, those patients will not be included in the analysis.

As sensitivity analysis, the imputed HbA1c values at 6 months and 12 months described in Section 2.4.4.1 will be used to repeat the change in HbA1c from baseline to 6 months and to

12 months analysis models. The estimates from the 100 fitted models will be combined to provide an overall estimate with corresponding CIs for the treatment differences in change from baseline in HbA1c to 6 months and 12 months.

Change in FPG and body weight

The change in FPG and body weight from baseline to 6 and 12 months will be analyzed in a similar fashion to the change in HbA1c.

Other binary endpoints

The proportion of patients with HbA1c <8% and <7% without documented symptomatic (BG ≤70 mg/dL) or severe hypoglycemia at 6 months and of patients with HbA1c <8% and <7% at 6 months will also be analyzed by the same logistic regression model as for the primary efficacy analysis.

The proportion of patients with individualized HbA1c target attainment without documented symptomatic (BG ≤70 mg/dL) or severe hypoglycemia at 12 months, and the proportion of patients with individualized HbA1c target attainment without documented symptomatic (BG <54 mg/dL) or severe hypoglycemia at 6 and 12 months be analyzed as described for the primary endpoint.

Health resource utilization

Detailed analysis methods for health resource utilization will be described in a separate analysis plan.

Subgroup factors

For subpopulations of sufficient size, additional subgroup analyses may also be performed.

The following treatment effects across different subgroups (baseline or screening factors), where appropriate according to subgroup size, will be explored on primary efficacy endpoint:
- Gender (male, female),
- Randomization strata of individualized HbA1c target (<8%/<7%),
- Randomization strata of GLP-1 RA use (y/n),
- Randomization strata of SU use (y/n),
- Baseline HbA1c category (<9%/≥9%),
- Age group (<65, [65–75] and ≥75 years of age),
- Baseline BMI categories (<25, [25-30], [30-40] and ≥40 kg/m²),
- Category of duration of diabetes (<10, ≥10 years).

For each subgroup, the primary efficacy variable will be analyzed in the ITT population using post-baseline HbA1c data available on the 6-month randomized treatment period. A similar logistic regression as described for primary analysis will be applied adding the corresponding subgroup factor, subgroup factor-by-treatment interaction and subgroup factor-by-visit-by treatment interaction (for 12-month analysis only).

When the subgroup considered is equal to one of randomization stratum or is part of one of the adjustment factor, this randomization stratum/adjustment factor is removed from the model.

The significance level of the treatment-by-subgroup factor interaction term at Month 6 will also be determined for each factor for descriptive purpose. Forest plots will be provided.

The individual components of the composite endpoint of the primary efficacy variable will also be assessed for each subgroup factor.

2.4.4.3 Multiplicity issues

All secondary endpoints will be exploratory and descriptive and therefore no adjustments for multiplicity are specified.

2.4.5 Analyses of safety endpoints

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

All safety analyses will be performed on the safety population using the following common rules:
- The baseline value is defined as the last available value prior to the first injection of IMP,
- There are no safety endpoints which are adjudicated.

The baseline values for all safety analyses are defined as the last available value prior to the first dose of randomized treatment. Given the short half-life for duration of action of insulin, the TEAE reporting period is up to one day after last administration of IMP.

The analysis of the safety variables is essentially descriptive and no systematic testing is planned.

In addition, for hypoglycemia events, relative risks and odd-ratios with their associated 95% CI will be provided.

2.4.5.1 Analyses of hypoglycemia

All safety analyses of hypoglycemia events will be performed by actual treatment arm in the safety population, using events occurring during TEAE period.
- At time of 6-month database lock during:
  - The 6-month on-treatment period.
- At time of 12-month database lock during:
  - The 12-month on-treatment period but also,
  - The 6-month on-treatment period,
  - The 6 to 12 month on-treatment period.

The number and proportion of patients experiencing at least one hypoglycemia event and the number and rate (per patient-year of exposure) of hypoglycemic events will be presented by treatment arm and type of hypoglycemia event (Appendix A) according to time of occurrence:

- Nocturnal (defined by clock time:00:00 ^{AM}-5:59 ^{AM}),
- Daytime (defined by clock time: 6:00 ^{AM}-23:59),
- And at any time of the day.

during each on-treatment period (as defined above).

Any hypoglycemia and un-classified ones (if any) will also be presented by treatment arm.

The incidence of hypoglycemia will be analyzed using logistic regression, adjusting for randomization strata: HbA1c target, SU use (y/n) and GLP-1 RA use (y/n), as well as baseline HbA1c (as continuous).

The rate per patient-year of hypoglycemic events will be analyzed using an over-dispersed Poisson regression model, adjusting for randomization strata. The model will include the logarithm of the duration of the considered on-treatment period (in years) as an offset variable and a log-link function.

Cumulative mean number of severe and/or symptomatic documented (≤70 mg/dL and <54 mg/dL) hypoglycemia will be summarized over time using Nelson-Aalen estimates and plotted.

Subgroup analyses

The proportion of patients with at least one:

- Severe and/or symptomatic documented (≤3.9 mmol/L [≤70 md/dL]) hypoglycemia,
- Severe and/or symptomatic documented (<3.0 mmol/L [<54 md/dL]) hypoglycemia.

during respectively the 6-month on-treatment period and the 12-month on-treatment period (any time of the day), will also be presented by the following subgroups:

- Gender (male, female),
- Randomization strata of individualized HbA1c target (<8%/<7%),
- Randomization strata of GLP-1 use (yes/no),
- Randomization strata of SU use (yes/no),
- Baseline HbA1c category (<9%/≥9%),
- Age group (<65, [65-75] and ≥75 years of age),
- Baseline BMI categories (<25, [25-30], [30-40] and ≥40 kg/m²),
- Category of duration of diabetes (<10, ≥10 years).

and analyzed using a similar logistic regression model approach as presented previously, but adding the corresponding subgroup factor and treatment arm-by-subgroup interaction factor. Odds ratio estimates and 95% CI will be provided in each subgroup category.

The interaction treatment arm-by-subgroup p-value will be provided for descriptive propose.

When the subgroup considered is equal to one of the randomization strata or is part of one of the adjustment factor, this randomization stratum/adjustment factor is removed from the model.

If the logistic regression model does not converge (eg, due to sparse data) some of the randomization strata may be removed.

Forest plots will be provided.

Relative Risks

For each hypoglycemic category of event (Appendix A) except relative and probable, the estimated proportions along with their corresponding 95% CI, as well as the relative risk of Toujeo arm versus SOC basal insulin arm and their corresponding 95% CI, will be estimated by the mean of a log-binomial regression model, using a log-link function and a binomial response distribution, and adjusting for:

- Randomization strata HbA1c target,
- Randomization strata of SU use (y/n),
- Randomization strata of GLP-1 RA (y/n),
- Baseline HbA1c (as continuous).

Risk difference

For each hypoglycemic category of event (Appendix A) except relative and probable, the adjusted risk differences of Toujeo arm versus SOC basal insulin arm and their corresponding 95% CI, will be estimated by the mean of a log binomial regression model, with an identity link function and a binomial response distribution, and adjusting for:

- Randomization strata HbA1c target,
- Randomization strata of SU use (y/n),
- Randomization strata of GLP-1 RA (y/n),
- Baseline HbA1c (as continuous).

Scatter plots

Scatter plots of Relative Risk versus Risk Difference presenting each type of analyzed hypoglycemia for patient experiencing at least one hypoglycemia event.

- During respectively the 6-Month on-treatment period and the 12-Month on-treatment period,
- According to time of occurrence (any time of the day, nocturnal [defined by clock time: 00:00 ^{AM}-5:59 ^{AM}]).

2.4.5.2 *Analyses of adverse events*

Generalities

All safety analyses of hypoglycemia events will be performed by actual treatment arm in the safety population. The primary focus of AE reporting will be on TEAEs. Pre-treatment and post-treatment AEs will be described separately.

The primary focus of AE reporting will be:
- At time of 6 month database lock: 6 month on-treatment AE, occurring during the 6 month on-treatment period.
- At time of 12 month database lock: on treatment-emergent adverse events, occurring during the 12 month on-treatment period.

Pre-treatment and post-treatment AEs will be described separately.

Post-treatment AE will only be displayed at time of 12-month database lock but also at time of 6 month database lock for patients who permanently discontinued treatment during the 6 month randomized period.

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-treatment 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 present by SOC, HLGT, HLT, and PT, sorted in alphabetical order for each treatment group, the number (n) and percentage (%) of patients experiencing an AE. 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 AEs within the observation period (pre-treatment, treatment-emergent, and post-treatment). For that purpose, the table of all treatment-emergent adverse events presented by SOC and PT sorted by the internationally agreed SOC order and decreasing frequency of PTs within SOCs will define the presentation order for all other tables unless otherwise specified. Sorting will be based on results for the IMP.

Analysis of all treatment-emergent adverse events

The following treatment-emergent adverse event summaries will be generated for the safety population.
- Overview of treatment-emergent adverse events, summarizing number (%) of patients with any,
  - Treatment-emergent adverse event,
  - Serious treatment-emergent adverse event,
  - Treatment-emergent adverse event leading to death,
  - Treatment-emergent adverse event leading to permanent treatment discontinuation,
  - Treatment-emergent adverse event related to IMP,
  - Treatment-emergent adverse event related to anti-diabetic NIMP,
  - Treatment-emergent adverse event related to pen device.

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

- Number (%) of patients experiencing TEAE(s) presented by PT, sorted by decreasing incidence of PT in the Toujeo arm.

- All treatment-emergent adverse events by primary SOC and PT, showing the number (%) of patients with at least 1 treatment-emergent adverse event, sorted by the internationally agreed SOC order and by decreasing incidence of PTs in the Toujeo arm within each SOC. This sorting order will be applied to all other tables, unless otherwise specified.

- All common TEAEs (HLT incidence ≥2% in any treatment arm) by primary SOC, HLT and PT, showing the number (%) of patients with at least 1 common TEAE, sorted by the internationally agreed SOC order. The other levels (HLT, PT) will be presented in alphabetical order.

- All treatment-emergent adverse events 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 adverse event, sorted by the internationally agreed SOC order. The other levels (HLGT, HLT, PT) will be presented in alphabetical order.

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

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

- All treatment-emergent adverse events by maximal severity, presented by primary SOC and PT, showing the number (%) of patients with at least 1 treatment-emergent adverse event by severity (ie, mild, moderate, or severe), sorted by the sorting order defined above,

- Number (%) of patients experiencing treatment-emergent adverse event(s) presented by primary and secondary SOC, HLGT, HLT, and PT sorted by the internationally agreed SOC order. The other levels (HLGT, HLT, PT) will be presented in alphabetical 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 serious treatment-emergent adverse event, 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 SAE, 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 treatment-emergent adverse events leading to treatment discontinuation, by primary SOC, HLGT, HLT, and PT, showing the number (%) of patients sorted by the internationally agreed SOC order. The other levels (HLGT, HLT, PT) will be presented in alphabetical order.

Injection site reaction and hypersensitivity reaction

- Number (%) of patients with events related to injection site reactions and hypersensitivity reactions will be provided separately.

Adverse event of special interest

Symptomatic overdose

- A listing of patients with symptomatic overdose with IMP/NIMP will be provided.

ALT increase

- A listing of patients with increase of ALT (collected as an AE) will be provided.

Pregnancy

- A listing of patients with pregnancy (including event occurring to a partner) will be provided.

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.
- All pre-treatment SAEs by primary SOC and PT, showing the number (%) of patients with at least 1 post-treatment SAE, sorted by the internationally agreed SOC order and decreasing incidence of PTs within each SOC.
- All pre-treatment AEs leading to study discontinuation by primary SOC and PT, showing the number (%) of patients, sorted by the internationally agreed SOC order and decreasing incidence of PTs within each SOC.
- 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.
- All post-treatment SAEs by primary SOC and PT, showing the number (%) of patients with at least 1 post-treatment serious AE, sorted by the internationally agreed SOC order and decreasing incidence of PTs within each SOC.

2.4.5.3 Deaths

The following deaths summaries will be generated on the safety population:

- Number (%) of patients who died by study period (pre-treatment, on-study, on-treatment, post-study) and reasons for death summarized by treatment arm,

- Death in non-randomized patients or randomized and not treated patients,

- TEAE leading to death (death as an outcome on the AE CRF 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 presented in alphabetic order within SOC.

- All pre-treatment AEs leading to death by primary SOC and PT, showing the number (%) of patients, sorted by the internationally agreed SOC order and decreasing incidence of PTs within each SOC (if more than 5 patient with the event),

- All post -treatment AEs leading to death by primary SOC and PT, showing the number (%) of patients, sorted by the internationally agreed SOC order and decreasing incidence of PTs within each SOC (if more than 5 patient with the event).

A listing of all deaths sorted by treatment arm will be performed.

2.4.5.4 Analyses of laboratory safety variables

Not applicable.

2.4.5.5 Analyses of vital sign variables

The summary statistics (including number, mean, median, SD, minimum and maximum) of blood pressure, heart rate and body weight (values and changes from baseline) will be calculated for each visit or study assessment (baseline, each post baseline time point, last on-treatment) by treatment arm.

2.4.5.6 Analyses of electrocardiogram variables

Not applicable.

2.4.5.7 Analyses of Patient-Support Program (PSP)

Participation to a PSP will be described by visit and overall per randomized treatment arm on the safety population.

Furthermore, the number and percentage of patients having or not initiated a Toujeo PSP at any time during respectively the 6-month randomized period at time of 6 month database lock and the 12-month randomized period at time of 12 month database lock will be described in the Toujeo randomized treatment arm, according to demographic and baseline characteristics.

2.4.6 Analyses of pharmacokinetic and pharmacodynamic variables

Not applicable.

2.4.7 Analyses of patient reported outcome variables

The patient report outcomes (PRO) analyses will be performed on the ITT population.

Self-reported treatment satisfaction

DTSQs and DTSQc scores at 6 and 12 months will be compared between treatment groups using an ANCOVA approach as described for the continuous efficacy variables. The ANCOVA model with satisfaction score as the dependent variable includes HbA1c target and other stratification factors used at randomization: SU use (y/n), GLP-1 RA use (y/n) and baseline HbA1c (as continuous) in the model.

The consistency between results of DTSQs change from baseline to 12 months and DTSQc scores at 12 months will be assessed.

Cumulative distribution function of score changes from baseline at 6 and 12-months will be plotted by treatment arm for each of the 3 DTSQs scores. Cumulative distribution function of 12-month DTSQc scores will be plotted by treatment arm.

Hypoglycemia Patient Questionnaire

Items of the Hypoglycemia Patient Questionnaire will be summarized by descriptive statistics. a graphical presentation will also be provided.

The proportion of patients reporting one or more episodes of either moderate or severe hypoglycemia will be compared between treatment groups using the same logistic regression model as for GES detailed in following paragraph.

Global Effectiveness Scale

At each study endpoint, the proportion of patients and providers who reported "complete control of diabetes" or "marked improvement of diabetes" at 6 and 12 months will be analyzed using a logistic regression same as the primary efficacy analysis. The dependent variable is the "complete control/marked improvement" (y/n). The logistic model includes HbA1c target and other stratification factors used at randomization: SU use (y/n), GLP-1 RA use (y/n) and baseline HbA1c (as continuous) in the model.

Additionally, at each time point (6 months and 12 months), all response choices for both patients and provider will be summarized by descriptive statistics.

2.4.8 Analyses of Health economic endpoints

Detailed analysis methods for health resource utilization will be described in a separate analysis plan.

2.5   DATA HANDLING CONVENTIONS

2.5.1 General conventions

Reference day

The reference day for the calculation of extent of exposure, time to onset and relative days is the day of the first IMP administration, denoted as Day 1.

Demographic formulas

Age (years) will be calculated as follows = Integer ([Date of informed consent – Date of birth + 1]/365.25)

BMI (kg/m²) will be calculated as follows = (Weight in kg)/(Height in meters²)

In case of partial dates, (Section 2.5.3).

Disease history

Diabetes duration of (years) will be calculated as follows = (Date of informed consent – Date of diagnosis of diabetes + 1)/365.25.

Age at diagnosis of diabetes (years) will be calculated as follows = (Date [MM-YYYY] of diagnosis of diabetes – date [MM-YYYY] of birth + 1)/365.25.

In case of unavailable date of birth, only the year of the date of diabetes diagnosis and the year of the date of birth (retrieve using the age recorded at screening) will be considered in the age at diagnosis of diabetes calculation.

Duration of previous non-insulin anti-hyperglycemic treatment (years) will be calculated as follows = (Date of informed consent – date of first dose of previous non-insulin anti-hyperglycemic treatment + 1)/365.25

IMP exposure

IIMP (U/Kg) will be determined as:

IMP dose (U)/Body weight (kg) obtained at a given visit.

Laboratory efficacy data conversion

HbA1c conversion:

IFCC in mmol/mol = (10.93× NGSP in %) - 23.5.

FPG and SMPG conversion:

From mg/dL to mmol/L: x 0.0555,

From mmol/L to mg/dL: x 18.0148.

Renal function formulas

Conversion for serum creatinine:

From mg/dL to µmol/L: x 88.402.

Conversion for total bilirubin:

From mg/dL to µmol/L: x 17.104.

Creatinine clearance value will be derived using the equation of Cockroft and Gault, using weight assessed at the same visit:

For Male:

$$\text{CLcr} \, (\text{mL/min}) = \frac{[(140 - age(years)) \times weight(kg)]}{[0.814 \times \text{serum creatinine}(\mu\text{mol/L})]}$$

For Female: result above multiplied by 0.85.

Hepatic function formulas

Conversion for total bilirubin:

From mg/dL to µmol/L: 17.10.

Patient reported outcomes

Minimal clinically important difference:
- Standard MCID threshold will be calculated as half- SD at baseline.

- Distribution-based MCID threshold will be calculated as $SD_{baseline} \times \sqrt{1-r}$ where r is the Cronbach's α reliability coefficient at baseline.

Effect sizes (ES)

- Within and between effect sizes (ES) interpretation is the following:
  - An ES <0.2 is considered negligible,
  - 0.2 ≤ES <0.5 is considered small,
  - 0.5 ≤ES <0.8 is considered moderate,
  - ES ≥0.8 is considered important.

- Within ES, calculated at treatment arm level, indicates if a change from baseline to the Month 6/Month 12 is clinically meaningful:

$$\frac{LS \ mean \ change \ from \ baseline \ at \ timepoint \ t}{Pooled \ baseline \ SD}$$

- Between ES, calculated when comparing treatment arm, indicates if the mean changes from baseline to Month 6/Month 12 are clinically meaningful between treatment arm,

$$\frac{\textit{Difference in LS means between groups at timepo}\,\text{int }t}{\textit{Pooled baseline SD}}$$

- Both LS means will be determined using the MMRM/ANCOVA whereas pooled baseline SD will be determine from descriptive statistic at baseline.

$$\frac{N_{Toujeo} \times SD_{Toujeo} + N_{standard\_of\_care\_basal\_insulin} \times SD_{standard\_of\_care\_basal\_insulin}}{N_{Toujeo} + N_{standard\_of\_care\_basal\_insulin}}$$

Table 5 - Unexpected Scoring

Scoring issues	DTSQs	DTSQc
In between score	If patients answer in between choices, score should be coded as mean of both scores (ie, if between 2 and 3 coded 2.5)	
Item answered twice using 2 consecutive scores	Midpoint score should be used (ie, if 1 and 2 are answered should be coded 1.5)	
Item answered twice but using non-consecutive scores	Score is set to missing	
Item is circled but accompanied by word	Only score should be taken into account	
If respondent as circle word of the extremes of the scale instead of value	Score should be the nearest value (ie, 0 or 6)	Score should be the nearest value (ie, -3 or 3)

ABBREVIATION: DTSQs = diabetes treatment satisfaction questionnaire status version.

Table 6 - PRO missing data handling

PRO scales	Score	Number of items included in score derivation	Minimum number of items required for score calculation
DTSQs	Total treatment satisfaction score	6	4
	Perceived frequency of hyperglycemia score	1	1
	Perceived frequency of hypoglycemia score	1	1
DTSQc	Total treatment satisfaction score	6	4
	Perceived frequency of hyperglycemia score	1	1
	Perceived frequency of hypoglycemia score	1	1

ABBREVIATION: DTSQs = diabetes treatment satisfaction questionnaire status version; DTSQc = diabetes treatment satisfaction questionnaire change version.

DTSQc and DTSQs total treatment satisfaction score

If no more than 2 out of the 6 questions comprising the treatment satisfaction score are missing, the treatment satisfaction score is imputed by calculating the average of the scores from the answered questions, dividing this sum by the number of answered questions and multiplying the average by six.

2.5.2 Data handling conventions for secondary efficacy variables

Fasting condition

FPG measurement not collecting in fasting condition will not be used in the analyses.

Invalid laboratory data

HbA1c or FPG measurements flagged as invalid by the laboratory will not be used in the analyses.

Treatment persistence/switch

A switch is defined by:
- A switch from Toujeo IMP to any commercial "standard of care" basal insulin or,
- A switch from "Lantus or Levemir" IMP basal insulin to commercial Toujeo.

Change in "standard of care" basal insulin to another basal insulin not being considered as a "switch".

Remain on assigned therapy is defined by,
- Neither permanently withdraw,
- Nor switch:
  - From IMP Toujeo to Lantus or Levemir insulin,
  - Or from "standard of care" basal insulin to commercial Toujeo,
- Assigned therapy refers to the treatment allocated by IRT: Toujeo or Lantus/Levemir,
- Commercial Toujeo not being allowed in both treatment arms.

Persistence is defined by,
- Remained on assigned therapy.

HbA1c target and severe and or symptomatic hypoglycemia

HbA1c target missing data handling,
- Patients with missing HbA1c data at respectively Month 6 and Month 12 will be regarded as not having reached the HbA1c target (ie, as failure) at each given time-point.

Symptomatic hypoglycemia target missing data handling

In case of at least one symptomatic event with glycemic value missing it will be considered that the patients experiences at least:
- One document symptomatic hypoglycemia (BG $\leq$70 mg/dL [$\leq$3.9 mmol/L]) and,
- One document symptomatic hypoglycemia (BG <54 mg/dL [<3.0 mmol/L]).

In case of at least one symptomatic event with timing missing it will be considered that the patients experiences at least:

- One symptomatic hypoglycemia event occurring at any time of the day and,
- One nocturnal symptomatic hypoglycemia event.

Composite endpoints missing data handling

- For each cases presented above regarding missing HbA1c values the patient will be considered as failure for each composite endpoint of the corresponding randomized period.
- For each case presented above regarding missing hypoglycemia event information the patient will be regarded as failure for each related composite endpoint.
- In case of premature end of study during the 6 month randomized period the patient will be considered as failure for each composite endpoint relative to the 6-month and the 12 month period.
- In case of premature end of study during the 12 month randomized period the patient will be considered as failure for each composite endpoint relative to the 12-month period.

Change in HbA1c category

Patients with missing:

- Baseline HbA1c will be regarded as not having reached the HbA1c target at Month 6 and Month 12.
- HbA1c data at respectively Month 6 and Month 12 will be regarded as not having reached the HbA1c target at each given time-point.

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. Number of patients/events with imputed missing data should be identified and presented in table (if applicable).

Handling missing first IMP date and reference start date

For patient randomized and exposed with missing IMP start date the reference day will be set to randomization date.

Handling of missing weight in the calculation of basal insulin dose in (u/kg)

For calculation of basal insulin dose expressed in U/kg if the body weight measurement is missing at a given visit, the last available measurement from previous visit will be used.

Handling partial date of birth

For calculation age, missing birth day is assumed as the 15 of the month. If day and month are missing, missing birth month and day are assumed as the 15 of June. If year is missing, age could not be calculated. If calculated age is less than 18 (but not less from 17) and corresponding exclusion criteria (E02) not met, 18 will be assumed.

Handling of partial date of diagnosis of Type 2 diabetes

Partial date of diagnosis of Type 2 diabetes will be imputed to calculate the duration of diabetes.

In order to calculate:

- Duration of diabetes (year),
- Age at diagnosis of diabetes (year).

If only day is missing, only month and year will be used.

If month is missing, only year will be used in calculation.

Handling of partial date of start date of treatment with OADs

Partial start date of treatment with OADs will be imputed to calculate the OAD duration.

- If the day is missing and month is present the 15 of the month will be used.
- If year is present and equal to year of date of diagnosis of Type 2 diabetes and day and month are missing then the date of diagnosis of Type 2 diabetes will be used.
- If year is present (and not equal to year of date of diagnosis of Type 2 diabetes) and day and month are missing then the 1st July will be used.

Handling of computation of treatment duration if investigational medicinal product end of treatment date is missing

For the calculation of the treatment duration, the date of the last dose of IMP is equal to the date of last administration reported on the end-of-treatment CRF page. If this date is missing, the exposure duration should be left as missing (ie, in extent of exposure duration).

The last dose intake should be clearly identified in the CRF 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 adverse events/hypoglycemia with missing or partial date/time of onset

Missing or partial AE/hypoglycemia event onset dates and times will be imputed so that if the partial AE/hypoglycemia event onset date/time information does not indicate that the event started prior to treatment or after the TEAE/12 month on-treatment period, the event will be classified as treatment-emergent.

Furthermore, if the hypoglycemia event onset dates and times does not indicate that the event belongs:

- At time of 6-month database lock to either the 3 month on-treatment period or to the 3 to 6 month on-treatment period the hypoglycemia will be classified in both periods,

- At time of 12-month database lock to either the 6-month on-treatment period or to the 6 to 12 month on-treatment period the hypoglycemia will be classified in both periods.

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/last investigational medicinal product administration is missing

When the date and time of the first IMP administration is missing, all AEs: hypoglycemia that occurred on or after the day of randomization should be considered as treatment-emergent.

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

Handling of missing assessment of relationship of adverse events to investigational medicinal product

If the assessment of the relationship to IMP is missing, then the relationship to IMP 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 of adverse events

If a severity is missing, estimate the missing severity/by the maximal severity/observed in the treatment arms.

Handling hypoglycemia with missing or partial time

No imputation for missing or partial time of hypoglycemia will be done.

However:

- If the minutes are missing and the hours comprised between 00 and 05, the event will be considered as nocturnal,
- If hours are missing and the patient was awaken during nocturnal, the event will be considered as nocturnal (defined by clock time).

Handling of hypoglycemia event classification when some classification items are missing

In case of information missing on the item "Assistance required", "Associated with clinical symptoms", or "Plasma glucose", the following algorithm was applied:

Figure 2 - Flow chart for hypoglycemia event classification in case of missing items
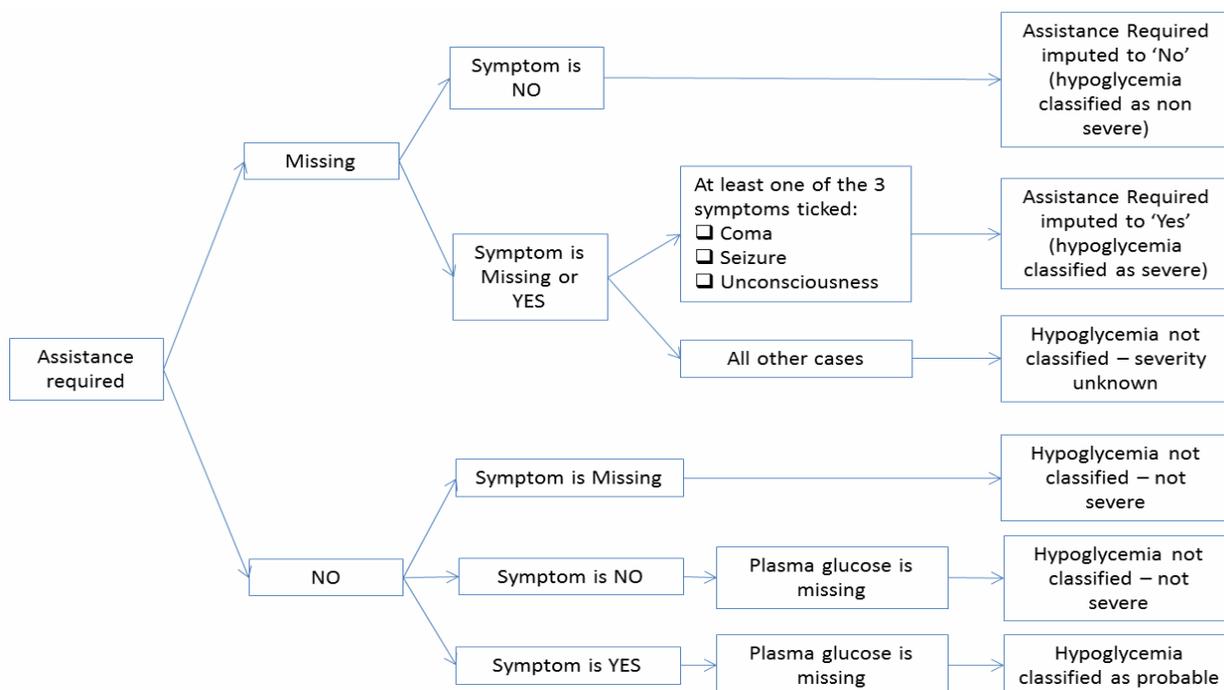
As a conclusion, for cases where hypoglycemia could not be fully classified, as described above, two categories will therefore be created:

- Hypoglycemia classified as Non-classified Hypoglycemia (severity unknown) if the variable "assistance required" is equal to "missing".
- Hypoglycemia classified as Non-classified Hypoglycemia (non-severe) if the variable "assistance required" is equal to "No".

Handling of PRO missing data

- Handling of missing data and multiple answers on DTSQ available in user guidelines: refer to Table 6.
- For Hypoglycemia Patient Questionnaire, if items 4 and 5 missing, patients will be considered as non-responders.
- For GES, if data are missing, patients will also be considered as non-responders.

2.5.4 Windows for time points

The following process will be applied for Month 6, Month 12, premature EOT visits and unscheduled visits re-allocation. Re-allocated visits will be used in statistical analyses for the following endpoints: HbA1c, FPG, Body weight, dose and PROs (descriptive statistics, graphs, and statistical models).

Assessments will be re-allocated to the closest visit using the time windows given in Table 7.

Table 7 - Re-allocation time windows definition

Re-allocated scheduled visit	Targeted study day/month	Re-allocation time windows in study days/month as 30-day period
Visit 3 (Month 6)	180 (Month 6)	(120-240)/end of Month 4 to end of Month 8
Visit 4 (Month 12)	360 (Month 12)	(300-420)/end of Month 10 to end of Month 14

Study days are calculated from the randomization day, being Day 1.

If more than one visits/measurements fall into a visit window, the one closest to the scheduled visit will be used.

For efficacy labs, only central lab data will be used.

2.5.5 Unscheduled visits

The determination of baselines and values at Month 6 and Month 12 for efficacy variables (measurements from the central laboratory only) is based on all measurements from both scheduled and unscheduled visits.

The determination of the last on-treatment value for safety parameters is also based on all assessments from both scheduled and unscheduled visits.

Unscheduled visit measurements of laboratory data, vital signs, will not be included in the by-visit summaries.

2.5.6 Pooling of centers for statistical analyses

Centers will be pooled to geographic US region and explored as a sensitivity analysis.

Note: this sensitivity analysis will concern US region only. It is not applicable to Canada due to low number of patients to be enrolled there, in comparison with US.

2.5.7 Statistical technical issues

Not applicable.

3   INTERIM ANALYSIS

One IA of primary composite endpoint is planned for the purpose of overwhelming efficacy to provide an early evaluation of Toujeo effectiveness, when approximately 1800 patients (54.2% information fraction of targeted sample size) have completed their 6-month (Day 180) visit.

Using a group sequential approach with a Gamma (-3) alpha-spending function and an overall two-sided alpha-level of 0.05, the two-sided nominal significance level is 0.010 at IA, and 0.046 at final analysis.

If the targeted sample size (1800) cannot be reached, it could be considered to include a lower number of patients, however not below 1700. In this case, the nominal significance level to be spent at interim and at final 6-month analyses would be recalculated accordingly before the database lock for IA.

Overwhelming efficacy will be shown at IA if the efficacy boundary is met, ie, if the two-sided p-value for comparison of primary composite endpoint incidence is below 0.010, corresponding to an estimated difference of 5.8% in favor of Toujeo between percentages of patients reaching the primary composite endpoint.

The corresponding conditional power to show overwhelming efficacy at IA under the alternative hypothesis is of 45%.

For the IA, analyses will be performed by an independent statistician, external to the Sponsor. Results will be delivered directly to an independent Data Monitoring Committee (DMC) for review. a DMC meeting will be organized for this IA.

- Scenario 1: If overwhelming efficacy is shown:
    - The recruitment would be stopped (total sample size at that time estimated to be between 2600 and 3000 patients),
    - The patients already enrolled would continue the study as planned.
    - The results will be disclosed to the Sponsor and a communication plan to protect clinical data integrity is detailed in the DMC charter.
- Scenario 2: If overwhelming efficacy is not shown (probability of 55% under the alternative hypothesis):
    - The study would continue as planned, up to the final 6-month and 12-month analyses.
    - The Sponsor would not have access to any unblinded data.
    - No results would be communicated.

4   DATABASE LOCK

The database is planned to be locked at approximately 2 months after last patient last visit.

Prior to the final analysis of the 12 months, database will be locked:

- For the need of IA, when the 1800 patients have completed their 6-month visit, and their data have been cleaned and reconciled,

- At 6 months, when 6-month efficacy and safety data are available for all patients.

For the 6-months analysis, efficacy and safety data will be summarized and analyzed. Those CRF forms, which are not related to any specific study visits, will not be locked until the end of 12 months.

5   SOFTWARE DOCUMENTATION

All summaries and statistical analyses will be generated using R and SAS Version 9.3 or higher.

6   REFERENCES

1. Toujeo Real World US Insulin Naïve Trial. CLINICAL TRIAL PROTOCOL. A randomized, open-label, parallel group real world pragmatic trial to assess the clinical and health outcomes of Toujeo compared to commercially available basal insulins for initiation of therapy in insulin naïve patients with uncontrolled type 2 diabetes mellitus.

2. American Diabetes Association Workgroup on Hypoglycemia. Defining and reporting hypoglycemia in diabetes. Diabetes Care. 2005;28:1245-9.

3. Bradley C. The diabetes treatment satisfaction questionnaire: (DTSQ). In: Bradley C, ed. Handbook of Psychology and Diabetes: a guide to psychological measurement in diabetes research and practice. Chur, Switzerland: Harwood Academic Publishers. 1994;111-32.

4. Bradley C, Plowright R, Stewart J, Valentine J, Witthaus E. The Diabetes Treatment Satisfaction Questionnaire change version (DTSQc) evaluated in insulin glargine trials shows greater responsiveness to improvements than the original DTSQ. Health Qual Life Outcomes. 2007;5:57.

5. Karve S, Cleves MA, Helm M, Hudson TJ, West DS, Martin BC. Good and poor adherence: optimal cut-point for adherence measures using administrative claims data. Curr Med Res Opin. 2009;25(9):2303-10.

6. Lipkovich I, Dmitrienko A, Denne J, Enas G. Subgroup identification based on differential effect search-a recursive partitioning method for establishing response to treatment in patient subpopulations. Stat Med. 2011;30(21):2601-21.