NCT04075513

STATISTICAL ANALYSIS PLAN

A 12-week randomized controlled trial to compare TOUJEO® and TRESIBA® in terms of glucose values in target range and variability during continuous glucose monitoring (CGM) in patients with type 1 diabetes mellitus

HOE901-U300-LPS14947

STATISTICIAN: [redacted]

STATISTICAL PROJECT LEADER: [redacted]

DATE OF ISSUE: 01-Jul-2021

Total number of pages: 64

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

ADA:	American Diabetes Association
AE:	Adverse Event
AESI:	Adverse Event of Special Interest
ANCOVA:	Analysis of Covariance
ATC:	Anatomical Therapeutic Chemical
AUC:	Area Under the Curve
CDG:	Customised Drug Grouping
CGM:	Continuous Glucose Monitoring
CI:	Confidence Interval
CV:	Coefficient of Variation
DTSQc:	Diabetes Treatment Satisfaction Questionnaire (Change)
DTSQs:	Diabetes Treatment Satisfaction Questionnaire (Status)
eCRF:	Electronic Case Report Form
EOS:	End of Study
GMI:	Glucose Management Indicator
HLT:	High-Level Term
ICF:	Informed Consent Form
IMP:	Investigational Medicinal Product
IRT:	Interactive Response Technology
ITT:	Intent-to-Treat
LLT:	Lower-Level Term
LS Means:	least squares means
MAR:	Missing at Random
MDI:	Multiple Daily Injections
MedDRA:	Medical Dictionary for Regulatory Activities
MI:	Multiple Imputation
MNAR:	Missing not at Random
NIMP:	Non-investigational Medicinal Product
PRO:	Patient-Reported Outcomes
PT:	Preferred Term
SAE:	Serious Adverse Events
SAP:	Statistical Analysis Plan
SD:	Standard Deviation
SE:	Standard Error
SMPG:	Self-Monitoring of Plasma Glucose
SOA:	Schedule of Activities
SOC:	System Organ Class
TEAE:	Treatment-Emergent Adverse Event
WHO-DD:	World Health Organization-Drug Dictionary

1 OVERVIEW AND INVESTIGATIONAL PLAN

1.1 STUDY DESIGN AND RANDOMIZATION

This is a multicenter, randomized, active-controlled, parallel-group, 12-week open-label study to compare the efficacy of TOUJEO® (HOE901-U300) with TRESIBA® using 20-day continuous glucose monitoring (CGM) glucose profiles at Week 12 in patients with type 1 diabetes mellitus.

Patients satisfying the study inclusion/exclusion criteria (including CGM performance requirements) will be randomized via an Interactive Response Technology (IRT) in a 1:1 ratio to 1 of the 2 treatment groups and treated for 12 weeks. Randomization will be stratified by HbA1c at screening (<8.0% versus ≥8.0%).

Approximately 338 patients will be randomly assigned to study intervention for an estimated total of 131 evaluable patients per treatment within approximately 50 sites.

1.2 OBJECTIVES

1.2.1 Primary objectives

The primary objective of this study is to demonstrate the non-inferiority of insulin glargine 300 U/mL (Toujeo®) in comparison to insulin degludec 100 U/mL (Tresiba®) on glycemic control and variability in patients with diabetes mellitus.

1.2.2 Secondary objectives

The secondary objectives are

- To evaluate the glycemic control and variability parameters in each treatment group at Week 12 using CGM
- To evaluate the safety of insulin glargine 300 U/mL in comparison to insulin degludec 100 U/mL

1.2.3 Tertiary/exploratory objectives

The tertiary objective is to further explore the glycemic control and variability parameters in each treatment group at Week 12 using CGM.

1.3 DETERMINATION OF SAMPLE SIZE

The sample size calculation is based on the percent time spent in glucose range of ≥70 to ≤180 mg/dL (≥3.9 to ≤10 mmol/L) at Week 12, assessed from CGM measurements obtained from Weeks 10 to 12.

For this criterion of percent time in range, there is no predefined non-inferiority margin. A relative non-inferiority margin of 10% is considered.

Let m1 and m0 be the true means for the Toujeo® and Tresiba® groups, respectively.

The non-inferiority null hypothesis would be H0: m1-m0 ≤-0.1*m0 or H0: m1-0.9*m0 ≤ 0 and the alternative hypothesis would be H1: m1=m0.

Toujeo would be considered non-inferior to Tresiba if the lower limit of the 95% confidence interval (CI) for the adjusted difference estimate of m1-0.9*m0 at Week 12 is > 0.

To ensure 90% power, the sample size should satisfy:

$$1.96 - \frac{(0.1 * m0)}{\sqrt{\frac{1.81 * SD^2}{n}}} = -1.282$$

where -1.282 is the 10th percentile of a standard normal distribution and standard deviation (SD) the common SD, which leads to:

$$n = \frac{(1.96 + 1.282)^2 * 1.81 * SD^2}{(0.1 * m0)^2}$$

Assuming a common SD of 14.7%, an average percent time in range in the Tresiba treatment group of 56% (value from LPS14587), no true difference between both arms under H1, and a relative non-inferiority of 10%, a sample size of 131 evaluable patients per treatment group would provide at least 90% power to show non-inferiority of Toujeo with respect to Tresiba on the percentage of time plasma glucose within the range of ≥70 to ≤180 mg/dL at Week 12.

With the assumption of a non-evaluability rate of 22%, 338 patients (169 per treatment group) will be randomized.

1.4 STUDY PLAN

The study consists of a 1- or 2-week screening period, followed by a 4-week run-in period, and a 12-week randomized treatment period.

During the first 2 weeks of the run-in period (Weeks -4 and -3), patients will undergo treatment stabilization of their current basal and mealtime insulin treatment and will be trained and tested on how to perform study-related procedures. During the (-3) – (-1) weeks of run-in, patients will wear a CGM device, blinded to both patients and Investigators. During these 20 consecutive days, CGM readings will be collected as CGM assessment at baseline.

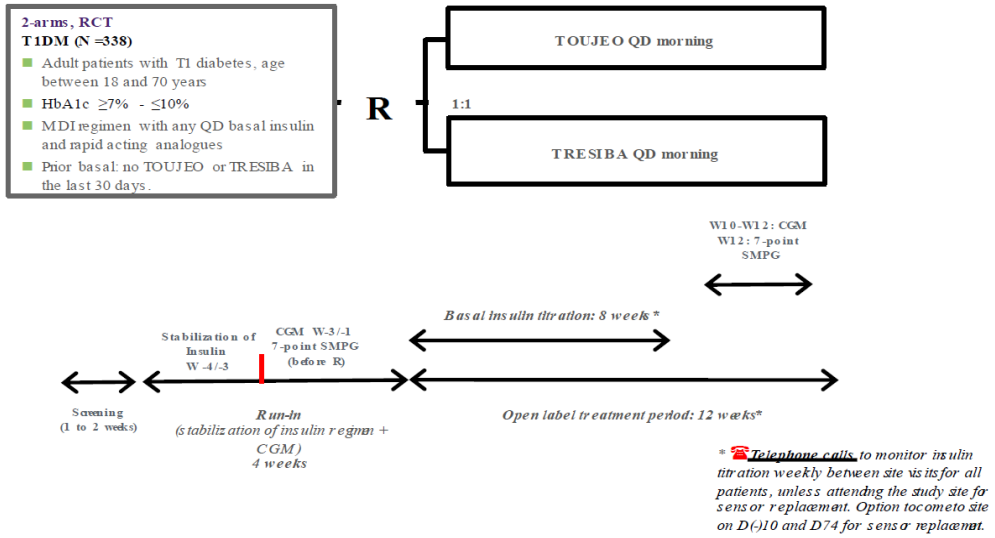
During the randomized treatment period, CGM readings will also be collected in the same fashion as in the run-in period during Weeks 10-12. This will be the CGM assessment at Week 12.

During the entire study starting after the screening period, patients will perform Self-Monitoring of Plasma Glucose (SMPG) as specified in the study flow chart. The SMPG readings will be used to guide insulin dose titration for optimal glycemic control.

At the end of the study, the patients will be followed-up for between 2 to 4 days after the last Investigational Medicinal Product (IMP) dose to collect post-treatment safety information.

Figure 1 is a representation of the study schema.

Figure 1 - Graphical study design



Abbreviations: CGM = continuous glucose monitoring, HbA1c = glycated hemoglobin, , MDI = multiple daily injection, R = randomized, SMPG = self-monitoring of plasma glucose, T1 diabetes = type 1 diabetes.

The schedule of activities (SOA) for the study is presented in Section 1.3 of the protocol.

1.5 MODIFICATIONS TO THE STATISTICAL SECTION OF THE PROTOCOL

Following the updated version of '*Clinical Targets for Continuous Glucose Monitoring Data Interpretation: Recommendations From the International Consensus on Time in Range'* in 2019 (1), some exploratory endpoints were added related CGM

- Mean glucose
- Glucose Management Indicator (GMI) (2) as defined by GMI (%) = 3.31+0.02392*(mean glucose in mg/dL)
- Percent time and mean hours per day in the glucose range of ≥54 to <70 mg/dL (≥3.0 to <3.9 mmol/L) (on all-time and during the night [00:00 to 05:59]),
- Percent time and mean hours per day in the glucose range of >180 to ≤250 mg/dL (>10 to ≤13.9 mmol/L) (on all-time and during the night [00:00 to 05:59]),

Following a misprint in '*International Consensus on Use of Continuous Glucose Monitoring*' in 2017 (3) , corrected in CGM consensus in 2019, an update of threshold for glucose within-day CV was done: ≤ 36 % instead of < 36%

For consistency between CGM endpoints and safety endpoints, the following endpoints related to hypoglycemia were added:

- Hypoglycemia (symptomatic or asymptomatic) documented by a measured plasma glucose concentration of <3.9 mmol/L (< 70 mg/dL) and ≥3.0 mmol/L (≥54 mg/dL), on all-time and during the night [00:00 to 05:59]

As per FDA & EMA guidelines provided during COVID-19 pandemic situation, several considerations (new population set call "Population without trial impact (disruption) due to COVID-19" and sensitivity analyses) were considered during the Statistical Analysis Plan (SAP) finalization.

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 baseline value is the last available value obtained up to the date and time of randomization or up to the time of first injection of the treatment if the measurement occurred on the same day as randomization.

All baseline safety and efficacy parameters (apart from those listed below) are presented along with the summary statistics in the safety and efficacy sections (Section 2.4.5 and Section 2.4.4).

Demographic characteristics

Demographic variables are:

- Gender (Male, Female),
- Age in years at date of informed consent (quantitative and qualitative variable: from 18 – 64 and 65 years and over),
- Weight in kg (quantitative),
- BMI in kg/m^2 (quantitative and qualitative variable: <25, ≥25 < 30, ≥30 <35, and ≥35 kg/m^2),
- Randomization strata of screening HbA1c categories (<8.0 and ≥8.0%).

Medical or surgical history

Medical (or surgical) history includes any past and/or concomitant diseases or past surgeries.

This information will be coded using the version of Medical Dictionary for Regulatory Activities (MedDRA) currently in effect at Sanofi at the time of database lock.

Disease characteristics at baseline

Specific disease history includes:

- Time from diagnosis to randomization (years) and age at diagnosis (years),
- History of gestational diabetes (yes, no),
- Time from first intake of basal insulin analog and of rapid acting insulin analog to randomization (years),
- Screening basal insulin analog (treatment preferred name according to the World Health Organization-Drug Dictionary [WHO-DD]). Distinction between insulin glargine U100 and insulin glargine U300 will be provided,

- Screening basal insulin dose, as reported on screening visit and in case of several lines reported the last available will be used (total daily dose in IU, total daily dose adjusted for body weight, in IU/kg),
- Screening mealtime insulin analog (Insulin glulisine, Insulin lispro, Insulin aspart),
- Screening mealtime insulin dose, as reported on screening visit and in case of several lines reported the last available will be used (total daily dose in IU, total daily dose adjusted for body weight, in IU/kg),
- Screening total daily insulin dose, ie, the day of screening visit (basal + mealtime) in IU and adjusted for body weight in IU/kg,
- Run-in basal insulin analog (treatment preferred name according to the WHO-DD). Distinction between insulin glargine U100 and insulin glargine U300 will be provided,
- Run-in basal insulin dose (total daily dose [in IU], total daily dose adjusted for body weight [in IU/kg]), as mean of previous basal daily dose during the 3 last days prior to the first IMP (or prior to randomization for not treated patients), in case of less than 3 days, the last dose will be used,
- Run-in mealtime insulin analog (Insulin glulisine, Insulin lispro, Insulin aspart),
- Run-in mealtime insulin dose (total daily dose, total daily dose adjusted for body weight) as mean of previous mealtime insulin total daily dose during the 3 last days prior to the first IMP (or prior to randomization for not treated patients), in case of less than 3 days, the last dose will be used. Of note, this will be considered as baseline mealtime insulin dose,
- Run-in total daily insulin doses (basal + mealtime) (total daily dose, total daily dose adjusted for body weight),
- Diabetes complications (at least one, Diabetic Retinopathy, Non-Proliferative Diabetic Retinopathy, Proliferative Diabetic Retinopathy, Diabetic Neuropathy, Diabetic Nephropathy),
- Method of mealtime insulin titration (Pattern of postprandial plasma glucose results of SMPG from the prior 3 days or Carbohydrate content of the meal [carb counting]).

Vital signs

Vital signs at baseline are:

- Systolic Blood Pressure (mmHg),
- Diastolic Blood Pressure (mmHg),
- Heart Rate (Beats/min).

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

2.1.2 Prior or concomitant medications

All medications taken since enrollment and until the end of the study are to be reported in the case report form.

Non-antidiabetic and antidiabetic medications will be considered separately.

All medications will be coded using the WHO-DD version currently in effect at Sanofi at the time of database.

- Prior medications are those the patient used prior to first injection of IMP intake. Prior medications can be discontinued before first administration or can be ongoing during treatment phase.
- Concomitant medications are any treatments received by the patient concomitantly to the IMP from first injection of IMP to the last injection of IMP + 2 days (0 day for anti-diabetic therapy). A given medication can be classified both as a prior medication and as a concomitant medication.
- Post-treatment medications are those the patient took in the period running from the 3rd day (1 day for anti- diabetic therapy) after last IMP intake up to the end of the study.

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

2.1.3 Efficacy endpoints

Observation period for efficacy endpoints

The 12–week randomized period for efficacy variables is defined as the time from randomization until Week 12 (Visit 18) for randomized patients, regardless of study treatment discontinuation.

CGM data

A patient is eligible for randomization if a minimum of 10 days (not necessarily consecutive) of useable CGM data are generated during 20 consecutive days at run-in period.

At Week 12, CGM is evaluable if there is at least 10 days of useable CGM data (not necessarily consecutives) within the analysis time windows (see Section 2.5.4) at Week 12 evaluation.

Useable CGM data for a day/a 24-hour CGM profile are defined as having:

- At least 80% time of records per 24 hours (ie, 231 data points among the 288 recorded in 24 hours).
- No gap (missing data) lasting for a total of ≥ 2 hours per 24 hours besides the warm-up time for the sensor after insertion.

For CGM endpoints, details on derivations are available in Section 2.5.

2.1.3.1 Primary efficacy endpoint(s)

The primary efficacy endpoint is the percent time in the glucose range of ≥70 to ≤180 mg/dL (≥3.9 to ≤10 mmol/L) at Week 12, obtained using CGM (for derivation see Section 2.5.1).

For any gap of less than 2 hours between two available CGM measurements, linear interpolation will be performed to impute the unobserved values at the time points (Section 2.5.1).

2.1.3.2 Secondary efficacy endpoint(s)

To address the secondary objective of evaluating the glycemic control and variability parameters in each treatment group, efficacy endpoints will be assessed using CGM, HbA1c, and FPG.

At Week 12 the following main secondary endpoints will be assessed using CGM (see Section 2.5.2.1).

- Glucose total coefficient of variation (CV) (%), main secondary endpoint,
- Glucose within-day and between-day CV (%).

Other secondary endpoints to be assessed are:

Based on central lab

- Change from baseline to Week 12 in HbA1c (% and mmol/mol),
- Change from baseline to Week 12 in FPG (mmol/L and mg/dL).

Based on CGM

- Percent time and mean hours per day with glucose <70 mg/dL (on all time [00:00 to 23:59] and during the night [00:00 to 5:59] obtained from CGM,
- Percent time and mean hours per day with glucose >180 mg/dL obtained from CGM.

2.1.3.3 Tertiary/exploratory efficacy endpoint(s)

To address the tertiary objective of exploring the glycemic control and variability parameters in each treatment group at Week 12 using CGM, the following endpoints will be assessed:

- Mean hours per day in the glucose range of ≥70 to ≤180 mg/dL (≥3.9 to ≤10 mmol/L),
- Percent time and mean hours per day in the glucose range of ≥70 to ≤140 mg/dL (≥3.9 to ≤7.8 mmol/L) (on all-time and during the night [00:00 to 05:59]),
- Nocturnal [00:00 to 05:59] percent time in glucose range of ≥70 to ≤180 mg/dL (≥3.9 to ≤10 mmol/L),
- Day time [06:00 to 23:59] percent time in glucose range of ≥70 to ≤180 mg/dL (≥3.9 to ≤10 mmol/L),
- Percent time and mean hours per day with glucose <54 mg/dL (<3.0 mmol/L) (on all time and during the night [00:00 to 5:59]),

- Percent time with glucose <70 mg/dL (<3.9 mmol/L) within the 12 hours after basal insulin injection,
- Percent time and mean hours per day with glucose >250 mg/dL (>13.9 mmol/L),
- Mean glucose profile over a 24-hour period,
- 24-hour Glucose Area Under the Curve (AUC mg/dL*minutes), derived using the trapezoidal rule. The derivation will be adjusted to be an hourly average (Section 2.5.2.1),
- Mean glucose,
- GMI,
- Percent time and mean hours per day in the glucose range of ≥54 to <70 mg/dL (≥3.0 to <3.9 mmol/L) (on all-time and during the night [00:00 to 05:59]),
- Percent time and mean hours per day in the glucose range of >180 to ≤250 mg/dL (>10 to ≤13.9 mmol/L) (on all-time and during the night [00:00 to 05:59]).

Other tertiary/exploratory endpoints than those related to CGM:

- Daily insulin doses (total daily insulin, basal insulin, mealtime insulin, ratio of basal to mealtime insulin dose) in units and U/kg body weight at each visit. Change at Week 8 and Week 12 from baseline will be derived.
- Change from baseline to Week 12 in patient-reported outcomes (treatment satisfaction, sleep quality and quantity, and productivity) (for detail on PROs, see Section 2.1.7).
- 7-point SMPG profiles (pre-prandial and 2-hour postprandial plasma glucose at breakfast, lunch, dinner, and bedtime) at Week 12 and change in 7-point SMPG profiles from baseline to Week 12.

2.1.4 Safety endpoints

To address the secondary objective of evaluating the safety of insulin glargine 300 IU/mL in comparison to insulin degludec 100 IU/mL, the following endpoints will be observed

- Number of patients with adverse events (see Section 8.3 of the protocol).
- Number of patients with at least one hypoglycemic event from baseline to Week 12.
- Number of hypoglycemic events per participant-year from baseline to Week 12.

The safety analysis will be based on the reported adverse events (AEs) and other safety information, such as vital signs. The specific safety parameter of interest for insulin treatment is hypoglycemia.

Observation period

For all safety data, the observation period will be divided as follows:

- The pre-treatment phase is defined as the time from when the patients sign the informed consent to first injection of IMP (excluded).
- The on-treatment phase is defined as the time from the first injection of IMP (included) up to 2 days after the last injection of IMP.
- The post-treatment phase is defined as the time after the on-treatment phase until the end of the study (Follow up Visit).

On-study observation period is defined as the time from start of treatment until the end of the study (defined as last protocol planned visit or the resolution/stabilization of all serious adverse events (SAEs) and AEs of specific interest (AESI) as defined in the protocol).

2.1.4.1 Adverse events variables

All AEs and SAEs will be collected from the signing of the informed consent form (ICF) until the follow-up visit (defined as patient's last visit planned per protocol or the resolution/stabilization of all SAEs and AESI). All events will be managed and reported in compliance with all applicable regulations and included in the final clinical study report.

Adverse event observation period

- Pre-treatment AEs are AEs that developed or worsened or became serious from the signed informed consent date up to first administration of IMP
- Treatment-emergent adverse events (TEAEs) are AEs that developed or worsened or became serious during the on-treatment period
- Post-treatment AEs are AEs that developed or worsened or became serious during the post-treatment period

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

AESI include the following terms, as defined by the protocol:

- Symptomatic overdose: Symptomatic overdose (serious or non-serious) with IMP and/or non-investigational medicinal product (NIMP) will be identified through a specific question on the AE page in the eCRF.
- Pregnancy: Pregnancy occurring, either in a female subject or in a female partner of a male subject entered in the study will be identified through a dedicated AE page in the e-CRF.

2.1.4.2 Deaths

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

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

2.1.4.3 Laboratory safety variables

Clinical laboratory data consist of blood analysis, including hematology, clinical chemistry, and serum pregnancy tests.

Blood samples for clinical laboratories will be taken at Visit 1 (screening, Week -6/-5). Additional urine samples will be collected on Visit 6 (randomization) and 18 (Week 12) in WOCBP for urine pregnancy test (to be confirmed by a serum test, if positive). The laboratory parameters will be classified as follows:

- **Hematology (Conventional Units and SI units)**:
  - Platelet count (×10^3/uL and GI/L),
  - Hemoglobin (g/dL and g/L),
  - Hematocrit (%),
  - Red blood cell count (×10^6/uL and TI/L),
  - White blood cell count × x10^3/uL and GI/L).
- **Clinical chemistry (Conventional Units and SI units)**:
  - Uric acid (mg/dL and µmol/L),
  - Potassium (mEq/L and mmol/L),
  - Aspartate aminotransferase/Serum glutamicoxaloacetic transaminase (AST/SGOT, U/L),
  - Alanine aminotransferase/Serum glutamicpyruvic transaminase (ALT/SGPT, U/L),
  - Total bilirubin (in case of values above the normal range, differentiation in conjugated and nonconjugate bilirubin mg/dL and umol/L),
  - Creatinine (mg/dL and umol/L),
  - Sodium (mEq/L and mmol/L),
  - Estimated creatinine clearance (Glomerular Filtration Rate, MDRD, mL/min/1.73m^2),
  - Blood urea nitrogen (BUN, mg/dL, and mmol/L),
  - Alkaline phosphatase (U/L).

- **Other tests (Conventional Units and SI units):**
  - C peptide (ng/mL and pmol/L),
  - Serum β-human chorionic gonadotropin (for women of childbearing potential at screening, mIU/mL),
  - Urine hCG pregnancy test for WOCBP at Visit 6 and 18 (to be confirmed with serum test, if positive).

2.1.4.4 Vital signs variables

Vital signs will be assessed according to the defined visit window (see Section 2.5.4).

The following vital signs will be assessed:

- Weight (Kg),
- Systolic blood pressure (mmHg) in sitting position,
- Diastolic blood pressure (mmHg) in sitting position,
- Heart rate (beats/min) in sitting position.

2.1.4.5 Electrocardiogram variables

Not applicable to the study.

2.1.4.6 Hypoglycemia events

Hypoglycemia events will be evaluated during pre-treatment and on-treatment period (see Section 2.1.4).

All reported hypoglycemia events will be classified according to American Diabetes Association (ADA) classification as follows:

- **Severe hypoglycemia**: an event that 'Required Assistance Because Subject Was Not Capable of Helping Self'.
- **Documented symptomatic hypoglycemia**: an event during which typical symptoms are accompanied by a measured plasma glucose concentration less than 3.9 mmol/L (<70 mg/dL).
- **Asymptomatic hypoglycemia**: an event with no symptoms of hypoglycemia but a documented measured plasma glucose concentration less than 3.9 mmol/L (<70 mg/dL).

In addition, the following events will be also considered:

- Documented symptomatic hypoglycemia with measured plasma glucose concentration less than 3.0 mmol/L (<54 mg/dL).
- Asymptomatic hypoglycemia with measured plasma glucose concentration less than 3.0 mmol/L (<54 mg/dL).

- Hypoglycemia (symptomatic or asymptomatic) documented by a measured plasma glucose concentration less than 3.9 mmol/L (<70 mg/dL).
- Hypoglycemia (symptomatic or asymptomatic) documented by a measured plasma glucose concentration of <3.9 mmol/L (< 70 mg/dL) and ≥3.0 mmol/L (≥54 mg/dL).
- Hypoglycemia (symptomatic or asymptomatic) documented by a measured plasma glucose concentration less than 3.0 mmol/L (<54 mg/dL).
- Severe and/or symptomatic hypoglycemia documented by plasma glucose concentration less than 3.9 mmol/L (<70 mg/dL).
- Severe and/or symptomatic hypoglycemia documented by plasma glucose concentration less than 3.0 mmol/L (<54 mg/dL).

In addition to anytime of the day (00:00 to 23:59), hypoglycemia events will be evaluated at the following time periods:

- Nocturnal hypoglycemia: any hypoglycemic event that occurs between 00:00 and 05:59.
- Diurnal hypoglycemia: any hypoglycemic event that occurs between 06:00 and 23:59.

Note: CGM values will not be considered for identifying hypoglycemia.

2.1.5 Pharmacokinetic variables

Not applicable to the study.

2.1.6 Pharmacodynamic/genomics endpoints

Not applicable to the study.

2.1.7 Quality-of-life endpoints

Several patient-reported outcomes (PROs) endpoints will be evaluated using the following questionnaires: Diabetes Treatment Satisfaction Questionnaire Status (DTSQs) and Change (DTSQc) versions, MOS Sleep, and Work Productivity and Activity Impairment Questionnaire (WPAI).

Details on handling of missing data for PROs are available in Section 2.5.2.4.

2.1.7.1 DTSQs

The DTSQs is used to assess patient satisfaction with treatment and patient perception of blood glucose control; it is a validated questionnaire comprised of the 8 questions answered on a Likert scale from 0 to 6. It will be completed at baseline and Week 12.

Responses to questions 1, 4, 5, 6, 7, and 8 will be summed to produce a Total Treatment Satisfaction score ranging from 0 (no satisfaction) to 36 (high satisfaction with treatment).

The 2 items of 'perceived frequency of hyperglycemia' (question 2) and 'perceived frequency of hypoglycemia' (question 3) are scored separately ranging from 0 (none of the time) to 6 (most of the time).

2.1.7.2 DTSQc

The Diabetes Treatment Satisfaction Questionnaire Change Version (DTSQc) was developed from the original DTSQ to evaluate the change in treatment satisfaction at a specific time point. It contains 8 questions answered on a Likert scale from -3 to 3. The DTSQc instructions and response options differ from those of the DTSQs to produce measures of relative change in satisfaction rather than measures of absolute satisfaction. The DTSQc will be completed at Week 12.

Responses to questions 1, 4, 5, 6, 7, and 8 will be summed to produce a Total Treatment Satisfaction score ranging from -18 to 18. Positive scores are indicative of improvement in treatment satisfaction, whereas negative scores are indicative of deterioration in treatment satisfaction since start of the study.

The 2 items of 'perceived frequency of hyperglycemia' (question 2) and 'perceived frequency of hypoglycemia' (question 3) are scored separately ranging from -3 (much less of the time) to 3 (much more of the time), meaning negative scores indicate fewer problems with blood glucose levels and positive scores indicate more problems than before.

2.1.7.3 MOS Sleep

The MOS Sleep used to measure patient perception on key aspects of sleep, consisting of 12 questions that will be completed by the patients at baseline and Week 12. The MOS Sleep 12 items acute revised version with the 'past week' recall period will be used to collect information.

The 12-item MOS Sleep Scale assessment yields scores on six subscales: Sleep Disturbance (4-item subscale on sleep initiation problems, and sleep maintenance problems), Snoring (single-item subscale), Shortness of Breath or Headache (single-item subscale asking about the frequency of awakening with shortness of breath or with a headache), Sleep Adequacy (2-item subscale asking about frequency of awakening fresh and rested in the morning and getting the amount of sleep needed), Daily Somnolence (3-item subscale), and Sleep Quantity (single-item subscale representing the average number of hours slept).

The Sleep Problems Index is also calculated based on 9 of the items.

Responses to all but the open-ended sleep quantity item are transformed to scores on a 0-100 metric, with higher item scores reflecting more of the attribute implied by the scale name.

Scores on the multi-item subscales and indexes also range from 0 to 100.

MOS-sleep scores will be calculated automatically via the electronic system provided by the owner of this questionnaire.

2.1.7.4 WPAI

The WPAI SHP assesses the amount of absenteeism, presenteeism and daily activity impairment attributable to a specific health problem (WPAI:SHP version V2.0). An adapted version of the WPAI:SHP will be used with the term PROBLEM replaced by DIABETES, as requested by the author.

The WPAI:SPH consists of 6 questions which will be completed by the patients at baseline and Week 12.

The 6 questions will be used to derive 4 scores (see Section 2.5.2.4) and scores are multiplied by 100 to express in percentages.

2.1.8 Health economic endpoints

Not applicable to the study.

2.2 DISPOSITION OF PATIENTS

Screened patients are defined as any patients who signed the informed consent.

Randomized patients consist of all patients with a signed informed consent form who have been allocated a treatment kit based on a randomization process, regardless of whether the treatment kit was used.

Patient study status

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 patients
- Screen failure patients and reasons for screen failure
- Treated but not randomized
- Randomized patients
- Randomized but not treated patients
- Randomized and treated patients
- Patients who completed the 12-week on-treatment period (patients who have performed Visit 18, who did not permanently discontinue)
- Patients who discontinued IMP by main reason for permanent treatment discontinuation
- Patients who did not complete the study
- Patients who discontinued the study by main reason for study discontinuation
- Status at last study contact (Alive/Death)
- Lost to follow-up

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 for each treatment group.

Reasons for treatment discontinuation will be supplied in tables giving numbers and percentages by treatment group.

Lost to follow-up will be based on the End of Study [EOS] page as 'Other' with specify 'Lost to Follow Up'.

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.

Same summary as above will be displayed for any critical or major deviation due to COVID-19 pandemic. Patients excluded from Population without trial impact (disruption) due to COVID-19 will be listed.

The patients with any visit impacted by COVID-19 crisis will be summarized in a separate table giving numbers and percentages by treatment group.

Additionally, the analysis populations for safety and efficacy, will be summarized in a table by number of patients in the randomized population:

- Efficacy population: intent-to-treat (ITT) population, per-protocol population and Population without trial impact (disruption) due to COVID-19,
- Safety population.

2.2.1 Randomization and drug dispensing irregularities

Randomization and drug-dispensing irregularities occur whenever:

- A randomization is not in accordance with the protocol-defined randomization method, such as a) an ineligible patient is randomized, b) a patient is randomized based on an incorrect stratum, c) a patient is randomized twice, or d) in a dynamic randomization scheme the treatment assignment is, in fact, not random, due to a computer program error. OR
- A patient is dispensed an IMP kit not allocated by the protocol-defined randomization, such as a) a patient at any time in the study is dispensed a different treatment kit than as randomized (which may or may not contain the correct-as-randomized IMP), or b) a 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 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). Nonrandomized, treated patients will be described separately.

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

- IMP dispensed without randomization,
- IMP dispensed prior to randomization,
- Patient randomized twice,
- IMP kit number actually dispensed to the patient is different from the IMP kit number allocated,
- Wrong stratum of randomization,
- IMP dispensed without allocation at re-supply visit,
- Non-existing patient randomized by error.

2.3 ANALYSIS POPULATIONS

Patients who are dispensed 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 safety.

2.3.1 Efficacy populations

The primary efficacy analysis population will be the ITT population.

2.3.1.1 Intent–to-treat population

The ITT population consists of all randomized patients, irrespective of the treatment received. Randomized patients are patients who signed the ICF and with a treatment group allocated before the first IMP administration and recorded in the IRT database.

Patients will be analyzed in the treatment group to which they are randomized.

2.3.1.2 Per-protocol population

The per-protocol population (PP) is a subset of the ITT population which will comprise all randomized patients, who are exposed, who do not permanently discontinued IMP allocated by IRT, who completed the study as per protocol (as reported in the EOS page) and who do not present any major protocol deviations that potentially impact the primary efficacy results, as presented below:

- Major or critical deviation linked to inclusion criteria:

- From inclusion criterion I02: Patients without T1DM.
- From inclusion criterion I03: Not treated with multiple daily injections (MDI) using basal insulin analog once daily and rapid acting insulin analogs for at least one year
- From inclusion criterion I04: HbA1c <7% or >10%

- Major or critical deviation linked to exclusion criteria:
  - With E09: Patients having received basal insulin dose ≥0.6 U/kg body weight within 30 days prior to screening.
  - With E13: Patients who will need real time, flash or implantable CGM for routine care any time during study participation.
  - With E23: Inappropriate CGM use during run-in period evidenced by failure to obtain a minimum of 10 days (not necessarily consecutive) of useable CGM data by the end of run-in.
- Other major or critical deviation potentially impacting primary efficacy results:
  - Not taken only the treatment allocated by IRT (ie, wrong treatment taken at randomization or switch during the study)
  - Wrong stratum of randomization
  - Taken any prohibited medications for more than 10 consecutive days during the study as follows:
    - Any glucose-lowering agents other than IMP, the patient's current background basal insulin analog during the run-in period (which may not be changed during the study), or the patient's existing mealtime insulin (which may not be changed during the study). This includes oral or injectable glucose lowering agents, other type of basal insulin (eg, NPH), pre-mixed insulin, and human regular insulin
    - Insulin pump therapy.
    - Systemic glucocorticoids (topical or inhaled applications are allowed).
  - Patient who did not receive the CGM at Week 9, or for whom there are issues with uploading the CGM data at Week 12. This will be evaluated for patients who completed the treatment period and the study

Of note: in case of CGM performed as planned in the protocol (20-days during Week 10, Week 11 and Week 12) but not evaluable (ie, less 10 days, [not necessarily consecutive] of useable CGM data among the 20-days of CGM), the patient is not excluded from PP population.

2.3.1.3 Population without trial impact (disruption) due to COVID-19

The "Population without trial impact (disruption) due to COVID-19" is defined as patient without.

- Premature treatment discontinuation due to COVID-19
- Premature end of study due to COVID-19
- Critical or major protocol deviations due to COVID-19

For the primary efficacy endpoint, analysis based on Population without trial impact (disruption) due to COVID-19 will be performed as sensitivity analysis.

2.3.2 Safety population

The safety population is defined as all randomized patients who received at least 1 dose or part of a dose of the IMP. Patients in the safety population will be analyzed according to the actual treatment received. In the event that different treatments are administrated to the same patient, the patient is to be analyzed using the treatment that he/she is on for the most time.

In addition

- Non-randomized but treated patients will not be part of the safety population; however, their safety data will be presented separately
- Randomized patients for whom it is unclear whether they took the IMP will be included in the safety population as randomized

2.4 STATISTICAL METHODS

2.4.1 Demographics and baseline characteristics

Continuous data will be summarized using the number of available data, mean, standard deviation (SD), median, minimum, first quartile (Q1), third quartile (Q3), and maximum for each treatment group.

Categorical and ordinal data will be summarized using the number and percentage of patients in each treatment group.

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

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

For medical and surgical history, non-prespecified diagnoses will be summarized by treatment group according to the MedDRA dictionary. Incidence tables will present by SOC and HLT – sorted by SOC internationally agreed order and by decreasing frequency of HLT based on incidence in the overall treatment group – the number (n) and percentage (%) of patients experiencing an event. Multiple occurrences of the same event in the same patient will be counted only once in the tables. The denominator for computation of percentages will be the randomized population within each treatment group.

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

2.4.2 Prior or concomitant medications

The prior, concomitant and post-treatment medications will be presented for the ITT population separately for anti-diabetic medications and other (ie, non-anti-diabetic) medications.

2.4.2.1 Non-antidiabetic medications

Non antidiabetic medications will be summarized by treatment group according to the WHO-DD dictionary, considering the first digit of the anatomical therapeutic chemical (ATC) class (anatomic category) 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 table for prior non-antidiabetic medications will be sorted by decreasing frequency of ATC followed by all other therapeutic classes based on the overall incidence across treatment group. In case of equal frequency regarding ATCs (anatomic or therapeutic categories), alphabetical order will be used.

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

2.4.2.2 Antidiabetic medications

The table for prior anti-diabetic medications other than basal and mealtime insulin analogs (permitted or not permitted per protocol) will be summarized per pre-defined and standardized medication name based on the overall incidence across treatment group. In case of equal frequency, alphabetical order will be used.

The table for concomitant anti-diabetic medications other than basal and mealtime insulin analogs (permitted or not permitted per protocol) will be summarized per pre-defined and standardized medication name based on the incidence in the Toujeo group. In case of equal frequency, alphabetical order will be used.

The following predefined classes (customised drug grouping [CDG]) will be considered for prior and concomitant:

At least one non-insulin therapy

- Biguanides
- Sulfonylureas
- Glinides
- Thiazolidinedione

- DPP-4 inhibitors
- SGLT-2 inhibitors
- GLP1-RA
- Alpha-glucosidase inhibitors
- Other blood glucose lowering drugs

At least one insulin therapy different from insulin analog (even if not permitted in prior or concomitant)

- Insulin NPH
- Premixed insulin
- Short acting regular human insulin

The table for post anti-diabetic medications will be summarized per pre-defined class and standardized medication name based on the incidence in the Toujeo group. In case of equal frequency, alphabetical order will be used. Insulin therapy could include insulin analog.

A list of codes to be used in determining these pre-defined classes is in a separate operational document.

Of note, basal and mealtime insulin analogs as prior medications are described in Section 2.1.1

Frequency statistics including number of patients and percentage will be provided.

2.4.3 Extent of investigational medicinal product exposure and compliance

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

2.4.3.1 Extent of investigational medicinal product exposure

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

The duration of treatment exposure will be the total number of days of administration of the open-label IMP, regardless of unplanned intermittent discontinuations (see Section 2.5.3 for calculation in case of missing or incomplete data). The duration of IMP exposure will be calculated as:

- (date of the last open-label IMP injection – date of the first open-label IMP injection) + 1.

Duration of IMP exposure will be summarized descriptively as a quantitative variable (number, mean, SD, median, minimum, and maximum). In addition, duration of treatment exposure will also be summarized categorically by numbers and percentages for each of the following categories and cumulatively according to these categories 1 to 21 days, 22 to 42 days, 43 to 63 days, 64 to 84 days and >84 days.

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

2.4.3.2 Daily insulin doses

Daily insulin dose will be summarized descriptively (number, mean, SD, median, Q1, Q3, minimum, and maximum) at each visit (baseline, and post baseline: Week 8 and Week 12).

Basal insulin daily dose (U and U/kg) at baseline is defined as the actual starting IMP dose of basal insulin as reported by the Investigator into the e-CRF (first IMP intake).

Mealtime insulin (ie, rapid acting insulin analog) daily dose (U and U/kg) at baseline is defined as the actual mealtime total daily dose in run-in (see Section 2.1.1) .

At each postbaseline visit, the daily basal, mealtime, and total (basal plus mealtime) insulin doses (U and U/kg) will be calculated as the mean of total daily insulin doses (as reported in the CRF) collected in the week before the corresponding visit.

The change from baseline daily doses variable at Week 8 and Week 12 will also be summarized descriptively.

2.4.3.3 Compliance

A given administration will be considered noncompliant if the patient did not take the IMP as required by the protocol, meaning one injection per day. No imputation will be made for patients with missing or incomplete data.

The percentage of compliance for a patient will be defined as the number of days with at least one administration divided by the duration of IMP exposure.

Treatment compliance will be summarized descriptively as quantitative variables (number, mean, SD, median, minimum, and maximum) by treatment group for the safety population. In addition, numbers and percentages of patients whose compliance is <80% will be summarized by treatment group.

2.4.4 Analyses of efficacy endpoints

For each patient, the number of days with useable CGM data within the analysis time window (see Section 2.5.4) will be derived in the following categories: < 10 days, 10 days, 11 days, 12 days, 13 days, 14 days, ≥14 days at baseline and at Week 12. The categories could be updated based on real data.

2.4.4.1 Analysis of primary efficacy endpoint(s)

2.4.4.1.1 Primary efficacy analysis

The primary efficacy endpoint is the percent time in glucose range is defined as the proportion of CGM readings falling into the range of ≥70 to ≤180 mg/dL [≥3.9 to ≤10 mmol/L] at the Week 12 evaluation.

The primary efficacy endpoint will be analyzed using available data from the 12-week randomized period in ITT population via an analysis of covariance (ANCOVA) including the fixed categorical effects of treatment group (Toujeo versus Tresiba), randomization stratum of screening HbA1c (<8.0% versus ≥8.0%), as well as, the continuous fixed covariate of baseline percent time in range value.

Multiple imputation (MI) assuming missing at random (MAR) will be used to replace missing data of percent time in glucose range at Week 12 and at baseline. The MAR assumption means that the missingness depends only on observed variables and not on the data that would have been observed (in the present case, the percent time in glucose range that would have been observed).

Moreover, the imputation procedure will also depend on the observed pattern of missingness:

A) **Monotone Pattern**: Patients with missing baseline CGM and missing Week 12 CGM, or patients with baseline CGM but without a Week 12 value.

B) **Non-Monotone Pattern**: Patients with missing baseline CGM but with a Week 12 CGM. Baseline CGM evaluable being used as a criterion to allow the randomization, this pattern is not expected to be represented in the study.

In order to explore missing data patterns for the primary efficacy analysis, number and percentage of patients in each of the following categories will be presented by treatment group.

- Pattern 1: patients with missing baseline CGM and missing Week 12 CGM,
- Pattern 2: patients with baseline CGM but missing Week 12 CGM,
- Pattern 3: patients with missing baseline CGM but with Week 12 CGM,
- Pattern 4: patients with baseline CGM and Week 12 CGM

Once missing data pattern will be assessed, the three-step programming process described in Section 2.5.3 will provide baseline-adjusted least-squares means estimates at Week 12 for both treatment groups, as well as, the differences of these estimates, with their corresponding standard errors (SEs), 95% CI, and one-sided P-values (H0: m1 - 0.9 × m0 ≤ 0, with m1 = true mean of Toujeo and m0 = true mean of Tresiba).

Toujeo would be considered non-inferior to Tresiba if the lower limit of the 95% CI for the adjusted difference estimate of m1-0.9*m0 at Week 12 is > 0 (with m1=true means of Toujeo and m0=true means of Tresiba).

In addition to the non-inferiority testing that is the primary analysis, and only if non-inferiority of Toujeo in comparison to Tresiba on the percent time in the range of 70-180 mg/dL is demonstrated and if non-inferiority of Toujeo versus Tresiba on the glucose total CV is demonstrated (see Section 2.4.4.3 for multiplicity issues), superiority of Toujeo over Tresiba on the percent time in the range of 70-180 mg/dL will be tested.

The point estimate of difference in mean of percent time in range and the corresponding 95% CI will also be estimated using the following ESTIMATE statement in PROC GLM (in Step 2, See Section 2.5.3).

```
ESTIMATE "TOUJEO-TRESIBA" TREAT 1 -1
```

and the Step 3 will be done to obtain baseline-adjusted least-squares means estimates at Week 12 for both treatment groups as well as, the differences of these estimates, with their corresponding standard errors (SEs), 95% CI. The lower bound of the two-sided 95% CI for the adjusted difference estimate of m1 - m0 at Week 12 will be compared to 0 and subjected to a superiority test (if lower limit of 95% CI is >0).

The superiority comparison is considered as a secondary analysis of the primary efficacy variable.

Assumption assessment

The studentized residuals from the 1000 ANCOVA models described in Step 2 above with imputed data will be inspected and plotted against the predicted values, all in the same graph, for a visual assessment to detect any homoscedasticity.

2.4.4.1.2 Sensitivity analyses

First sensitivity analysis: MI on the PP population

A first sensitivity analysis will be performed on the PP population (see Section 2.3.1.2) with the same multiple imputation process as for primary analysis.

Second sensitivity analysis: MI under missing not at random (MNAR**) assumption on ITT population**

In addition, to investigate possible violations of the MAR assumption, a second sensitivity analysis with multiple imputation in conjunction with MNAR assumption based on patient status at Week 12 will be performed for ITT population.

Under the MNAR assumption, patients with study treatment discontinuation without any CGM done at Week 12 and patients with missing CGM endpoints at Week 12 due to other reasons (ie, non-evaluable CGM) will be imputed differently.

This MI process for the missing data can be carried out in two parts:

1. For patients who prematurely discontinue the study treatment, as it is not expected to have enough patient who discontinued with CGM at Week 12 done, hypothesis of return to baseline will be used. The imputation for Week 12 will be from the following model ~ N (BASE, SD) (BASE is the baseline value for the patient, SD estimated for all patients at baseline)
2. For patients who did not prematurely discontinue the study treatment, the multiple imputation method described as the primary efficacy analysis under MAR assumption will be used.

Missing values will be imputed 1,000 times. For each imputation, the two completed datasets from the two parts detailed above will be combined into one single completed dataset. Each completed dataset will then be analyzed though an ANCOVA model which will include fixed categorical effects of treatment group, randomization stratum screening HbA1c value [<8.0% versus ≥8.0%], as well as fixed continuous fixed covariate of baseline percent time in range value. The final results will be obtained by combining the least squares (LS Means) means and LS mean differences from these 1,000 analyses, using Rubin's formula.

Third sensitivity analysis: MI on Population without trial impact (disruption) due to COVID-19

A third sensitivity analysis will be performed on Population without trial impact (disruption) due to COVID-19 (see Section 2.3.1.3) with the same multiple imputation process as for primary analysis.

2.4.4.1.3 Subgroup analyses

Where appropriate for the size of a subgroup (at least 30 patients in each category), exploratory analyses will be performed on the primary endpoint to summarize the treatment effects across subgroups defined by the following baseline and screening factors:

- Duration of diabetes (<10/>= 10 years),
- Gender,
- Randomization strata of screening HbA1c categories (<8.0% versus ≥8.0%),
- BMI (<30 and ≥30 kg/m²).

Categories for subgroup analyses may be adapted according to data.

For each subgroup, the primary efficacy endpoint using available data from the 12-week randomized period in ITT population will be analyzed via an analysis of covariance (ANCOVA) including the fixed categorical effects of treatment group (Toujeo, Tresiba), randomization stratum of screening HbA1c (<8.0% versus ≥8.0%), the continuous fixed covariate of baseline percent time in range value, and adding the corresponding subgroup factor and subgroup factor-by-treatment interaction. LSmeans treatment difference (with the corresponding SEs and 95% CI) will be presented for each subgroup. The P-value for the interaction fixed term will be

obtained in order to evaluate whether the corresponding subgroup is behaving as a response modifier.

The ESTIMATE statement in PROC GLM, which reflects the subgroup analyses by baseline and screening factors, will be used. An example of SAS code for the subgroup analysis by gender is provided below:

Female group:

```
ESTIMATE "TOUJEO-TRESIBA" ARM 1 -1 SEX*ARM 1 0 -1 0;
```

Male group:

```
ESTIMATE "TOUJEO-TRESIBA" ARM 1 -1 SEX*ARM 0 1 0 -1;
```

Forest plot will be provided.

Further subgroup analyses may be performed if deemed necessary for the interpretation of results.

2.4.4.2 Analyses of secondary efficacy endpoints

Secondary efficacy endpoints analysis will be done on the ITT population, with no multiple imputation except for the main secondary endpoint 'Glucose total CV', for which a similar procedure as for the primary endpoint applies as it is part of the multiplicity procedure. No sensitivity analyses are planned for secondary endpoints.

2.4.4.2.1 CGM secondary endpoints

The CGM endpoints will be analyzed using an ANCOVA model

- Glucose total CV (main secondary endpoint)
- Glucose within-day and between-day CV
- Percent time and mean hours per day with glucose <70 mg/dL (on all time and during the night [00:00 to 05:59])
- Percent time and mean hours per day with glucose >180 mg/dL

The ANCOVA model to perform the analysis includes the fixed categorical effects of treatment group (Toujeo, Tresiba), randomization stratum of screening HbA1c (<8.0% versus ≥8.0%), as well as the continuous fixed covariate of baseline value for the corresponding endpoints.

For glucose total CV (main secondary endpoint), and if the non-inferiority on the primary endpoint is demonstrated, the non-inferiority will be tested. But, unlike TIR, for CV the lower the better. Therefore, the non-inf comparison will be based on:

H0: $m1 - 1.1 \times m0 \geq 0$

For the other secondary endpoints, LSmeans treatment difference (with the corresponding SEs and 95% CI) will be provided and the p-value (of the difference between the two treatments) will be provided for descriptive purpose only (see Section 2.4.4.4).

2.4.4.2.2 HbA1c

Change in HbA1c will be analyzed using an ANCOVA model including the fixed categorical effects of treatment group (Toujeo, Tresiba) as well as the continuous fixed covariate of baseline HbA1c value; LSmeans treatment difference (with the corresponding SEs and 95% CI) will be provided and p-value (of the difference between the two treatments) will be provided for descriptive purpose.

2.4.4.2.3 FPG

Change in FPG will be analyzed using an ANCOVA model including the fixed categorical effects of treatment group (Toujeo, Tresiba) and the randomization stratum of HbA1c at screening (<8.0%, ≥8.0%) as well as the continuous fixed covariate of baseline FPG value. LSmeans treatment difference (with the corresponding SEs and 95% CI) will be provided and p-value (of the difference between the two treatments) will be provided for descriptive purpose.

2.4.4.2.4 Subgroups analyses of secondary efficacy endpoints

Subgroup analyses will be performed on glucose within-day CV to summarize the treatment effects across subgroup defined by glucose within day CV≤36% vs. >36% at baseline. The ANCOVA model will be used, the randomization stratum of HbA1c at screening, CV subgroup at baseline, as well as the interaction of treatment group by CV subgroup at baseline are included as fixed effects. LS means treatment difference (with the corresponding SEs and 95% CI) will be presented for each subgroup. The p-value for the interaction fixed term will be obtained in order to evaluate whether the corresponding subgroup is behaving as a response modifier.

2.4.4.3 Multiplicity issues

To control the type I error, a hierarchical step-down testing procedure described by Hochberg and Tamhane (4) will be applied for the primary efficacy endpoint and the main secondary endpoint:

- Step 1 assesses non-inferiority of Toujeo versus Tresiba on the percent time in the range of 70-180 mg/dL,
- Step 2: only if Step 1 is demonstrated, then non-inferiority of Toujeo versus Tresiba on the glucose total CV will be tested with a relative non-inferiority margin of 10%,
- Step 3: only if Step 2 is demonstrated, then the difference between Toujeo and Tresiba on the percent time in the range of 70-180 mg/dL will be tested.

No multiplicity adjustment will be made on other secondary/other efficacy variables; 95% CI and p-values presented for these endpoints will be done for descriptive purpose only.

2.4.4.4 Additional efficacy analysis(es)

The tertiary/exploratory variables (see Section 2.1.3.3) will be summarized by treatment group, using descriptive statistics for continuous variables for ITT population, except for PROs.

A plot with the mean glucose profile over 24-hour period will be presented. For each patient, the data will be summarized for each 1-hour interval over 24-hour period, eg, [00:00, 01:00), [01:00, 2:00), …, [23:00, 24:00) taking the average of CGM readings from usable CGM days. Then, the mean profile will be calculated taking the average across all patients at each interval and plotted at the midpoint of each 1-hour interval.

For PRO, the analyses are described in Section 2.4.7.

Change from baseline in 7-point SMPG profiles (pre-prandial and 2-hour postprandial plasma glucose at breakfast, lunch and dinner, and at bedtime) to Week 12 will be described by treatment group using mean and SD.

A figure representing the mean (+/- SE) at baseline and Week 12 per time point as well as change from baseline at Week 12 will be produced by treatment group.

2.4.5 Analyses of safety data

The summary of safety results will be presented by treatment group classifying each patient in the group corresponding to the actual treatment received.

General common rules

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

- Safety data in patients who do not belong to the safety population (eg, exposed but not randomized) will be listed separately,
- The baseline value is defined as the last available value before the first injection of open-label IMP,
- The analysis of the safety variables will be essentially descriptive.

2.4.5.1 Analyses of hypoglycemia events

Incidence of patients with at least one hypoglycemia event

A frequency table showing the number and % of patients experiencing at least one hypoglycemic event will be presented for any hypoglycemia and for each type of hypoglycemia (see

Section 2.1.4.6), overall and according to time of occurrence (nocturnal [ie, 00:00 to 05:59], for any time of day and daytime [ie, 06:00 to 23:59] during the on-treatment period.

The Odds Ratio and its corresponding 95% CI for the Toujeo arm over the Tresiba arm for each hypoglycemic event (nocturnal and any time of the day) during the on treatment period will be estimated by a logistic regression model adjusted on randomization strata of screening HbA1c.

Number and rate of hypoglycemia event

The number and rate of hypoglycemic events (in patient-year of exposure) will be determined for any hypoglycemia and for each type of hypoglycemic event, overall and according to time of occurrence (nocturnal [ie, 00:00 to 05:59], any time of the day) during the on treatment period and in the pre-treatment period.

For each type of hypoglycemic event during the on treatment period, the rate ratio, and its corresponding 95% CI, for the Toujeo arm over the Tresiba arm will be estimated using an over-dispersed Poisson regression model adjusted on randomization strata of screening HbA1c.

2.4.5.2 Analyses of adverse events

Generalities

The primary focus of AE reporting will be on TEAEs. Pre-treatment and post-treatment AEs will be described separately.

If an adverse event date/time of onset (occurrence, worsening, or becoming serious) is incomplete, an imputation algorithm will be used to classify the adverse event as pre-treatment, treatment-emergent, or post-treatment. The algorithm for imputing date/time of onset will be conservative and will classify an adverse event as treatment emergent unless there is definitive information to determine if it is pre-treatment or post-treatment. Details on classification of adverse events with missing or partial onset dates are provided in Section 2.5.4.

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

Sorting within tables ensures the same presentation for the set of all adverse events within the observation period (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 Toujeo treatment group.

A listing of SAEs and listing of withdrawals due to AEs will be presented.

Analysis of all treatment-emergent adverse events

The following treatment-emergent adverse event (TEAE) summaries will be generated for the safety population.

- Overview of TEAEs, summarizing the number (%) of patients with any
  - TEAE
  - Serious TEAE
  - TEAE leading to death
  - TEAE leading to permanent treatment discontinuation
  - TEAE related to IMP
- The number (%) of patients experiencing TEAE(s) presented by PT, sorted by decreasing incidence of PT in the Toujeo treatment arm.
- All TEAEs 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.
- All TEAEs 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 treatment arm within each SOC.
- All TEAEs related to IMP 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 treatment arm within each SOC.

Analysis of all treatment emergent serious adverse event(s)

- All treatment-emergent serious adverse events 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 serious adverse events by primary SOC and PT, showing the number (%) of patients with at least 1 serious treatment-emergent adverse event, sorted by the internationally agreed SOC order and by decreasing incidence of PTs in Toujeo treatment arm within each SOC.
- All treatment-emergent serious adverse events related to IMP, by primary SOC and PT, showing the number (%) of patients with at least 1 treatment-emergent serious adverse event, sorted by the internationally agreed SOC order and by decreasing incidence of PTs in Toujeo treatment arm within each SOC.
- All treatment-emergent serious adverse events leading to treatment discontinuation, by primary SOC and PT, showing the number (%) of patients sorted by the internationally agreed SOC order and by decreasing incidence of PTs in Toujeo treatment arm within each SOC.

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

- All treatment-emergent adverse events leading to treatment discontinuation, by primary SOC and PT, showing the number (%) of patients sorted by the internationally agreed SOC order and by decreasing incidence of PTs in the Toujeo treatment arm within each SOC.

Analysis of adverse events of special interest

- A listing of patients with symptomatic overdose with IMP/NIMP. Pregnancy will be listed separately.

Analysis of pre-treatment and post-treatment adverse events

An overview of pre-treatment and/or post-treatment adverse events will be presented as a listing if only a few, otherwise as a summary (presenting number of patients with any pre-/post-treatment adverse events, serious pre-/post-treatment adverse events, pre-/post-treatment adverse events leading to death).

- All pre-treatment adverse events by primary SOC and PT, showing the number (%) of patients with at least 1 pre-treatment adverse event, sorted by the internationally agreed SOC order and PT will be presented by alphabetical order.
- All post-treatment adverse events by primary SOC and PT, showing the number (%) of patients with at least 1 post-treatment adverse event, sorted by the internationally agreed SOC order and PT will be presented by alphabetical order.

In case of a low number of events (pre-treatment and post-treatment), a only listing will be provided.

2.4.5.3 Deaths

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

- Number (%) of patients who died by study period (on-study, on-treatment, poststudy).
- TEAEs leading to death (death as an outcome on the adverse event case report form page as reported by the Investigator) by primary SOC and PT showing number (%) of patients sorted by internationally agreed SOC order and by decreasing incidence of PTs in the Toujeo treatment arm within each SOC. In case of a low number of events, only a listing will be provided.

Deaths in non-randomized patients or randomized but not treated patients will be also presented.

2.4.5.4 Analyses of laboratory variables

As most of safety laboratory parameters are only collected at screening and not during the study (see Section 2.1.4.3), no analyses will be performed.

2.4.5.5 Analyses of vital sign variables

The summary statistics (including number, mean, median, Q1, Q3, standard deviation, minimum and maximum) of all vital signs variables as well as changes from baseline will be calculated for each study assessment (baseline, Week 8 and Week 12, see Section 2.5.4) by treatment group.

2.4.5.6 Analyses of electrocardiogram variables

Not applicable to the study.

2.4.6 Analyses of pharmacokinetic and pharmacodynamic variables

Not applicable to the study

2.4.7 Analyses of quality of life/health economics variables

The analyses will be done on ITT population.

For DTSQs, MOS-Sleep, and WPAI, a descriptive summary of each score (see Section 2.1.7) will be done at baseline and at Week 12.

Change in DTSQs, MOS-Sleep, and WPAI scores from baseline to Week 12 will be analyzed using an ANCOVA model including the fixed categorical effects of treatment group (Toujeo, Tresiba), and the randomization stratum of HbA1c at screening (<8.0%, ≥8.0%), as well as, the continuous fixed covariate of baseline score.

LS means and the corresponding SEs and 95% confidence interval by treatment for the change from baseline will be presented for descriptive purposes.

The DTSQc score will be analyzed using descriptive statistics at Week 12.

The DTSQc scores at Week 12 will be analyzed using an ANCOVA model including the fixed categorical effects of treatment group (Toujeo, Tresiba), DTSQs score at baseline, and the randomization stratum of HbA1c at screening (<8.0%, ≥8.0%).

LS means differences in scores between treatment groups and its SEs and their corresponding 95% confidence interval of will also be evaluated.

2.5 DATA HANDLING CONVENTIONS

The following formulas will be used for computation of parameters:

2.5.1 General conventions

CGM days will be defined as starting/ending at 00:00/23:59 on the 24-hour clock. Only useable CGM days within the time windows (see Section 2.5.4) are used for derivation of CGM endpoints (primary and secondary).

Percentage time in range

The proportion of CGM readings falling into the range of ≥70 to ≤180 mg/dL [≥3.9 to≤10 mmol/L] from useable CGM (at least 10 useable days among the 20-day CGM within the time-window, see Section 2.5.4).

If there is missing data between two available measurements, linear interpolation will be performed to impute the time points with missing data. For example, if a glucose value is less than 70 mg/dL at 5 min, missing at 10 and 15 min then higher than 180 mg/dL at 20 min, the data at 10 and 15 minutes will be imputed using linear interpolation. Missing data at the start of the day and end of day will not be imputed.

Demographic formulas

- If the date of T1DM diagnosis is not missing, age at diagnosis will be calculated extracting the year from the date of the T1DM diagnosis and the year of birth:
  a) Age at diagnosis (years) = YEAR (date of diagnosis) – year of birth.
- Time since diabetes diagnosis will be derived:
  a) Time since diabetes diagnosis (years) = Integer [(date of Informed Consent – Date of Diagnosis + 1)/365.25].

Disease characteristics at baseline formulas

- Total daily insulin dose adjusted for body weight: Total daily dose (IU) / body weight in kg.

Physical examination formulas

- BMI (Quetelet index formula):
  a) Weight (kg) / (Height (m) × Height (m))

  where using height assessed at Screening Visit and body weight collected over the study.

2.5.2 Data handling conventions for secondary efficacy variables

2.5.2.1 CGM secondary and tertiary endpoints

- mean hours per day = percent time × 24 hours

  24-hour Area under the Curve (AUC24), derived using the trapezoidal rule based on AUCt (AUC over the CGM assessment period) and then divided by the number of days with usable CGM data. Trapezoidal rule:

  a) 0.5 × (glucose reading at time i + glucose reading at time i-1) × (time i - time i-1).

Note: Missing values for calculation of AUC will be treated as not collected.

The following main secondary endpoints will be calculated as follows:

- Glucose total CV = 100% times square root of $s^2{}_{total}$ divided by the mean glucose $\bar{y}$. That is:

$$\mathrm{CV}_{total} = 100 \times \frac{s^2_{total}}{\bar{y}}$$

with

$$s^2{}_{total} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(y_{ij} - \bar{y}\right)^2}{N-1}$$

and

$$\bar{y} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij}}{N},$$

where N is the number of total useable CGM data for each patient; k is the number of CGM assessment days; y_{ij} is the j^{th} CGM data from the Day i and n_i is the number of total useable CGM data within the Day i.

- Glucose between days CV = 100% times square root of $s^2{}_{between}$ divided by the mean glucose $\bar{y}$. That is:

$$\mathrm{CV}_{between} = 100 \times \frac{s^2_{between}}{\bar{y}}$$

with

$$s^2{}_{between} = \frac{\sum_{i=1}^{k}(\bar{y}_i - \bar{y})^2}{K-1}$$

where $\bar{y}_i$ is the mean glucose for Day i:

$$\bar{y}_i = \frac{\sum_{j=1}^{n_i} y_{ij}}{n_i}.$$

- Glucose within days CV = 100% times square root of $s^2{}_{within}$ divided by the mean glucose $\bar{y}$. That is:

$$\mathrm{CV}_{within} = 100 \times \frac{s^2_{within}}{\bar{y}}$$

with

$$s^2{}_{within} = \frac{1}{K}\sum_{i=1}^{k} \frac{\sum_{j=1}^{n_i}\left(y_{ij} - \bar{y}_i\right)^2}{n_i - 1}.$$

A Proc Varcomp procedure will be used to compute Glucose within days CV and Glucose between days CV for each patient where time point variability within each day is also accounted for

```
proc varcomp data =dsn;
 by USUBJID avisit;
 class ADY TIMEBIN;
 model AVAL = ADY TIMEBIN;
 ods output estimates=outstat;
run;
```

where:

AVAL: glucose value
AVISIT: Analysis visit including Baseline and Week 12
ADY: Analysis day
TIMEBIN: Timepoints

The CGM secondary/exploratory endpoints related to percent in time in range whatever the threshold(s) will be derived on a similar way as the primary endpoint.

- GMI (%) = 3.31 + 0.02392 × (mean glucose in mg/dL)

2.5.2.2 Absolute change

Absolute change from baseline for all secondary endpoints will be calculated as follows:

- Absolute change from baseline on X at Visit i = (X at Visit i) – (X at baseline)

where X is a given secondary endpoint value, i is the post-baseline visit i and baseline is the last assessment before first administration of IMP.

2.5.2.3 HbA1c and FPG

HbA1c (mmol/mol) = (HgbA1C% − 2.15) × 10.929

FPG (mmol/mol) = FPG mg/dL × 0.0555

2.5.2.4 WPAI

Employment status will be derived based on the coding rules for self-administration. That is, if Q1 = YES or Q1 = NO or missing and hours missed or worked > 0, then employed. If Q1 is missing and hours missed and worked = 0, then not employed.

The following scoring instruction is used to score the four subscales from the 6-item validated questionnaire WPAI-diabetes.

Four subscales:

- Absenteeism (Percent work time missed due to diabetes): Q2 / (Q2 + Q4)
- Presenteeism (Percent impairment while working due to diabetes): Q5 / 10
- Work productivity loss (Percent overall work impairment due to diabetes):

  $$Q2 / (Q2 + Q4) + [(1 - (Q2 / (Q2 + Q4))) \times (Q5 / 10)]$$
- Activity Impairment (Percent activity impairment due to diabetes): Q6 / 10

2.5.2.5 MOS Sleep

Values for Sleep Quantity (ie, the average number of hours sleep each night during the past 4 weeks) will be considered aberrant if they are >24. For these cases, the variable will be considered as missing.

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.

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 Completion of End of Study/Follow-up case report form page. If this date is missing, the exposure duration should be left as missing.

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 with missing or partial date/time of onset

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

Handling of adverse events when date and time of first investigational medicinal product administration is missing

When the date and time of the first IMP administration is missing, all adverse events that occurred on or after the day of randomization should be considered as treatment-emergent adverse events. The exposure duration should be kept as missing.

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 adverse event considered as such in the frequency tables of possibly related adverse events, but no imputation should be done at the data level.

Handling of missing responses of DTSQc and DTSQs

The total treatment satisfaction score is the sum of items 1, 4, 5, 6, 7 and 8.

If no more than 2 out of the 6 questions comprising the total treatment satisfaction score are missing, the total 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.

If items 2 or 3 are not answered, the frequency of hypoglycemia or hyperglycemia scores is considered as missing.

Handling of missing responses of WPAI

If the patient is employed, and Q2, Q4 Q5, and Q6 is missing then the scores will be imputed using average of all scores from other patients who answer the questions. Otherwise, only Q6 will be imputed using average of all scores from other patients.

Multiple imputation for primary analysis

Multiple imputation under MAR assumptions

Once the missing data pattern will be assessed, the programming process for this primary analysis will follow three steps (if monotone missing pattern is observed for all patients, then proceed directly with Step 1b):

- Step 1a: Use PROC MI to produce a monotone missing data pattern. A monotone Monte Carlo Markov Chain MCMC imputation model will be used to impute interim missing data. A sample SAS code for this imputation is shown below:

```
PROC MI NIMPUTE=1000 DATA=orig OUT=impute1 SEED=14947 MIN=0 MAX=100;
 BY ARM;
 MCMC IMPUTE=MONOTONE;
 VAR BASE PTIR;
RUN;
```

- Step 1b: Use PROC MI to execute the second imputation using the data set produced in the first imputation. A sample SAS code for this imputation is shown below (if Step 1a was not done due to monotone missing, then NIMPUTE=1000 and DATA=ORIG, BY statement not needed):

```
PROC MI NIMPUTE=1 DATA=impute1 OUT=impute2 SEED=14947 MIN=0 MAX=100;
 CLASS ARM HBACRIT1;
  BY _IMPUTATION_;
 MONOTONE METHOD=REG
 VAR ARM HBACRIT1 BASE PTIR;
RUN;
```

- Step 2: Each of the 1000 imputed datasets will undergo an ANCOVA using fixed categorical effects of treatment group (Toujeo, Tresiba), randomization stratum of screening HbA1c (<8.0% versus ≥8.0%), and the baseline percent time in range value 1 using the following ESTIMATE statement in PROC GLM

```
PROC GLM DATA=impute2;
 BY _IMPUTATION_;
 CLASS ARM HBACRIT1;
 MODEL PTIR = ARM HBACRIT1 BASE;
 ESTIMATE "TOUJEO-0.9 TRESIBA" INTERCEPT 0.1 TREAT 1 -0.9;
RUN;
```

- Step 3: The 1000 analysis results will be combined using PROC MIANALYZE (Rubin's rule). A sample SAS code for this combined analysis is shown below:

```
PROC MIANALYZE PARMS=est;
 MODELEFFECTS estimate;
RUN;
```

In this SAS code above:

- orig: input dataset including one observation per patient in the ITT population
- BASE: Percentage of time in the glucose range of ≥70 to ≤180 mg/dL all time at baseline
- PTIR: Percentage of time in the glucose range of ≥70 to ≤180 mg/dL all time at Week 12

- ARM: treatment group as per randomized (TOUJEO/ TRESIBA)
- seed: the seed (14947) has been chosen arbitrarily and is based on the study code (LPS14947).
- HBACRIT1: randomization stratum of screening HbA1c (<8.0% versus ≥8.0%)

2.5.4 Windows for time points

The windows below will be employed for the main primary and secondary efficacy endpoints as well as vital sign which will also be collected during the study for both scheduled and unscheduled visits.

CGM

CGM values measured outside the visit window will be flagged for exclusion from analyses. No derivations of CGM endpoints will be performed for CGM data collected outside of the analysis CGM window. Study days for CGM evaluations are calculated from the randomization date, and Day 1 indicates the day of randomization.

	Target CGM visit in Study day	Analysis CGM window in study day
CGM study visit		
Week- 3 to Week -1	Day -20 to Day -1	-28 to -1
Week 10 to Week 12	Day 64 to 84	57 to 91

For 7-point SMPG, HbA1c, FPG and PROs: Day indicates days relative to randomization date, and Day 1 indicates the day of randomization.

7-point SMPG

	Target Study day	Analysis window in study day
Time point		
Week 12	84	77 to 91

HbA1c, FPG

	Targeted Study Day	Analysis window in study days
Time Point		
Week 12	84	77 to 91

PROs

	Targeted Study Day	Analysis window in study days
Time Point		
Week 12	84	77 to 91

Vital signs - Safety

Study days for vital signs evaluations are calculated from the day of first IMP injection, the day of first IMP injection being Day 1. The last assessment before the first study drug intake will be considered as Baseline.

	Target visit in Study day	Analysis window in study day
Time Point		
Week 8	56	49 to 63
Week 12	84	77 to 91

One or more results for a required parameter may be recorded within a given visit window. In such an event, the assessment with the closest date to the scheduled visit date will be used. In the event of 2 observations within a given visit window being equidistant to the scheduled visit date, the first observation will be used for descriptive tabulations or figures.

2.5.5 Unscheduled visits

Unscheduled visit measurements of vital signs will not be included in the by-visit summaries but will be used for computation of baseline and analysis visit definition.

2.5.6 Pooling of centers for statistical analyses

No pooling of centers is planned for statistical analyses: center and country will not be included in statistical analysis.

2.5.7 Statistical technical issues

Not applicable for the study.

3 INTERIM ANALYSIS

No interim analysis is planned.

4 DATABASE LOCK

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

5 SOFTWARE DOCUMENTATION

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

6 REFERENCES

1. Battelino T, Danne T, Bergenstal RM, Amiel SA, Beck R, Biester T, et al. Clinical targets for continuous glucose monitoring data interpretation: recommendations from the international consensus on time in range. Diabetes Care. 2019 Aug;42(8):1593-603.

2. Bergenstal RM, Beck RW, Close KL, Grunberger G, Sacks DB, Kowalski A, et al. Glucose management indicator (GMI): a new term for estimating A1C from continuous glucose monitoring. Diabetes Care. 2018 Nov;41(11):2275-80.

3. Danne T, Nimri R, Battelino T, Bergenstal RM, Close KL, DeVries JH, et al. International consensus on use of continuous glucose monitoring. Diabetes Care. 2017 Dec;40(12):1631-40.

4. Hochber Y, Tamhane AC. Multiple comparison procedures, John Wiley & Sons (print 1987). Published online: 27 May 2008.

7 LIST OF APPENDICES

Appendix A: Summary of statistical analyses

Appendix B: Questionnaires (PROs)

APPENDIX A: SUMMARY OF STATISTICAL ANALYSES

EFFICACY ANALYSIS

Endpoint	Analysis population	Primary analysis	Supportive analysis	Subgroup analysis	Other analyses
Primary endpoint					
Percent of time in glucose range of ≥70 to ≤180 mg/dL (≥3.9 to ≤10 mmol/L) using continuous glucose monitoring (CGM) at Week 12	ITT, PP, ITT COVID 19 free	For ITT population: ANCOVA with fixed categorical effects Trt and randomization stratum HbA1c and continuous fixed covariate of baseline percent in range value with multiple imputation of missing data Non inferiority with a relative non-inferiority margin of 10 % is the primary analysis	1st Sensitivity analyses: same modeling (ANCOVA) with MI on PP population 2nd Sensitivity analyses same modeling (ANCOVA) with multiple imputation of missing data being imputed in conjunction with MNAR assumption from ITT population. 3rd sensitivity analyses: same modeling (ANCOVA) with MI on Population without trial impact (disruption) due to COVID-19	Yes, subgroup by screening and baseline factors such as duration of diabetes, Gender, Randomization strata, BMI.	In addition to the non-inferiority testing that is the primary analysis, and only if non-inferiority of Toujeo in comparison to Tresiba on the percent time in the range of 70-180 mg/dL is demonstrated and if non-inferiority of Toujeo versus Tresiba on the glucose total CV is demonstrated, superiority of Toujeo over Tresiba on the percent time in the range of 70-180 mg/dL will be tested

Endpoint	Analysis population	Primary analysis	Supportive analysis	Subgroup analysis	Other analyses
Secondary endpoints					
Glucose total CV using CGM at Week 12 (Main secondary endpoint)	ITT	ANCOVA model to perform the analysis includes the fixed categorical effects of treatment group (Toujeo, Tresiba),and randomization stratum of screening HbA1c (<8.0%versus ≥8.0%), as well as, the continuous fixed covariate of baseline value for the corresponding endpoint As this is the main secondary endpoint: Multiple imputation (as for primary endpoint) and non-inferiority with a relative non-inferiority margin of 10 % tested in case of non-inferiority demonstrated for primary analysis of primary endpoint	No	No	No
Glucose within-day CV using CGM at Week 12	ITT	The ANCOVA model to perform the analysis includes the fixed categorical effects of treatment group (Toujeo, Tresiba), and randomization stratum of screening HbA1c (<8.0%versus ≥8.0%), as well as, the continuous fixed covariate of baseline value for the corresponding endpoint.	No	Yes, subgroup defined by glucose within day CV ≤36% vs. >36% at baseline.	No
Glucose between-day CV using CGM at Week 12	ITT	The ANCOVA model to perform the analysis includes the fixed categorical effects of treatment group (Toujeo, Tresiba), randomization stratum of screening HbA1c (<8.0%versus ≥8.0%), as well as, the continuous	No	No	No

Endpoint	Analysis population	Primary analysis	Supportive analysis	Subgroup analysis	Other analyses
		fixed covariate of baseline value for the corresponding endpoint.			
HbA1c Change from baseline at Week 12	ITT	Change in HbA1c will be analyzed using an ANCOVA model including the fixed categorical effects of treatment group (Toujeo, Tresiba), as well as, the continuous fixed covariate of baseline HbA1c value. 95% CI and P-values presented for these endpoints will be done for descriptive purposes only.	No	No	No
FPG Change from baseline at Week 12	ITT	Change in FPG will be analyzed using ANCOVA model including the fixed categorical effects of treatment group (Toujeo, Tresiba), and the randomization stratum of HbA1c at screening (<8.0%, ≥8.0%), as well as, the continuous fixed covariate of baseline FPG value. 95% CI and P-values presented for these endpoints will be done for descriptive purposes only.	No	No	No
Tertiary/exploratory endpoints except PROs					
Change from baseline in 7-point SMPG profiles (pre-prandial and 2-hour postprandial plasma glucose at breakfast, lunch and dinner, and at bedtime) to Week 12	ITT	Descriptive summary statistics by treatment group	No	No	No

Endpoint	Analysis population	Primary analysis	Supportive analysis	Subgroup analysis	Other analyses
Mean hours per day in the CGM glucose range of ≥70 to ≤180 mg/dL (≥3.9 to ≤10 mmol/L)	ITT	Descriptive summary statistics by treatment group	No	No	No
Percent time and mean hours per day in the CGM glucose range of ≥70 to ≤140 mg/dL (≥3.9 to ≤7.8 mmol/L) (for all-time and during the night [00:00 to 05:59])	ITT	Descriptive summary statistics by treatment group	No	No	No
Nocturnal percent time in CGM glucose range of ≥70 to ≤180 mg/dL (≥3.9 to ≤10 mmol/L) at Week 12, obtained using CGM	ITT	Descriptive summary statistics by treatment group	No	No	No
Percent time and mean hours per day with CGM glucose: <54 mg/dL, (for all time and during the night [00:00 to 05:59])	ITT	Descriptive summary statistics by treatment group	No	No	No
Percent time with CGM glucose <70 mg/dL within the 12 hours after basal insulin injection	ITT	Descriptive summary statistics by treatment group	No	No	No
Percent time and mean hours per day with CGM glucose >250 mg/dL	ITT	Descriptive summary statistics by treatment group	No	No	No
Mean CGM glucose profile over 24-hour period	ITT	Graphic presentation	No	No	No

Endpoint	Analysis population	Primary analysis	Supportive analysis	Subgroup analysis	Other analyses
24h-AUC of CGM glucose values	ITT	Descriptive summary statistics by treatment group	No	No	No
Mean glucose	ITT	Descriptive summary statistics by treatment group	No	No	No
GMI	ITT	Descriptive summary statistics by treatment group	No	No	No
Percent time and mean hours per day in the CGM glucose range of ≥54 to <70 mg/dL (for all-time and during the night [00:00 to 05:59]),	ITT	Descriptive summary statistics by treatment group	No	No	No
Percent time and mean hours per day in the glucose range of ≥180 to <250 mg/dL (on all-time and during the night [00:00 to 05:59]),	ITT	Descriptive summary statistics by treatment group	No	No	No
Daily insulin doses (total daily insulin, total basal insulin, total mealtime insulin, ratio of basal to mealtime insulin dose) in units and U/kg body weight	Safety	Descriptive summary statistics (including change from baseline to Week 12) by treatment group	No	No	No
PROs endpoints					
DTSQs	ITT	Change in DTSQs scores from baseline to Week 12 will be analyzed using an analysis of covariance (ANCOVA) model including the fixed categorical effects of treatment group (Toujeo, Tresiba), and the	No	No	No

Endpoint	Analysis population	Primary analysis	Supportive analysis	Subgroup analysis	Other analyses
		randomization stratum of HbA1c at screening (<8.0%, ≥8.0%), as well as, the continuous fixed covariate of baseline score. LS means and its corresponding 95% confidence interval by treatment for the change from baseline will be presented for descriptive purposes			
DTSQc	ITT	The DTSQc scores at Week 12 will be analyzed using an analysis of covariance (ANCOVA) model including the fixed categorical effects of treatment group (Toujeo, Tresiba),and the randomization stratum of HbA1c at screening (<8.0%, ≥8.0%). LS means and its corresponding 95% confidence interval by treatment for the change from baseline will be presented for descriptive purposes	No	No	No
MOS Sleep	ITT	Change in MOS-Sleep from baseline to Week 12 will be analyzed using an analysis of covariance (ANCOVA) model including the fixed categorical effects of treatment group (Toujeo, Tresiba), and the randomization stratum of HbA1c at screening (<8.0%, ≥8.0%), as well as, the continuous fixed covariate of baseline score. LS means and its corresponding 95% confidence interval by treatment for the change from baseline will be presented for descriptive purposes	No	No	No

Endpoint	Analysis population	Primary analysis	Supportive analysis	Subgroup analysis	Other analyses
WPAI	ITT	Change in WPAI scores from baseline to Week 12 will be analyzed using an analysis of covariance (ANCOVA) model including the fixed categorical effects of treatment group (Toujeo, Tresiba), and the randomization stratum of HbA1c at screening (<8.0%, ≥8.0%), as well as, the continuous fixed covariate of baseline score LS means and its corresponding 95% confidence interval by treatment for the change from baseline will be presented for descriptive purposes	No	No	No

ANCOVA = analysis of covariance; ITT = intent-to-treat; PP = per-protocol

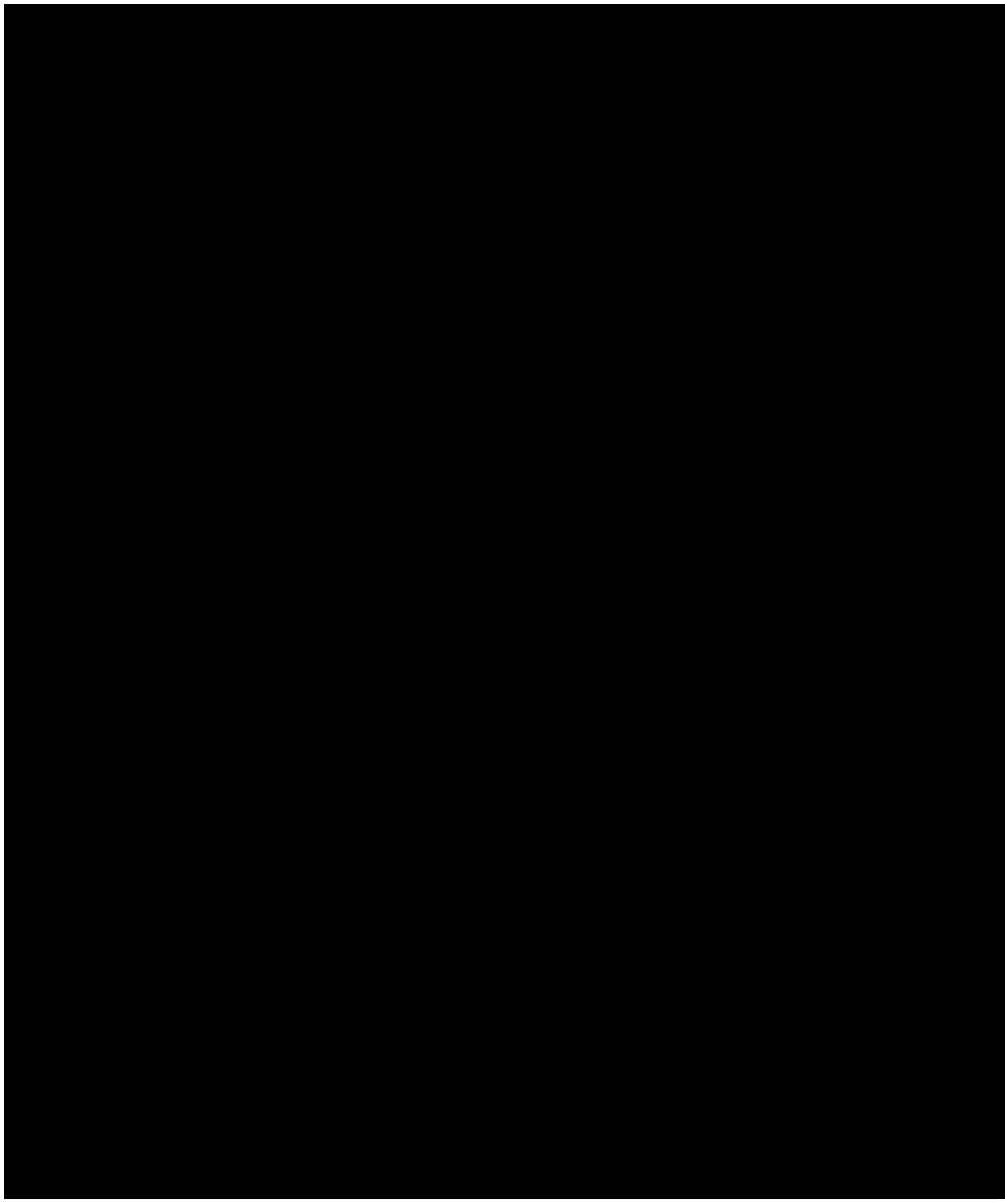
SAFETY ANALYSES

Endpoint	Analysis population	Primary analysis	Supportive analysis	Subgroup analysis	Other analyses
Number of patients with adverse events	Safety	Summary by treatment group using number and % of patients experiencing at least one adverse event, focusing on TEAE	No	No	No
Number of patients with at least one hypoglycemic event from baseline to Week 12	Safety	Summary by treatment group using number and % of patients experiencing at least one hypoglycemic event will be presented for any hypoglycemia and for each type of hypoglycemia, overall and according to time of occurrence (nocturnal [ie, 00:00 to 05:59], any time of day and daytime [ie, 06:00 to	No	No	No

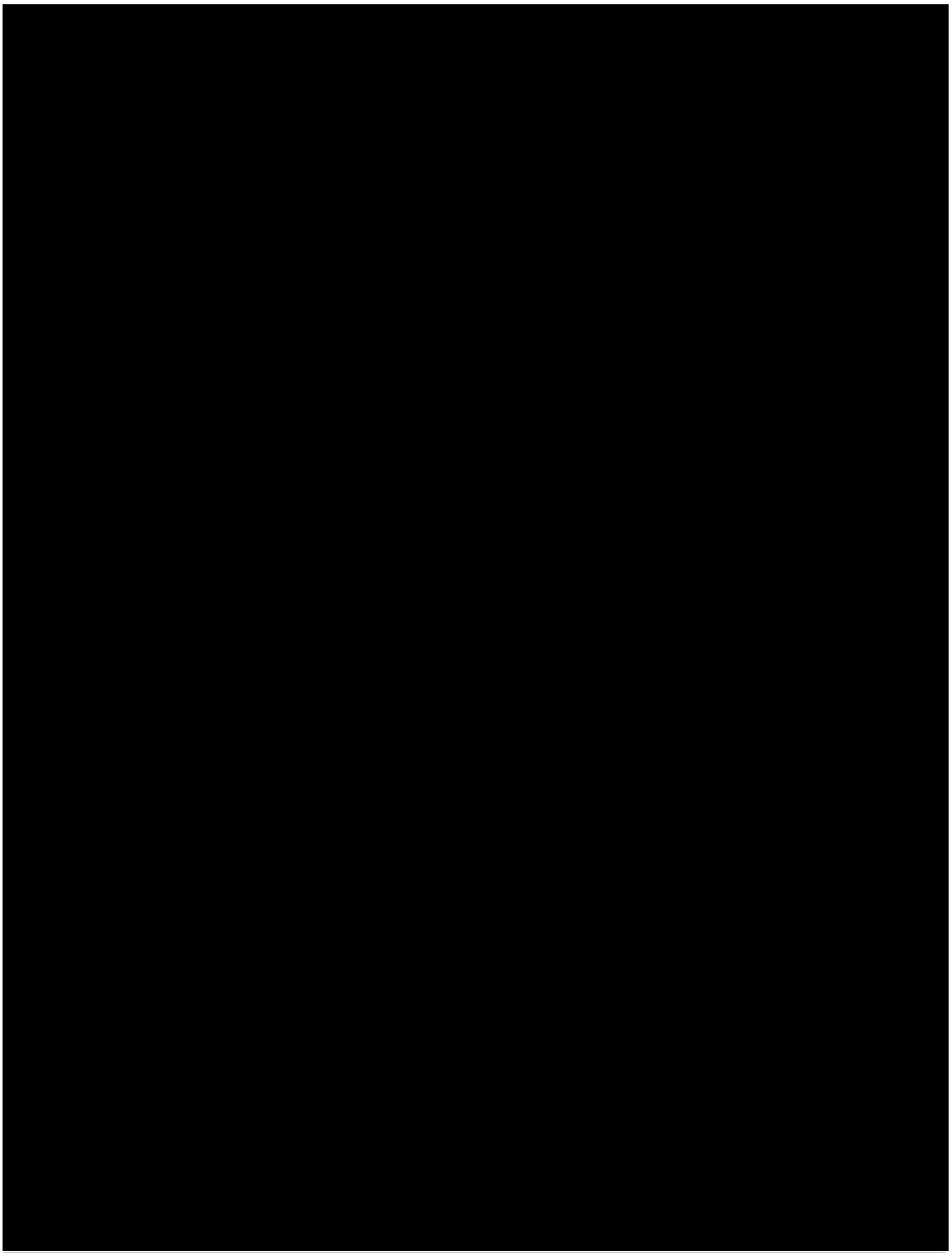
Endpoint	Analysis population	Primary analysis	Supportive analysis	Subgroup analysis	Other analyses
		23:59] during the 12-week on treatment period Logistic regression model adjusted on randomization strata of screening HbA1c to estimate Odds Ratio and its corresponding 95% CI for the Toujeo arm over the Tresiba arm for each hypoglycemic event			
Number of hypoglycemic events per participant-year from baseline to Week 12	Safety	Number and rate of hypoglycemic events (in patient-year of exposure) will be determined for any hypoglycemia and for each type of hypoglycemic event, overall and according to time of occurrence (nocturnal [ie, 00:00 to 05:59], any time of the day) during the 12-week on treatment period. Over-dispersed Poisson regression model adjusted on randomization strata of screening HbA1c to estimate for each hypoglycemic event, the rate ratio, and its corresponding 95% CI, of Toujeo arm over Tresiba arm	No	No	No

APPENDIX B: QUESTIONNAIRES (PROS)

Diabetes Treatment Satisfaction Questionnaire: DTSQs

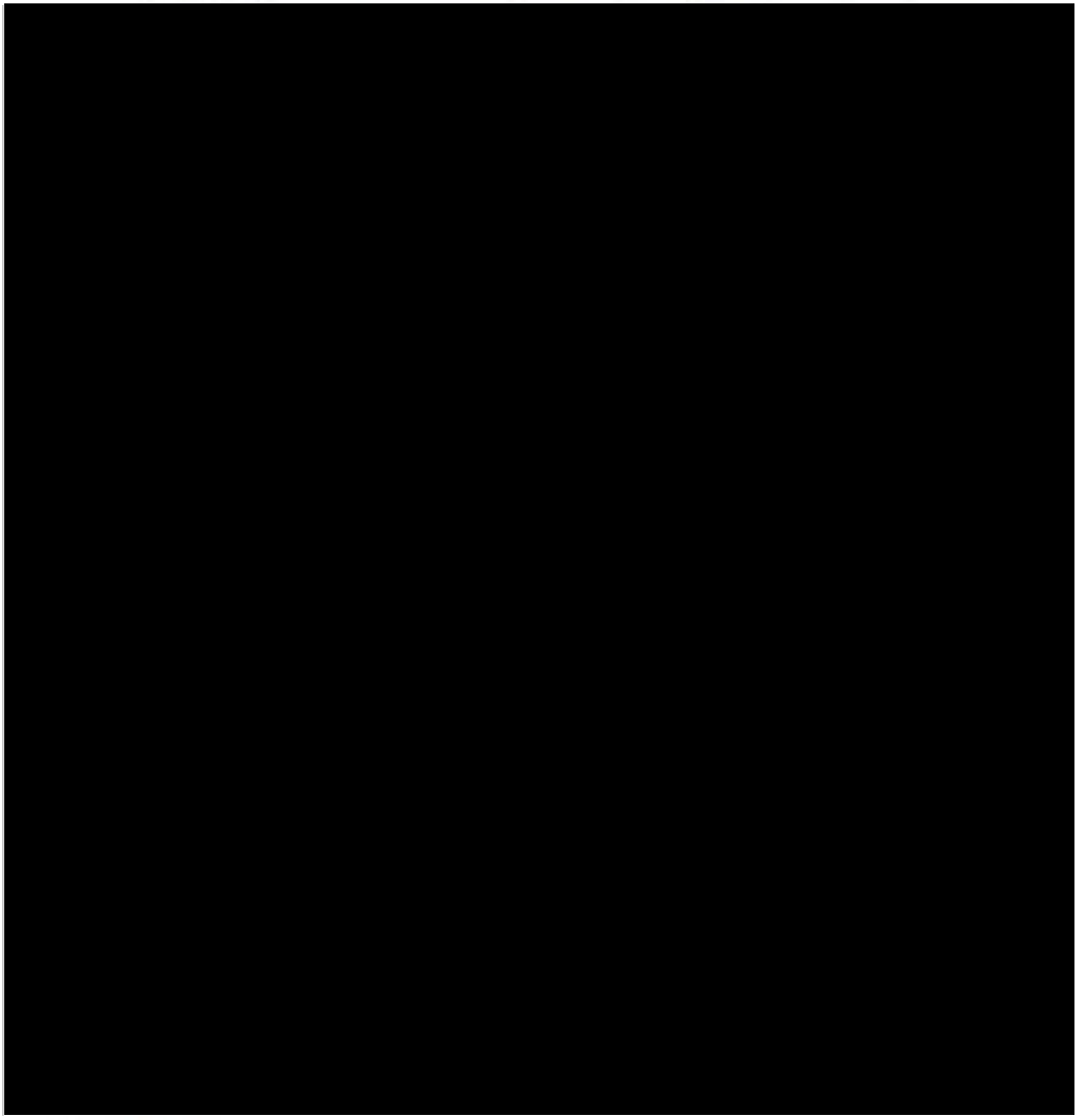


Diabetes Treatment Satisfaction Questionnaire (change): DTSQc

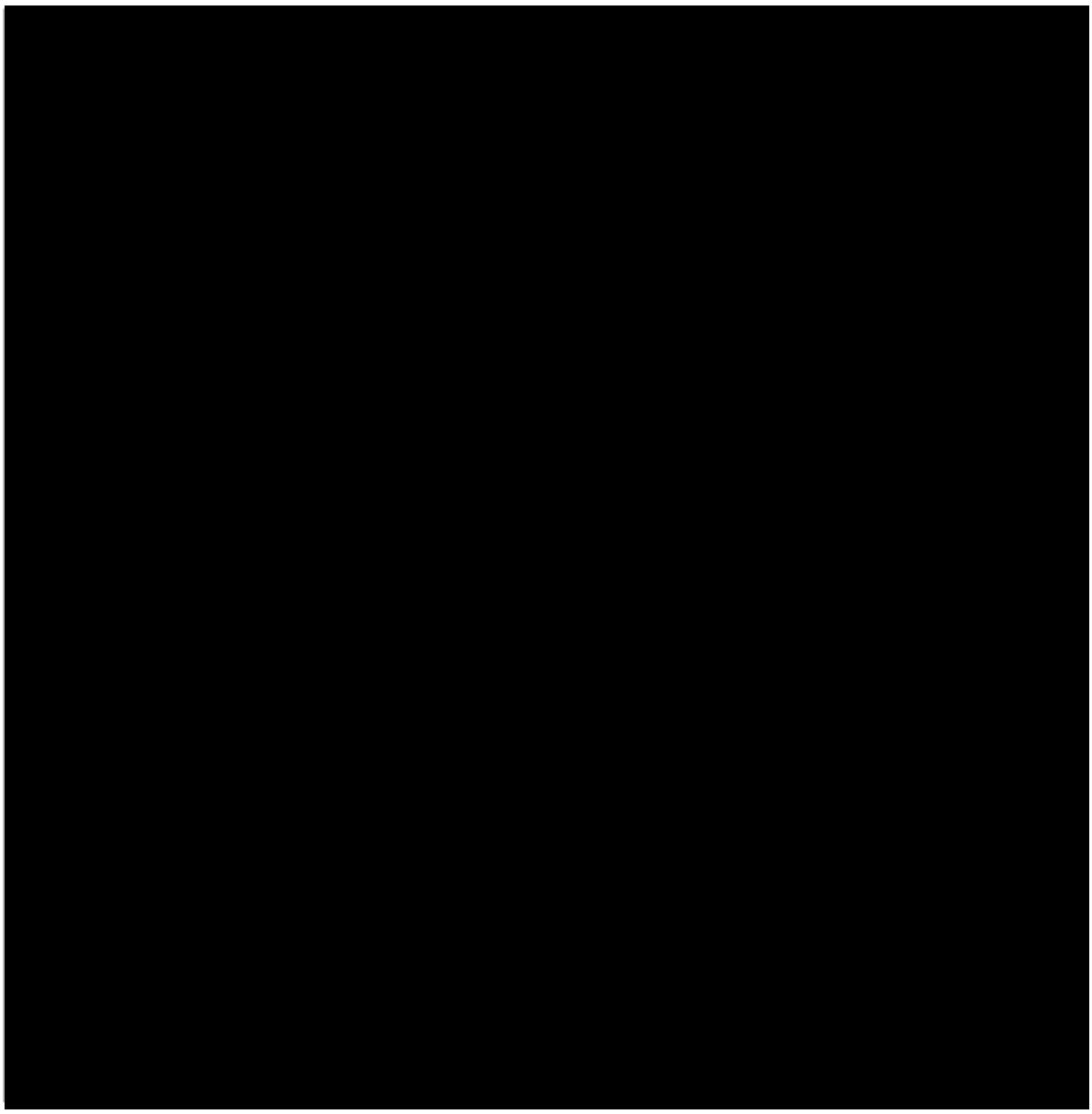


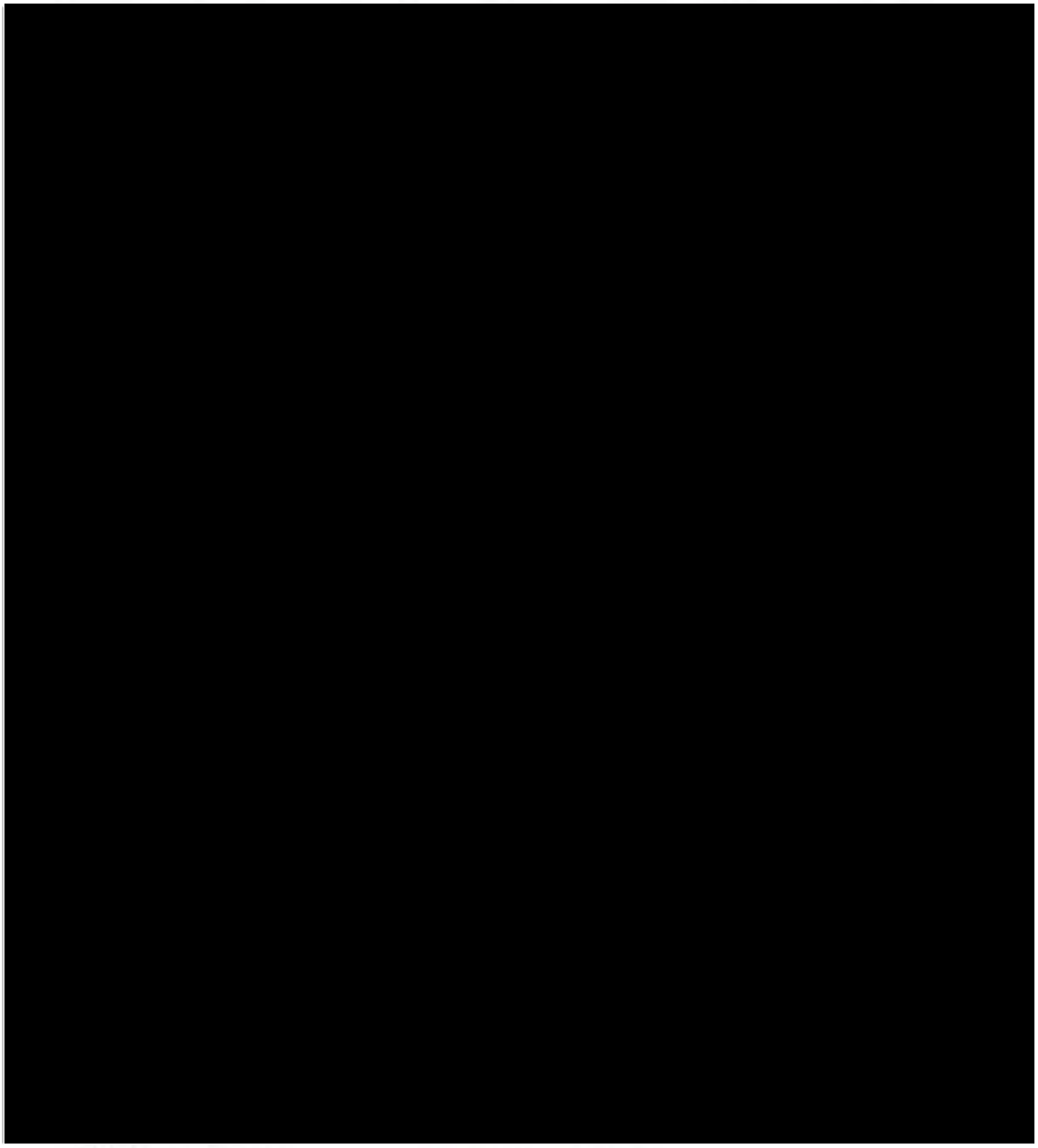
NOT FOR USE: for review by sanofi-aventis, ref HPR1649
DTSQc © Prof Clare Bradley 11.9.96 Standard UK English (rev. 4.3.98; generic intro. rev. 28.2.02)
Health Psychology Research, Dept of Psychology, Royal Holloway, University of London, Egham, Surrey, TW20 0EX, UK

Your Sleep



Work Productivity and Activity Impairment Questionnaire: Diabetes (WPAI:Diabetes)





WPAI:SHP V2.0 (US English)

Reilly MC, Zbrozek AS, Dukes E: The validity and reproducibility of a work productivity and activity impairment measure. Pharmacoeconomics 1993; 4(5):353-365. 2

Signature Page for VV-CLIN-0612739 v1.0
lps14947-16-1-9-sap

Approve & eSign	[redacted]
Approve & eSign	
Approve & eSign	