Statistical Analysis Plan Version 2 -I8F-MC-GPIL

An Open-Label, Single-Arm, Phase 4 Study to Assess Glycemic Control When Adults with Type 2 Diabetes Switch from a GLP-1 RA to Tirzepatide (SURPASS-SWITCH-2)

NCT05706506

Approval Date: 19-Sep-2023

Title Page

Protocol Title: An Open-Label, Single-Arm, Phase 4 Study to Assess Glycemic Control When Adults with Type 2 Diabetes Switch from a GLP-1 RA to Tirzepatide (SURPASS-SWITCH-2)

Protocol Number: I8F-MC-GPIL

Compound Number: LY3298176

Short Title: A study to investigate glycemic control in adults with type 2 diabetes switching to tirzepatide from a GLP-1 RA

Acronym: SURPASS-SWITCH-2

Sponsor Name: Eli Lilly and Company

Legal Registered Address: Indianapolis, Indiana, USA 46285

Regulatory Agency Identifier Number(s)

Registry	ID
IND	128801
EudraCT	2022-002708-18

Document ID:

Template Version: 1.0

Eli Lilly and Company

I8F-MC-GPIL

An Open-Label, Single-Arm, Phase 4 Study to Assess Glycemic Control When Adults with Type 2 Diabetes Switch from a GLP-1 RA to Tirzepatide (SURPASS-SWITCH-2)

September 19, 2023

Statistical Analysis Plan Version 2.0

Prepared by:

PPD, part of Thermo Fisher Scientific

7551 Metro Center Drive

Austin, TX 78744, USA

Issued by: PPD

Upon review of this document the undersigned approves the statistical analysis plan. The analysis methods and data presentation are acceptable, and the table, listing, and figure production can begin.

Approved by: PPD

Table of Contents

Version History

This Statistical Analysis Plan (SAP) for study I8F-MC-GPIL is based on the protocol dated 16AUG2022.

The second version of SAP (SAP V2.0) for study I8F-MC-GPIL incorporates the updates jointly decided by Eli Lilly & Co. and PPD. Please see Revision History for details.

SAP Version History Summary

SAP Version	Approval Date	Change	Rationale
1.0	13 Feb 2023	Not Applicable	Original version
2.0	19 Sep 2023	Post-Dry Run 1 update	

Revision History

The key changes of SAP V2.0 for study I8F-MC-GPIL are summarized below:

- Section 3 (p. 13). Add the definition of inadvertent enroller.
- Section 4.1 (p. 15). Clarify the CGM data from a non-compliant session will be excluded from CGM-related analysis.
- Section 4.5 (p. 22). Remove the formula of mean 24-hour daily blood glucose.
- Section 4.6.2 (p. 25). Explain how the AE of interest is defined.
- Section 6.1.1 (p. 37). Add additional criteria to distinguish prior and con med.
- Section 6.1.1 (p. 38). Defines OAM.
- Section 6.3 (p. 39). List all SMQs requested in the TLF outputs.

1. Introduction

This study aims to investigate the effects of switching from glucagon-like peptide-1 receptor agonist (GLP-1 RA) therapy to tirzepatide in people with type 2 diabetes (T2D). Tirzepatide is a recently approved single molecule that acts as glucose-dependent insulinotropic polypeptide (GIP) and GLP-1 RA, which has demonstrated an ability to reduce HbA1c and body weight in adults with T2D across all 3 therapeutic doses: 5, 10, and 15 mg (Del Prato et al. 2021; Frías et al. 2021; Ludvik et al. 2021; Rosenstock et al. 2021; Dahl et al. 2022).

The study will assess the efficacy and safety characteristics of the study intervention by using a starting dose of 5 mg tirzepatide and will evaluate changes in glycemic control, body weight, and gastrointestinal (GI) tolerability occurring in the first 12 weeks for the targeted study population.

There are no changes to the analyses described in the protocol. Table, figure, and listing specifications are contained in a separate document.

1.1. Objectives, Endpoints, and Estimands

Objectives	Endpoints
Primary	
• To demonstrate that participants with T2D switching from a stable dose of GLP-1 RA to tirzepatide 5 mg QW have an improvement in HbA1c at Week 12 compared to baseline	• Change from baseline in HbA1c
Secondary	
• To evaluate the changes in CGM-assessed TAR from baseline at Week 4 and at Week 12	• Change from baseline in ▪ Percentage of time per day that CGM-derived values are >180 mg/dl (10 mmol/L) ▪ Duration of time in minutes per day that CGM-derived values are >180 mg/dl (10 mmol/L)
• To evaluate the change in FSG from baseline at Week 12	• Change from baseline in FSG
• To evaluate the changes from baseline in weight at Week 12	• Change from baseline in weight
Exploratory	

• To evaluate changes in CGM-derived measures from baseline over the course of 12 weeks	• Change from baseline in ▪ standard CGM metrics (TIR, TBR, GV) ▪ time in TITR ▪ composite of all in-ranges ▪ mean 24-hr daily blood glucose ▪ estimated post-prandial glucose excursion ▪ estimated fasting glucose
Abbreviations: CGM = continuous glucose monitoring; FSG = fasting serum glucose; GLP-1 RA = glucagon-like peptide-1 receptor agonist; GV = glycemic variability; HbA1c = hemoglobin A1c; QW = once-weekly; T2D = type-2 diabetes; TAR = time above range (>180 mg/dl or >10 mmol/L); TBR = time below range (≤70 mg/dl or ≤3.9 mmol/L); TIR = time-in-range (71 to 180 mg/dL or 3.9 to 10 mmol/L); TITR = time in tight target range (71 to 140 mg/dL or 3.9 to 7.8 mmol/L);	

Primary estimand

The primary estimand in this study is the efficacy estimand.

The efficacy estimand focuses on the treatment effect if participants continued to receive the study treatment without prohibited medication. This estimand will be used in publications to inform prescribers or physicians.

The primary clinical question of interest is

> What is the change in HbA1c from baseline to Week 12 in participants with T2D, regardless of changes in background oral antihyperglycemic medication (OAM), and assuming that participants continue to take treatment without use of prohibited medication?

This estimand is described by the following attributes:

- Population: participants with T2D on a GLP-1 RA who may benefit from additional therapy. Further details can be found in Protocol Section 5.
- Endpoint: changes from baseline to Week 12 in HbA1c.
- Treatment condition: treatment with tirzepatide regardless of changes to the background OAM. Further details on study treatment can be found in Protocol Section 6.
- Intercurrent events of interest: "Treatment discontinuation for any reason" and "Initiation of prohibited medication" will be addressed using the following strategies:
    - As if participants stay on the treatment (hypothetical strategy)
    - As if participants did not receive prohibited medication (hypothetical strategy).
- Population-level summary: Mean HbA1c change from baseline
- Rationale for estimand: Estimating the effect of study intervention without confounding from prohibited mediations or treatment discontinuation

Secondary Endpoints based on Primary Estimand

1) The first clinical question of interest is for the secondary objective regarding continuous glucose monitoring levels:

   What is the change in CGM-assessed time above range (TAR) from baseline in participants with T2D, regardless of changes in background OAM, and assuming that participants continue to take treatment without use of prohibited medication?

   There are two endpoints defined for the above question based on the primary estimand:

   a) The first endpoint is

      - the change from baseline to Week 12 in percentage of time per day that CGM-derived values are >180 mg/dl (10 mmol/L)

   b) The second endpoint is

      - the change from baseline to Week 12 in duration of time in minutes per day that CGM-derived values are >180 mg/dl (10 mmol/L)

   Of note, the primary estimand is described by the following attributes:

- Population: participants with T2D on a GLP-1 RA who may benefit from additional therapy. Further details can be found in Protocol Section 5.
- Treatment condition: treatment with tirzepatide regardless of changes to the background OAM. Further details on study treatment can be found in Protocol Section 6.
- Intercurrent events of interest: "Treatment discontinuation for any reason" and "Initiation of prohibited medication" will be addressed using the following strategies:
  - As if participants stay on the treatment (hypothetical strategy)
  - As if participants did not receive prohibited medication (hypothetical strategy).
- Population-level summary: Mean HbA1c change from baseline

Rationale for estimand: Estimating the effect of study intervention without confounding from prohibited mediations or treatment discontinuation.

2) The second clinical question of interest is for the secondary objective regarding fasting serum glucose levels:

   What is the change in fasting serum glucose (FSG) from baseline to Week 12 in participants with T2D, regardless of changes in background OAM, and assuming that participants continue to take treatment without use of prohibited medication?

   The endpoint corresponding to this question is the change from baseline to Week 12 in FSG.

3) The third clinical question of interest is for the secondary objective regarding weight:

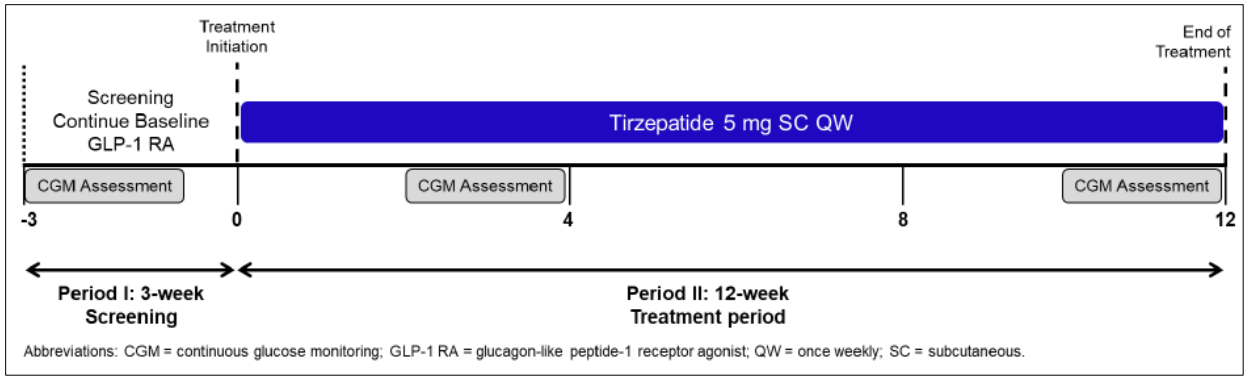
   What is the change in weight from baseline to Week 12 in participants with T2D, regardless of changes in background OAM, and assuming that participants continue to take treatment without use of prohibited medication?

   The endpoint corresponding to this question is the change from baseline to Week 12 in weight.

1.2. Study Design

SURPASS-SWITCH-2 is a study to assess the changes in glycemic control when switching from a stable dose of a GLP-1 RA directly to tirzepatide 5 mg in participants with T2D. It has the following main design attributes:

- Phase 4 study
- Multicenter, single arm
- Open label
- The study consists of 2 periods:
  - Screening period: 3 weeks, starting from Visit 1 to Visit 2.
  - Treatment period: 12 weeks, starting from Visit 2 to Visit 5, where Visit 2 (on Week 0) and Visit 5 (on Week 12) are fasting visits.
- No dose escalation will be employed during the trial. The goal is to initiate tirzepatide at 5 mg on Visit 2 (on Week 0) and maintain until Visit 5 (on Week 12).
- The study will enroll approximately 150 participants.



2. Statistical Hypotheses

For the primary estimand with primary endpoint, mean change in HbA1c percentage from baseline to Week 12, the following hypothesis is planned to be tested for tirzepatide 5 mg. Let the mean treatment difference be defined as μ = mean of ([HbA1c percentage at Week 12] – [baseline HbA1c percentage]), the statistical hypotheses are formulated as follows:

$$H_0: \mu \geq 0.0 \text{ vs. } H_a: \mu < 0.0$$

Operationally the hypothesis will be evaluated by 2-sided test, at a 2-sided alpha level of 0.05.

2.1. Multiplicity Adjustment

Multiplicity adjustment will not be implemented in this study.

3. Analysis Sets

For the purposes of analysis, the following analysis sets are defined:

Population/Analysis Set	Description
Screened population	• All participants who signed informed consent
Modified intent-to-treat population (mITT)	• All participants who are exposed to at least 1 dose of study intervention
Efficacy Analysis Set (EAS)	• This analysis set will be used to estimate the efficacy estimand for the primary objective, and for all other (secondary, exploratory) endpoints. • Data obtained during the treatment period from the mITT population, excluding (1) participants who were inadvertently enrolled* and (2) data after permanent discontinuation of treatment or initiation of prohibited medication.
Safety Analysis Set (SS)	• This analysis will be used to assess the safety of study treatment • Data obtained during the Treatment Period from the mITT population, regardless of adherence to treatment or initiation of prohibited medication.

* An inadvertent enroller refers to a subject who does not meet inclusion/exclusion criteria but receives study intervention.

4. Statistical Analyses

4.1. General Considerations

Statistical analysis of this study will be the responsibility of Lilly or its designee.

Any change to the data analysis methods described in the protocol will require an amendment *only if* it changes a principal feature of the protocol. Any other change to the data analysis methods described in the protocol, and the justification for making the change, will be described in the statistical analysis plan or clinical study report. Additional exploratory analyses of the data will be conducted as deemed appropriate.

Nominal significance levels:

All tests of treatment effects will be conducted at a 2-sided alpha level of 0.05, unless otherwise stated, and all confidence intervals will be given at a 2-sided 95% level.

Timing of Primary Analysis:

Primary analysis will be performed at database release after Last Participant Last Visit (LPLV).

Definition of baseline:

Baseline is defined as the last non-missing measurement recorded on or before the date/time of first dose. If the measurement time is not specified, it is assumed that any measurement taken on the date of first dose occurred prior to dosing.

Populations for analysis:

- Analyses will use the EAS to evaluate the primary efficacy estimand for all (secondary and exploratory) endpoints.
- Safety analysis will be performed on the SS.

Pooling of data:

Study data will be pooled across centers.

Analysis Windows:

Visit windows will be defined for by-visit summary and analysis purposes. Summary data such as adverse events and concomitant medications which are not reported by visit will not use visit windows. Visit labels will be assigned to each post-baseline record based on the windows or intervals for study day relative to the date of first dose. A record with the study day closest to the target visit day (week × 7 days + 1) will be chosen when multiple records fall in the same visit window. In case of a tie, the record with the earlier date will be chosen. If Day 1 is missing, the closest visit on or before the date of first dose will be used as baseline. All visits will be included in the summary tables for worst post-baseline values.

All efficacy and safety analyses will be based on the visit windows in Table 1.

Table 1. Visit Windows

Study Day Relative to First Dose	Visit	Target Day
Study day ≤ 1	Baseline	1
2 ≤ study day ≤ 43	Week 4 (Visit 3)	29
44 ≤ study day ≤ 81	Week 8 (Visit 4)	57
82 ≤ study day ≤ 88	Week 12 (Visit 5)	85

General details of obtaining CGM-derived endpoints:

Participants will wear the study-provided CGM sensor to complete three 14-day sessions. The 14-day session should be conducted over the period indicated below:

1) Right after Visit 1, ideally Weeks -3 and -2.
2) Before Visit 3, ideally Weeks 3 and 4.
3) Before Visit 5, ideally Weeks 11 and 12.

Compliance is defined as the percentage of expected data that are actually collected (i.e., number of actually collected measurements / number of expected measurements × 100%). For each of the session, the compliance must reach at least 70%, which equals to at least 10 out of 14 days. If a participant fails to reach the 70% compliance threshold on the session before Visit 3, the participant can attempt another session right after Visit 3 (ideally Weeks 5 and 6). CGM data collected from a non-compliant session will not be included in efficacy analysis.

CGM measurements over the 14-day session will be downloaded from the sensor on Visits 2, 3 (or 4 if non-compliance occurs), and 5, and are used to compute CGM-derived endpoints.

General statistical methods:

Summary statistics for continuous measures will include n, mean, standard deviation, median, minimum, and maximum. Categorical data (including categorized continuous measures) will be described using the sample size, participant count and percentage in each category. For the summary statistics of all numerical variables unless otherwise specified, minimum and maximum will be displayed to the same level of precision as reported. Mean and median will be displayed to one level of precision greater than the data collected. Standard deviation / standard error will be displayed to two levels of precision greater than the data collected. P-values will be rounded to three decimal places. If a P-value is less than 0.001 it will be reported as '<0.001.' If a P-value is greater than 0.999 it will be reported as '>0.999.'

Data will be displayed in all listings sorted by participant ID. Participants will be identified in the listings by the participant identification number concatenated with the investigator number.

When count data are presented, the percentage will be suppressed when the count is zero in order to draw attention to the non-zero counts. A row denoted "Missing" will be included in count tabulations where specified on the shells to account for dropouts and missing

values. The denominator for all percentages will be the number of subjects in that treatment within the analysis set of interest, unless otherwise specified.

The analysis model to make comparisons within treatment relative to continuous measurements assessed over time will be an MMRM with terms for:

- Visits, as fixed effects
- Baseline HbA1c (≤8.0%, >8.0%), as a covariate
- Baseline measurement, as a covariate

For analyses of HbA1c, continuous baseline HbA1c will be included in the model instead of baseline HbA1c category. An unstructured covariance structure will model the relationship of within-participant errors.

Logistic regression may be used to examine the within-treatment effect from baseline at 12 weeks in binary efficacy outcomes. The negative binomial regression model will be used to measure the within-treatment effect of discrete count measures if deemed appropriate.

Summary of intercurrent events:

The number of participants who experience an intercurrent event will be summarized in a table by event type (treatment discontinuation for any reason, or initiation of prohibited medication) and visit. For each visit, the number of intercurrent events which happened before or during the measurements taken on that visit will be calculated, excluding the events which happened during or before the measurements taken at the previous visits.

Furthermore, the number of participants without any intercurrent events will be presented by visit too. At each visit the number and percentage of participants without intercurrent events will be shown.

All analyses will be conducted using SAS Version 9.4 or higher.

4.2. Participant Dispositions

Participant disposition will be summarized in a table for the Screened population. A disposition of participants includes the number and percentage of participants for the following categories: participants who were screened, participants who were screen failures, participants who completed treatment, participants who permanently discontinued treatment, participants who completed the study, participants who discontinued from the study. All percentages will be based on the mITT population.

The reasons for permanent discontinuation of treatment and study participation will also be summarized in this table. The reason for permanent discontinuation of study participation will be also summarized by study period (Screening Period, Treatment Period) The reason for either discontinuation of study treatment or study participation may include any of the following:

- Adverse Event
- Death
- Protocol Deviation
- Pregnancy
- Non-Compliance With Study Drug

- Withdrawal by Subject
- Physician Decision
- Study Terminated by Sponsor
- Site Terminated by Sponsor
- Study Terminated by IRB / ERB
- Lost to follow up
- Other

4.3. Primary Endpoint Analysis

4.3.1. Definition of endpoint

The primary endpoint for this study is change from baseline in HbA1c at Week 12. This endpoint will be used to evaluate the primary objective of the study.

4.3.2. Main analytical approach

The primary objective based on the efficacy estimand defined in Section 1.1 will be evaluated using the EAS dataset (Section 3). The primary analysis model for HbA1c measurements over time will be a mixed model for repeated measures (MMRM).

The observed changes in HbA1c level from baseline (Visit 2 on Week 0) to each of the scheduled post-baseline visits (Visit 3 on Week 4, Visit 4 on Week 8, and Visit 5 on Week 12) of each participant are the dependent variable. The baseline HbA1c value is incorporated as a covariate, and the visits are treated as fixed effects.

Mean change from baseline at Week 12 will be estimated by least square means (LS means), with accompanying P-value and 95% CI.

Operationally, the hypothesis defined in Section 2 will be tested at a two-sided 0.05 level by performing the test on the least square mean of changes from baseline at Week 12.

An unstructured covariance pattern will be used to estimate the variance-covariance of the within-participant repeated measures. Parameters will be estimated by REML with the Newton-Raphson algorithm and using the Kenward-Roger method for calculating the denominator degrees of freedom.

In case the model fails to converge, the following covariance structures will be tested in the following order:

1. Heterogenous Toeplitz
2. Heterogeneous First Order Autoregressive
3. Heterogeneous Compound Symmetry
4. Toeplitz
5. First Order Autoregressive, and
6. Compound Symmetry

The first covariance structure that converges will be used.

Missing data will not be imputed because the maximum likelihood estimate is unbiased for this model in case of a Missing at Random (MAR) mechanism. The P-value obtained for the least

square means for change from baseline will be used to determine the statistical significance to achieve the primary objective.

HbA1c values and change from baseline in HbA1c will also be summarized by visit using descriptive statistics.

4.4. Secondary Endpoint Analysis

Endpoints for secondary objectives will be evaluated based on the primary efficacy estimand. Missing data will be addressed by the MMRM model.

4.4.1. Definition of endpoints

The following secondary endpoints are defined:

- Change from baseline in %TAR:
  - Change from baseline to Week 12 in percentage of time per day that CGM-derived values are >180 mg/dl (10 mmol/L)
    - Percentage of time above range for a given day will be calculated as 100 times the number of CGM readings where a participant's blood glucose falls above the specified range during a day (00:00 to 23:59, inclusive), divided by the total number of CGM readings in the time interval (00:00 to 23:59, inclusive). This will be averaged over the days from Day 1 to Day 14 to obtain the %TAR for a participant.
- Change from baseline in TAR:
  - Changed from baseline to Week 12 in duration of time in minutes per day that CGM-derived values are >180 mg/dl (10 mmol/L)
    - Duration of time in minutes will be calculated as the percentage of time above range (%TAR) multiplied by the length of the time period (1440 minutes).
- Change from baseline to Week 12 in Fasting Serum Glucose (FSG)
- Change from baseline to Week 12 in weight

4.4.2. Main analytical approach

The same analytical approach will be applied for all secondary endpoints as for the primary endpoint, which is detailed under Section 4.3.2.

For secondary endpoints, MMRM will include baseline HbA1c (≤8.0%, >8.0%) category as a covariate instead of continuous baseline HbA1c.

Table 2. Summary of the analysis of Secondary Endpoints.

Objective	Efficacy Measure	Analysis Conducted	Additional information
CGM parameters	Change from baseline in %TAR at Week 12	MMRM as specified in Section 4.3.2.	MMRM with terms for: • Visits • Baseline HbA1c category (≤8.0%, >8.0%) • Baseline %TAR
CGM parameters	Change from baseline in TAR at Week 12	MMRM as specified in Section 4.3.2.	MMRM with terms for: • Visits • Baseline HbA1c category (≤8.0%, >8.0%) • Baseline TAR
FSG	Change from baseline at Week 12 in FSG	MMRM as specified in Section 4.3.2.	MMRM with terms for: • Visits • Baseline HbA1c category (≤8.0%, >8.0%) • Baseline FSG
Weight	Change from baseline at Week 12 in weight	MMRM as specified in Section 4.3.2.	MMRM with terms for: • Visits • Baseline HbA1c category (≤8.0%, >8.0%) • Baseline weight

4.5. Exploratory Endpoints Analysis

The objective of the exploratory analysis is to evaluate changes in CGM-derived measures from baseline over the course of 12 weeks. Endpoints for exploratory objectives will be evaluates based on the efficacy estimand, using the EAS.

The following exploratory endpoints will be evaluated:

- TIR – Time in Range (71 to 189 mg/dL or 3.9 to 10 mmol/L)
  - Calculated as a percentage of time per day:
    - Percentage of time in range for a given day will be calculated as 100 times the number of CGM data readings where a participant's blood glucose falls within the specified range during a day (00:00 to 23:59, inclusive), divided by the total number of CGM readings in the time interval (00:00 to 23:59, inclusive). This will be averaged over the days from Day 1 to Day 14 to obtain the TIR for a participant at a visit.
- TBR – Time Below Range (≤70 mg/dl or ≤3.9 mmol/L)
  - Calculated as a percentage of time per day:
    - Percentage of time below range for a given day will be calculated as 100 times the number of CGM data readings where a participant's blood glucose falls below the specified range during a day (00:00 to 23:59, inclusive), divided by the total number of CGM readings in the time interval (00:00 to 23:59, inclusive). This will be averaged over the days from Day 1 to Day 14 to obtain the TBR for a participant at a visit.
- GV – Glycemic Variability
  - Calculated as glycemic standard deviation (SD):
    - Within-day SD (SD_w)

      Each participant should have one SD_w per CGM session. SD_w is calculated by calculating the SD individually on all valid days within a session and then averaging across those days. Specifically,

      $$SD_w = \frac{\sum_{j=1}^{m} SD_j}{\mathrm{m}}$$

      where m stands for the number of valid CGM days in the session and SD_j is the standard deviation of blood glucose on day j,

      $$SD_j = \sqrt{\frac{\sum_{i=1}^{n_j}\left(BG_{ij} - \overline{BG_j}\right)^2}{n_j - 1}}$$

      where BG_{ij} is the blood glucose measure at time i on day j, n_j is the number of blood glucose measures on day j, and

      $$\overline{BG_j} = \frac{\sum_{i=1}^{n_j} BG_{ij}}{n_j}$$

- Between-day SD of the daily average (SD_b)

  Each participant should have one SD_b per CGM session. SD_b is calculated by averaging the glucose values individually on each valid day within a session and then calculating the SD across those days. Specifically,

$$SD_b = \sqrt{\frac{\sum_{j=1}^{m}\left(\overline{BG_j} - \frac{\sum_{j=1}^{m}\overline{BG_j}}{\mathrm{m}}\right)^2}{m-1}}$$

- Calculated as glucose coefficient of variation (CV):
  - Within-day CV (CV_w)

    Each participant should have one CV_w per CGM session. CV_w is calculated by calculating the CV individually on all valid days withing a CGM session and then averaging those CVs across all days in the session:

$$CV_w = \frac{\sum_{j=1}^{m} CV_j}{m}$$

    where m stands for the number of valid CGM days in a session and

$$CV_j = \frac{SD_j}{\overline{BG_j}} \times 100$$

  - Between-day CV (CV_b)

    Each participant should have one CV_b per CGM session. CV_b is calculated by averaging the glucose values individually on each valid day within a session and then calculating the CV across those days:

$$CV_b = \left[SD_b \Big/ \frac{\sum_{j=1}^{m}\overline{BG_j}}{m} \right] \times 100$$

- TITR – Time in Tight Target Range (71 to 140 mg/dL or 3.9 to 7.8 mmol/L)
  - Calculated as percentage of time per day in tight target range:
    - Percentage of time in tight target range for a given day will be calculated as 100 times the number of CGM data readings where a participant's blood glucose falls within the specified range during a day (00:00 to 23:59, inclusive), divided by the total number of readings in the time interval (00:00 to 23:59, inclusive). This will be averaged over the days from Day 1 to Day 14 to obtain TITR for a participant at a visit.
- Composite of all in ranges
  - Composite endpoint of the recommended time in all three International CGM Consensus Guidelines ranges (<4% in TBR [≤70 mg/dL], >70% in TIR, and <25% in TAR [>180 mg/dL]):

    - A binary endpoint will be created. If for a participant at a visit all three CGM measures (TAR, TIR, TBR) are within the recommended time-ranges then the composite endpoint will be set to "Y" for that participant, otherwise it will be set to "N".

- Mean 24-hr daily blood glucose

  Calculated by averaging all non-missing glucose readings a day (00:00 to 23:59, inclusive) and averaging across the 14 days to obtain mean daily glucose for a participant at a visit.

- Fasting glucose and post-prandial glucose excursion

  Both the endpoints are algorithmically derived from the entire CGM dataset. An algorithm reviews the entire CGM dataset, evaluates excursions, and time periods then estimates either the fasting glucose or post-prandial glucose excursion.

Summary statistics will be provided for actual values by visit (Baseline, Week 4, and Week 12) and changes from baseline values by each post-baseline visit (Week 4 and Week 12). The odds for a subject to reach a composite endpoint of the recommended time in all three International CGM Consensus Guideline ranges will be compared between baseline and Week 12 by a logistic regression model. The model will also include, baseline HbA1c (≤8.0%, >8.0%), and baseline composite endpoint as fixed effect.

4.5.1. Summary for the Analysis of Exploratory Measures

Table 3. Exploratory measures not controlled for Type 1 error

Objective	Relative to the efficacy measure	Analysis Conducted	Additional information
CGM parameters	Mean change in TIR from baseline at week 12	MMRM model as specified in Section 4.3.2.	MMRM with terms for: • Visit • Baseline HbA1c category (≤8.0%, >8.0%), and • Baseline TIR as a covariate
CGM parameters	Mean change in TBR from baseline at week 12	MMRM model as specified in Section 4.3.2.	MMRM with terms for: • Visit • Baseline HbA1c category (≤8.0%, >8.0%), and • Baseline TBR as a covariate
CGM parameters	Mean change in TITR from baseline at week 12	MMRM model as specified in Section 4.3.2.	MMRM with terms for: • Visit • Baseline HbA1c category (≤8.0%, >8.0%), and • Baseline TITR as a covariate
CGM parameters	Mean change in GV from baseline at week 12	MMRM model as specified in Section 4.3.2.	Separate MMRM model for all measures of GV. MMRM with terms for: • Visit • Baseline HbA1c category (≤8.0%, >8.0%), and • Baseline GV measure as a covariate
CGM parameters	Change in odds for reaching composite endpoint of all in ranges from baseline to week 12	Logistic regression model as specified in Section 4.5	Logistic Regression with terms for: • Visit

			• Baseline HbA1c category (≤8.0%, >8.0%), and • Baseline composite endpoint as a covariate
CGM parameters	Mean change in mean 24-hr daily blood glucose from baseline at week 12	MMRM model as specified in Section 4.3.2.	MMRM with terms for: • Visit • Baseline HbA1c category (≤8.0%, >8.0%), and • Baseline mean 24-hr daily blood glucose as a covariate
CGM parameters	Mean change in estimated post-prandial glucose excursion from baseline at week 12	MMRM model as specified in Section 4.3.2.	MMRM with terms for: • Visit • Baseline HbA1c category (≤8.0%, >8.0%), and • Baseline estimated post-prandial glucose excursion as a covariate
CGM parameters	Mean change in estimated fasting glucose at week 12	MMRM model as specified in Section 4.3.2.	MMRM with terms for: • Visit • Baseline HbA1c category (≤8.0%, >8.0%), and • Baseline estimated fasting glucose as a covariate

4.6. Safety Analyses

Safety assessments will be done using the Safety Analysis Set (SS) irrespective of adherence to study intervention or initiation of prohibited medication, unless indicated otherwise.

4.6.1. Extent of Exposure

Duration of exposure is defined as the total number of days a subject is exposed to any study intervention and will be presented as the total number of days from the first dose date (Day 1) to the last dose date (date of last dose minus the date of first dose + 1) as recorded on the Disposition Event: Treatment / Dosing pages on the CRF. If the last dose date on Disposition Event: Treatment / Dosing pages is missing, or if a participant is lost to follow-up, but the drug accountability log confirms that the participant has taken study drug, the visit date following the last completed drug accountability log will be used.

Missing doses will be ignored in this calculation.

The duration of exposure to study intervention will be summarized for all participants in the SS and will be presented in a table by summary statistics. The duration of exposure will then be classified into one of the following categories: > 0 Week, >= 4 Weeks, >= 8 Weeks, and >= 12 Weeks and will be presented as the number and percentage of participants in each duration category. Percentages will be computed from the number of participants in the SS.

4.6.2. Adverse Events

All adverse events will be classified by System Organ Class (SOC) and Preferred Term (PT) according to the Medical Dictionary for Regulatory Activities. Selected notable AEs of interest may be reported using high-level terms or Standardized Medical Dictionary for Regulatory Activities Queries. As an example, 'renal safety event', defined in the protocol as an AE of interest, is defined by the high-level group term 'Renal disorders (excl nephropathies)'. Please refer to Section 6.3 for the SMQs of interest.

A treatment-emergent AE (TEAE) is defined as an AE that meets any of the following conditions:

- begins on or after the first dose of study intervention
- begins before the first dose of study drug and worsens in severity on or after the first dose of study drug
- is completely missing an onset date and end date
- is completely missing an onset date and the end date is on or after the first dose of study intervention.

An overview summary of the number and percentage of participants with any TEAE, serious TEAE, study drug-related TEAE, study drug-related serious TEAE, TEAE leading to treatment discontinuation, TEAE leading to study termination, and AE leading to death will be provided. The number of events will also be summarized.

Local Adverse Events (i.e., injection site reactions) will be tabulated separately in addition to the overall TEAE tables.

Summaries of the total number of TEAEs and the number and percentage of participants with at least one TEAE will be provided. The number and percentage of participants and the number of events will also be presented by SOC and PT. At each level of participant summarization, a participant is counted once if the participant reported one or more events. Percentages will be calculated out of the number of participants in the SS. The number of events at each level of SOC and PT will also be summarized.

A summary of TEAEs will also be presented in descending order from the SOC with the highest total incidence (that is, summed across treatment group) to the SOC with the lowest total incidence. If the total incidence for any two or more SOCs is equal, the SOCs will be presented in alphabetical order. Within each SOC, the PTs will be presented in alphabetical order.

A summary of AEs where the answer to "*Outcome*" is "*Fatal*" will be presented in a table. At each level of participant summarization, a participant is counted once if the participant reported one or more events. Percentages will be calculated out of the number of participants in the Safety set within the subgroup category.

A summary of TEAEs with a study intervention action taken of "*Drug Withdrawn*" will be presented in a table. At each level of participant summarization, a participant is counted once if the participant reported one or more events. Percentages will be calculated out of the number of participants in the Safety set within the subgroup category.

A summary of TEAEs which are linked to study discontinuation via the "Disposition Event: Treatment Phase" CRF page will be presented in a table. At each level of participant summarization, a participant is counted once if the participant reported one or more events. Percentages will be calculated out of the number of participants in the Safety set within the subgroup category.

A summary of TEAEs by severity will be presented in a table. The severity that will be presented represents the most extreme severity captured on the CRF page. The possible severities are *"Mild," "Moderate," and "Severe."* In the TEAE severity table, if a participant reported multiple occurrences of the same TEAE, only the most severe will be presented. Treatment-emergent AEs that are missing severity will be presented in tables as "*Severe*" but will be presented in the data listing with a missing severity. Percentages will be calculated out of the number of participants in the Safety set.

A summary of TEAEs by relationship (i.e., "*Related*" and "*Not Related*") to study drug will be presented in a table by incidence of occurrence. The investigator will provide an assessment of the relationship of the event to the study drug. The possible relationships are *"Not Related", "Related"*. Treatment –emergent AEs that are missing a relationship will be presented in the summary table as "Related" but will be presented in the data listing with a missing relationship. Percentages will be calculated out of the number of participants in the Safety set.

The seriousness of an AE should be assessed by the Investigator independently from the severity of the AE. A serious AE (SAE) is defined as any untoward medical occurrence that at any dose results in death, is life-threatening, is a congenital anomaly/birth defect, requires in-patient hospitalization or prolongation, or results in significant disability/incapacity.

Important medical events that may not result in death, be life-threatening, or require hospitalization may be considered SAEs when, based upon medical judgment, they may jeopardize the participant and may require medical or surgical intervention to prevent one of the outcomes listed above.

Serious treatment-emergent adverse events (SAEs) will be presented in a table. Treatment-emergent SAEs by relationship to study drug will be presented in a table. A treatment-related treatment-emergent SAE is a treatment-emergent SAE with the answer "Yes" to the question "Is this event related to study treatment?" on the SAE CRF page. Treatment-emergent SAEs that are missing a relationship will be presented in the table as "Related" but will be presented in the data listing with a missing relationship. At each level of participant summarization, a participant is counted once if the participant reported one or more events. Percentages will be calculated out of the number of participants in the Safety set.

The treatment-emergent SAE data will be categorized and presented by SOC and PT in a manner similar to that described for TEAEs.

The number and percentage of participants who prematurely discontinue the study due to an AE will be summarized using MedDRA PT nested within SOC. Events will be ordered by decreasing

frequency within SOC. A listing of all the discontinuation from Study due to Adverse Event will be provided.

The number and percentage of participants who prematurely discontinue study drug due to an AE will be summarized using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency within SOC. A listing of all the discontinuation from Study Treatment due to Adverse Event will be provided.

All participants who have an AE related to study treatment with an outcome of "Death Related to Adverse Event" will be presented in a listing. The listing will include patient identification including the treatment, site number, date of death, age at the time of enrollment, gender, MedDRA PT of associated AE, time from first dose of study drug to death, time from last dose of study drug to death (if patient had discontinued study drug), cause of death as reported by investigator, cause of death as adjudicated by Clinical Endpoint Committee (CEC).

Patient narratives will be provided for all patients who experience any of the following "notable" events:

- Deaths
- SAEs
- Pregnancy
- Permanent Discontinuations of study treatment due to AEs

All AEs will be presented in a listing.

Handling of missing TEAE start and end dates:

For the purpose of inclusion in TEAE tables, incomplete AE onset and end dates will be imputed as follows:

Missing onset dates (where UK and UKN indicate unknown or missing day and month respectively):

- UK-MMM-YYYY: If the month and year are different from the month and year of the first dose of study drug, assume 01-MMM-YYYY. If the month and year are the same as the first dose of study drug month and year, and the end date (after any imputation) is on or after the first dose of study drug, then assume the date of the first dose of study drug. If the month and year are the same as the first dose of study drug month, and year and the end date (after any imputation) is prior to the first dose of study drug, then assume the end date for the onset date.
- DD-UKN-YYYY/UK-UKN-YYYY: If the year is different from the year of first dose of study drug, assume 01-JAN-YYYY of the collected year. If the year is the same as the first dose of study drug year, and the end date (after any imputation) is on or after the first dose of study drug, then assume the date of the first dose of study drug. If the year is the same as the first dose of study drug, and the end date (after any imputation) is prior to the first dose of study drug, then assume the end date for the onset date.

Missing end dates (where UK and UKN indicate unknown or missing day and month respectively):

- UK-MMM-YYYY: Assume the last day of the month.

- DD-UKN-YYYY/UK-UKN-YYYY: Assume 31-DEC-YYYY.

4.6.3. Product Complaints

A product complaint is any written, electronic, or oral communication that alleges deficiencies related to the identity, quality, durability, reliability, safety, effectiveness, or performance of a study intervention. When the ability to use the study intervention safely is impacted.

Product complaints related to study interventions used in clinical studies are collected in order to ensure the safety of participants, monitor quality, and to facilitate process and product improvements.

Investigators will instruct participants to contact the site as soon as possible if he or she has a product complaint or problem with the study intervention so that the situation can be assessed.

An event may meet the definition of both a product complaint and an AE/SAE. In such cases, it should be reported as both a product complaint and as an AE/SAE.

A listing of all device product complaints, inclusive of device product complaints that lead to an AE or that could have led to an SAE had intervention not been taken will be provided. Additional summaries will be provided as deemed appropriate.

4.6.4. Additional Safety Assessments

4.6.4.1. Central Laboratory Measures and Vital Signs

Summary tables will be presented for clinical laboratory tests with numeric values for participants in the Safety set. Observed results at each visit will be presented.

Summary tables presenting observed values and changes from baseline (for those parameters which are measured post-baseline) will be presented for clinical laboratory tests with numeric values for participants in the Safety set. Changes from baseline to each scheduled post-baseline visit will be presented.

A listing of abnormal findings will be created for laboratory analyte measurements. The listing will include patient ID, laboratory collection date, study day, analyte name, and analyte finding.

For clinical chemistry, the following laboratory tests will be included:

Sodium, Potassium, Chloride, Bicarbonate, Total bilirubin, Direct bilirubin, Alkaline phosphatase (ALP), Alanine aminotransferase (ALT), Aspartate aminotransferase (AST), Gamma-glutamyl transferase (GGT), Blood urea nitrogen (BUN), Creatinine, Creatine kinase (CK), Uric acid, Total protein, Albumin, Calcium, Phosphorus, Glucose.

All chemistry data by participant will be presented in a listing.

For hematology, the following laboratory tests will be included:

Hemoglobin, Hematocrit, Erythrocyte count (RBCs - red blood cells), Mean cell volume, Mean cell hemoglobin, Mean cell hemoglobin concentration, Leukocytes (WBCs - white blood cells), Differential (Neutrophils, Lymphocytes, Monocytes, Eosinophils, Basophils, Platelets, Cell morphology (RBCs and WBCs) if indicated.

All hematology data by participant will be presented in a listing.

For lipid panel, the following laboratory tests will be included:

Total cholesterol, Triglycerides, High-density lipoprotein (HDL-C), Low-density lipoprotein (LDL-C), Very low-density lipoprotein (VLDL-C).

All lipid panel data will be presented in a listing.

As for urine chemistry, the following tests will be included:

Albumin, Creatinine

All urine chemistry data will be presented in a listing.

Additional tests:

Calcitonin

All additional tests will be presented in a listing.

Calculations:

Urinary albumin/creatinine ratio (UACR), eGFR (CKD-EPI) calculated using creatinine. All calculations will be listed.

Laboratory Tests for Hepatic Evaluation

- Hepatic Hematology Panel:
  Hemoglobin, Hematocrit, Erythrocytes (RBCs - red blood cells), Leukocytes (WBCs - white blood cells), Differential (Neutrophils, segmented Lymphocytes, Monocytes, Basophils, Eosinophils, Platelets, Cell morphology (RBC and WBC).
- Hepatic Clinical Chemistry Panel:
  Total bilirubin, Direct bilirubin, Alkaline phosphatase (ALP), Alanine aminotransferase (ALT), Aspartate aminotransferase (AST), Gamma-glutamyl transferase (GGT), Creatine kinase (CK)
- Hepatic Coagulation Panel:
  Prothrombin time, INR (PT-INR)
- Urine Chemistry:
  Drug screen
- Haptoglobin
- Hepatitis A virus (HAV) testing:
  HAV total antibody, HAV IgM antibody
- Hepatitis B virus (HBV) testing:
  Hepatitis B surface antigen (HBsAg), Hepatitis B surface antibody (anti-HBs), Hepatitis B core total antibody (anti-HBc), Hepatitis B core IgM antibody, HBV DNA
- Hepatitis C virus (HCV) testing:
  HCV antibody, HCV RNA
- Hepatitis D virus (HDV) testing:
  HDV antibody

- Hepatitis E virus (HEV) testing:
  HEV IgG antibody, HEV IgM antibody, HEV RNA
- Anti-nuclear antibody (ANA)
- Anti-smooth muscle antibody (ASMA)
- Anti-actin antibody
- Epstein-Barr virus (EBV) testing:
  EBV antibody

Laboratory assessments will be performed by a central laboratory. All summaries will be based on the units provided by the central laboratory, no conversion will be done.

All relevant clinical laboratory tests will be classified as treatment emergent Low, Normal, and High, or Normal/Abnormal according to the normal ranges. This categorical data will be summarized in shift tables comparing the results at each scheduled post-baseline visit with those at the baseline visit. The percentages of participants with treatment-emergent abnormal, high, or low laboratory measures at each post-baseline visit will be summarized.

When there are multiple values within a visit for a particular laboratory variable, the worst value will be taken (worst being the smallest value for criteria below a certain threshold or the largest value for criteria above a certain threshold). If a participant has a value below the threshold and above the threshold, the value furthest from the threshold will be chosen.

A treatment-emergent abnormal value is defined as a change from normal value at baseline to an abnormal value at any time during the follow-up.

Definition of treatment emergent Low/High results:

- A treatment-emergent high result is defined as a change from a value less than or equal to the high limit at baseline to a value greater than the high limit at any time during the Treatment and Follow-up periods.
- A treatment-emergent low result is defined as a change from a value greater than or equal to the low limit at baseline to a value less than the low limit at any time during the Treatment and Follow-up periods.
    - Laboratory normal ranges will be provided by the Central Lab.

An MMRM using REML model will be used to analyze the changes in lipid panel parameters from baseline to all scheduled post-baseline visits. The model will include visits and baseline HbA1c category (≤8.0%, >8.0%) as fixed effects and baseline value of the dependent variable as a covariate.

<u>Vital signs:</u>

Summary tables presenting observed values and changes from baseline will be presented for vital sign data, *systolic blood pressure (mmHg), diastolic blood pressure (mmHg), pulse rate (bpm),* weight (kg), Waist Circumference measurement (cm) and BMI for participants in the Safety set. For height (cm), only observed values will be summarized at baseline.

Height and weight measurements will be used to calculate BMI, with following formula:

$$BMI = weight\ (kg) / [height\ (m)]^2$$

Descriptive summaries by nominal visit will be provided for baseline and postbaseline values as well as change from baseline values. If 2 records are taken at the same visit, they will be averaged prior to being used for data summaries and analyses.

An MMRM using REML model will be used to analyze the changes in vital signs from baseline to all scheduled post-baseline visits. The model will include visits and baseline HbA1c category (≤8.0%, >8.0%) as fixed effects and baseline value of the dependent variable as a covariate.

Counts and percentages of patients with treatment-emergent abnormal sitting systolic blood pressure (BP), sitting diastolic BP, and pulse will be presented. The criteria for identifying patients with treatment-emergent vital sign abnormalities are stated in Table 4.

Table 4. Categorical Criteria for Abnormal Treatment-Emergent Blood Pressure and Pulse Measurement

Parameter	Low	High
Systolic BP (mm Hg) (Supine or sitting – forearm at heart level)	Observed value ≤ 90 and decrease from baseline ≥ 20	Observed value ≥ 140 and increase from baseline ≥ 20
Diastolic BP (mm Hg) (Supine or sitting – forearm at heart level)	Observed value ≤ 50 and decrease from baseline ≥ 10	Observed value ≥ 90 and increase from baseline ≥ 10
Pulse (bpm) (Supine or sitting)	Observed value ≤ 50 and decrease from baseline ≥ 15	Observed value ≥ 100 and increase from baseline ≥ 15

4.6.4.2. Hypoglycemic events

The following hypoglycemic events (HEs) will be summarized respectively:

- Level 1 hypoglycemia
  - Glucose <70 mg/dL (<3.9 mmol/L) and ≥54 mg/dL (≥3.0 mmol/L)
  - Level 1 hypoglycemia can alert a person to take action such as treatment with fast-acting carbohydrates. Providers should continue to counsel participants to treat hypoglycemia at this glucose alert value.
- Level 2 hypoglycemia
  - Glucose <54 mg/dL (<3.0 mmol/L)
  - Level 2 hypoglycemia is also referred to as documented or BG-confirmed hypoglycemia with glucose <54 mg/dL (<3.0 mmol/L). This glucose threshold is clinically relevant regardless of the presence or absence of symptoms of hypoglycemia.
- Level 3 hypoglycemia (Severe)

        - A severe hypoglycemic event is characterized by altered mental or physical status requiring assistance of another person to actively administer carbohydrate, glucagon, or other resuscitative actions for the treatment of hypoglycemia.
        - The determination of an episode of severe hypoglycemia is made by the investigator based on the medical need of the participant to have required assistance and is not predicated on the report of a participant simply having received assistance. Those hypoglycemic events for which the answer is "yes" to the question "Did the subject experience a severe hypoglycemic episode with neurological (cognitive) impairment requiring assistance?" on the Hypoglycemic Events eCRF page and/or it is recorded as a SAE will be classified as Level 3.
- Severe hypoglycemia
    - A severe hypoglycemic event is characterized by altered mental or physical status requiring assistance of another person to actively administer carbohydrate, glucagon, or other resuscitative actions for the treatment of hypoglycemia
    - Examples of severe hypoglycemia in adults are
        - altered mental status and the inability to assist in their own care
        - semiconscious or unconscious, or
        - coma with or without seizures.
- Nocturnal hypoglycemia
    - Nocturnal hypoglycemia is a hypoglycemia event, including severe hypoglycemia, which occurs at night and presumably during sleep.
    - Defined as a hypoglycemia event, including severe hypoglycemia, which occurs at night and presumably during sleep. If for a hypoglycemia event the answer to the question "When did the hypoglycemic event occur?" is "Between bedtime and waking (while asleep)" on the Hypoglycemic Events eCRF page, then that event will be classified as nocturnal hypoglycemia.

To avoid duplicate reporting, all consecutive BG values occurring within a 1-hour period may be considered to be a single hypoglycemic event and the one with the lowest BG value can be selected to be the representative. If two records with the same lowest values within a hour, the earlier occurrence will be selected.

Statistical summaries and analyses will exclude hypoglycemic events occurring after initiation of a new antihyperglycemic therapy. For severe hypoglycemia and level 2 hypoglycemia incidence as well as rate per patient per year of exposure will be provided at specified time intervals.

For each HEs defined above, the total number of the respective HEs occurring between Week -3 and Day 1 will be used to calculate baseline frequency of total HEs respectively. Similarly, for each HEs defined above, the total number of respective HEs occurring in the 14 calendar days prior to Visit 3 (Week 4), Visit 4 (Week 8), and Visit 5 (Week 12) will be used to calculate frequency of total HEs for each HE defined above at each visit respectively.

Hypoglycemia events will be summarized by descriptive statistics. Summaries will include the number and percentage of participants experiencing at least 1 event and the total number of events for each visit. Where possible, the summaries will be presented by Level as well.

The incidence of hypoglycemic event will be analyzed using logistic regression with baseline HbA1c category (≤8.0%, >8.0%) as fixed effects.

The rate of hypoglycemic episodes per patient month may be analyzed using a generalized linear MMRM assuming the number of hypoglycemic episodes follows a negative binomial distribution with the mean modeled using visits and baseline HbA1c category (≤8.0%, >8.0%) as fixed effects if data warrants. The logarithm of months in specified time interval will be adjusted as an offset to account for possible unequal treatment duration in the specified time interval between patients. When the number of hypoglycemic events is less than 10, the listing of hypoglycemic events will be provided instead.

Data collected on the Hypoglycemic Events CRF page will be presented in listings.

4.6.4.3. Severe Persistent Hyperglycemia

Defined as Fasting Blood Glucose (via SMBG) >270 mg/dL (>15 mmol/L) during the first 4 weeks or >240 mg/dL (>13.3 mmol/L) from Weeks 5-12 occurring for at least 2 consecutive weeks. The number and percentage of participants with severe, persistent hyperglycemia will be summarized and data will be listed.

4.6.4.4. Gastrointestinal Events

The following gastrointestinal events will be summarized respectively:

- Nausea
- Vomiting
- Diarrhea
- Abdominal Pain/Abdominal Cramps
- other

These events are recorded on the "*Hypersensitivity, Anaphylactic and Infusion/Injection Related Reaction Follow-up*" eCRF page. If the answer to the question "Does subject have gastrointestinal signs or symptoms?" is "Yes" then the applicable event type is recorded as well ("Nausea", "Vomiting", "Diarrhea", "Abdominal Pain/Abdominal Cramps", "Other"). The events are defined based on this information. Additionally, this information is linked to the AE page where their severity is also captured.

A summary of the number of gastrointestinal events by severity and event type will be presented in a table, and the number and percentage of participants with at least one event will be provided. The severity that will be presented represents the most extreme severity captured on the CRF page. The possible severities are *"Mild," "Moderate," and "Severe."* If a participant reported multiple occurrences of the same event, only the most severe will be presented. Events that are missing severity will be presented in tables as "*Severe*" but will be presented in the data listing with a missing severity. Percentages will be calculated out of the number of participants in the Safety set. The time courses of prevalence and incidence (newly-occurring episodes) of nausea, vomiting, diarrhea, and combined will be plotted and maximum severity.

The maximum severity and duration of treatment-emergent nausea, vomiting, diarrhea, and combined through the end of the study will be summarized.

4.6.4.5. Hypersensitivity, Anaphylactic and Infusion/Injection Related Reaction Follow-up

Data collected on Hypersensitivity, Anaphylactic and Infusion/Injection Related Reaction Follow-up CRF page will be listed. Information to be summarized includes the timing of the reaction relative to study drug administration, and characteristics of the injection site reaction: erythema, induration, pain, pruritis, and edema. Patient based and event-based summaries will be created.

Data collected on this page for gastrointestinal events will be summarized (See Section 4.6.4.4). Additionally, summaries will be created for mucocutaneous, cardiovascular, central nervous system, respiratory and general signs or symptoms in a similar way as for gastrointestinal events. Percentages will be calculated out of the number of participants in the Safety set.

4.6.4.6. Hepatic evaluation

Liver Biopsy Assessment, Hepatic Monitoring Procedures, Hepatic Risk Factor Assessment and Liver Related Signs and Symptoms data will be presented in listings.

- Hepatobiliary Disorders:

  The AE database will be searched using Standardized MedDRA Queries (SMQs) to identify events consistent with hepatobiliary disorders. Detailed searching criteria can be found in Section 6.3 (Appendix 3). A summary by PT within SMQ will be provided.

- Acute Gallbladder Disease

  The AE database will be searched using predefined SMQs to identify events consistent with acute gallbladder diseases. Detailed searching criteria can be found in Section 6.3 (Appendix 3). A summary by PT within SMQ will be provided.

- Liver Enzymes:

  Analyses for laboratory analyte measurements are described in Section 4.6.4.1. This section describes additional analyses of liver enzymes. In addition, the following will be provided:

  - A shift table of maximum to maximum alanine aminotransferase (ALT) measurement from baseline (≤1 × ULN, > 1 × ULN) to postbaseline with the following categories - ≤1 × ULN, >1 to <3 × ULN, ≥3 to <5 × ULN, ≥5 to <10 × ULN, ≥10 × ULN
  - A shift table of maximum to maximum aspartate transaminase (AST) measurement from baseline (≤1 × ULN, > 1 × ULN) to postbaseline with the following categories - ≤1 × ULN, >1 to <3 × ULN, ≥3 to <5 × ULN, ≥5 to <10 × ULN, ≥10 × ULN.
  - Shift tables of maximum to maximum total bilirubin and direct bilirubin from baseline to postbaseline with the following categories - ≤1 × ULN, > 1 to <2 × ULN, ≥2 × ULN.
  - Shift tables of serum alkaline phosphatase from baseline to postbaseline with the following categories - ≤1 × ULN, >1 to <2 × ULN, ≥2 × ULN.

  Of note, maximum baseline will be the maximum non-missing observation in the baseline period. The maximum postbaseline value will be the maximum non-missing

value from the postbaseline period. Planned and unplanned measurements will be included.

4.6.4.7. Acute pancreatitis

Acute pancreatitis is an AE of special interest in all studies with tirzepatide, including this study.

AE listing of acute pancreatitis will be provided separately from other AEs. Determination of reported events will be through the pre-defined SMQ search for acute pancreatitis and MedDRA PT of pancreatitis chronic. Detailed searching criteria can be found in Section 6.3 (Appendix 3).

4.7. Other Analyses

4.7.1. Subgroup analyses

Subgroup analyses of the primary and secondary endpoint will be made to assess consistency of the improvement from baseline across subgroups.

The primary and secondary endpoints will be examined using the same MMRM model defined in Section 4.3.2. but additionally, a SUBGROUP term will be included as fixed effects. LS-means with their corresponding 95% CIs and P-values will be reported for the subgroups.

The following subgroups are defined:

- Age group: < 65 years vs ≥ 65 years
- Sex: female vs male
- Race: white vs black vs other
- Ethnicity: Hispanic vs not Hispanic
- Baseline GLP-1 RA: liraglutide vs semaglutide vs dulaglutide
- Baseline GLP-1 RA equivalent doses (adapted from Almandoz et al. 2020):
    - dulaglutide 0.75 mg and liraglutide 1.2 mg vs.
    - dulaglutide 1.5 mg, semaglutide 0.5 mg, and liraglutide 1.8 mg vs.
    - dulaglutide 3 mg and semaglutide 1 mg vs.
    - dulaglutide 4.5 mg and semaglutide 2 mg
- Baseline HbA1c (≤8.0%, >8.0%)

If the number of participants is too small (less than [10%]) within a subgroup, then the subgroup categories may be redefined prior to unblinding the study.

4.8. Interim Analyses

No interim analyses are planned for this study. If an unplanned interim analysis is deemed necessary for reasons other than a safety concern, the protocol must be amended.

4.8.1. Data Monitoring Committee (DMC) or Other Review Board

No DMC or other review board meeting is planned.

4.9. Changes to Protocol-Planned Analyses

There are no changes to protocol planned analyses.

5. Sample Size Determination

A sample size of 150 participants will provide at least 90% power to show an improvement in HbA1c from baseline at Week 12. This sample size is using the following assumptions:

- a 1-sample t-test with 2-sided significance level of 0.05
- within treatment difference from baseline at Week 12 is 0.3%, and
- intra-subject variability of 1.1%.

6. Supporting Documentation

6.1. Appendix 1: Demographic and Baseline Characteristics

A summary of demographics and baseline information will be presented. The demographic characteristics consist of age (in years), sex, race, and ethnicity. The baseline characteristics consist of baseline height (cm), baseline weight (kg), and Waist Circumference measurement (cm), fundoscopy exam results and substance use. All demographic and baseline characteristics will be presented in listings.

A participant's age in years is calculated using the date of the informed consent and date of birth. Age (years), baseline height (cm), baseline weight (kg), Waist Circumference measurement (cm), Systolic Blood Pressure (mmHg), Diastolic Blood Pressure (mmHg), and Pulse Rate (bpm) will be summarized using descriptive statistics.

The number and percentage of participants by age category (< 65 years vs ≥ 65 years), sex (Male, Female), race (White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander) and ethnicity (Hispanic or Latino, Not Hispanic or Latino), will also be reported.

Baseline fundoscopy results will be presented in a table for both eyes. For each eye respectively, number and percentage of participants with non-proliferative diabetic retinopathy will be presented by severity. Number and percentage of participants with proliferative diabetic retinopathy and macular edema will also be presented.

Percentages will be based on the total number of participants in the mITT set.

Collected data about substance use, including alcohol, caffein, recreational drugs, and tobacco will be listed.

6.1.1. Prior and Concomitant Medications/Therapy

All medications will be coded according to the World Health Organization drug dictionary

A prior medication/therapy is defined as those for which the end date is prior to the date of first study intervention administration or confirmed as 'not ongoing' by the participant before the start of study intervention. A concomitant medication/therapy is defined as any medication that has a stop date that is on or after the date of first dose of study intervention or confirmed as 'ongoing' by the participant at the start of study intervention. If a medication is not clearly indicated as ongoing nor the end date is documented, it will be treated as a concomitant medication.

For the purpose of inclusion in prior and/or concomitant medication tables, incomplete medication start and stop dates will be imputed as follows:

Missing start dates (where UK, UKN and UNKN indicate unknown or missing day, month and year respectively):

- UK-MMM-YYYY: Assume 01-MMM-YYYY. If the month and year are the same as the first dose of study intervention month and year, then assume the date of first dose of study drug.

- UK-UKN-YYYY: If the year is prior to the year of first dose of study intervention, assume 01-JAN-YYYY of the collected year. If the year is the same as the first dose of study intervention year, then assume the date of first dose of study intervention.
- UK-UKN- UNKN: Assume date of first dose of study intervention.

Missing stop dates for concomitant medications data will be handled as follows (where UK, UKN and UNKN indicate unknown or missing day, month, and year respectively):

- UK-MMM-YYYY: Assume the last day of the month.
- UK-UKN-YYYY: Assume 31-DEC-YYYY.
- UK-UKN- UNKN: Assume ongoing and leave it missing.

The total number of concomitant medications and the number and percentages of participants with at least one concomitant medication will be summarized. The number and percentages of all concomitant medications will be summarized and listed by Anatomical Therapeutic Chemical (ATC) level 4 and preferred term. Prespecified concomitant medications Acetaminophen / Paracetamol will be summarized in the same way.

The same summary will be presented for the prior medications/therapies as for concomitant medications.

The same summary will be presented for the background oral antihyperglycemic medications (OAMs), defined as medications belonging to the category of 'antihyperglycemics', taken via oral route, and starting prior to the start of study intervention.

All summaries will be performed using the Safety set.

6.1.2. Medical History

For Medical History, medical condition or event terms will be coded using the Medical Dictionary for Regulatory Activities (MedDRA).

The number and percentage of participants with any Pre-Existing Conditions and Medical History will be summarized overall and for each system organ class and preferred term. Participant Pre-Existing Conditions and Medical History data including specific details will be presented in a listing**.**

General Associated Person Medical History data will be listed for each participant. The number and percentage of participants having first degree relative with medical history will be summarized overall and for each system organ class and preferred term.

Associated Person Medical History data for Liver Disease will also be listed and the number (percentage) of participants associated with a certain medical condition or event (Alpha-1 antitrypsin Deficiency, Autoimmune disorder, Hemochromatosis) will be summarized by preferred term.

Prespecified Medical History including the date of diabetes diagnosis, Liver and Concurrent Disease will be listed. The number and percentage of participants with any Liver and Concurrent Disease will be summarized overall and for each system organ class and preferred term.

Percentages will be calculated based on number of participants in the mITT set.

6.2. Appendix 2: Treatment Compliance

Participants should administer Tirzepatide once weekly by SC injection.

Treatment overdose is defined as administration of more than a total of 3 SDPs (15mg) of Tirzepatid within a 72-hour period. If a dose is missed, the participant should take their dose as soon as possible, within 4 days (96 hours) after the missed dose. If more than 4 days have passed, that dose should be skipped, and the next dose should be taken at the appropriate time.

For each participant, study treatment compliance will be determined for each week by taking into account whether a participant takes all doses of study treatment as instructed based on the subject diary. If on a given week a participant overdosed treatment or a dose had to be skipped then that will count as non-compliance for that week. Weekly compliance along with weekly actual dose and skipped doses will be listed for each participant. Number and percentage of participants who were compliant, non-compliant (due to overdosing or due to skipping dose) will be summarized by week.

Furthermore, compliance will be characterized by accountability data for each visit. The number of SC injections taken will be calculated by subtracting the number of SDPs returned from the number of SDPs dispensed at a visit.

The study intervention compliance (%) for each visit will be calculated by dividing the total number of SDPs used up to the next visit by the total number of SDPs prescribed at each visit and then multiplying by 100. The overall study intervention compliance (%) will be calculated by dividing the total number of SDPs used at all visits by the total number of SDPs prescribed for all visits and then multiplying by 100.

Compliance (%) = [(total number of SDPs dispensed – total number of unused SDPs returned) / (Number of weeks in visit interval × Number of SDPs prescribed per week)] × 100.

Treatment compliance will be summarized descriptively. Furthermore, for overall compliance a categorical summary of whether participants were compliant (yes/no) will be presented. A participant is considered compliant if overall study intervention compliance is greater than or equal to 75%.

Percentages will be calculated out of the number of participants who were dosed at that dosing period in the Safety set.

6.3. Appendix 3: Search Criteria for Standardized MedDRA Queries

The Standardized MedDRA Queries of interest are listed as follows:

- Systemic hypersensitivity and anaphylactic reactions:
  - Anaphylactic reaction (SMQ), narrow search
  - Hypersensitivity (SMQ), narrow search
  - Angioedema (SMQ), narrow search
  - Severe cutaneous adverse reactions (SMQ), narrow search
- Hepatic disorders (SMQ)
- Gallbladder related disorders (SMQ)
- Acute pancreatitis (SMQ)
- Cardiac arrythmias (SMQ)

- Biliary disorders (SMQ)

The search criteria for each of the SMQs are stored in PPD Biostatistics Technology Infrastructure (BTI) in:

\\wilbtib\wilbtia00\Global Bios Library\Dictionaries\MedDRA\Current\MedAscii\smq_list.asc

7. References

Del Prato S, Kahn SE, Pavo I, et al. Tirzepatide versus insulin glargine in type 2 diabetes and increased cardiovascular risk (SURPASS-4): a randomised, open-label, parallel-group, multicentre, phase 3 trial. Lancet. 2021;398(10313):1811-1824. https://doi.org/10.1016/S0140 6736(21)02188-7

Frías JP, Davies MJ, Rosenstock J, et al. Tirzepatide versus semaglutide once weekly in patients with type 2 diabetes. N Engl J Med. 2021;385(6):503-515.https://doi.org/10.1056/NEJMoa2107519

Ludvik B, Giorgino F, Jódar E, et al. Once-weekly tirzepatide versus once-daily insulin degludec as add-on to metformin with or without SGLT2 inhibitors in patients with type 2 diabetes (SURPASS-3): a randomised, open-label, parallel-group, phase 3 trial. Lancet. 2021;398(10300):583-598. https://doi.org/10.1016/S0140-6736(21)01443-4

Rosenstock J, Wysham C, Frías JP, et al. Efficacy and safety of a novel dual GIP and GLP-1 receptor agonist tirzepatide in patients with type 2 diabetes (SURPASS-1): a double-blind, randomised, phase 3 trial. Lancet. 2021;398(10295):143-155. https://doi.org/10.1016/S0140-6736(21)01324-6. Erratum in: Lancet. 2021;398(10296):212. https://doi.org/10.1016/S0140-6736(21)01556-7

Dahl D, Onishi Y, Norwood P, et al. Effect of subcutaneous tirzepatide vs placebo added to titrated insulin glargine on glycemic control in patients with type 2 diabetes: The SURPASS-5 randomized clinical trial. JAMA. 2022;327(6):534-545.https://doi.org/10.1001/jama.2022.0078