Statistical Analysis Plan I8F-MC-GPGH (V2)

A Randomized, Phase 3, Open-Label Trial Comparing the Effect of LY3298176 versus Titrated Insulin Degludec on Glycemic Control in Patients with Type 2 Diabetes (SURPASS-3)

NCT03882970

Approval Date: 27-Jan-2021

1. Statistical Analysis Plan I8F-MC-GPGH: A Randomized, Phase 3, Open-Label Trial Comparing the Effect of LY3298176 versus Titrated Insulin Degludec on Glycemic Control in Patients with Type 2 Diabetes (SURPASS-3)

LY3298176 for Type 2 Diabetes Mellitus

Phase-3 randomized, 4-arm parallel design open label trial comparing 3-doses of LY3298176 to Insulin Degludec in patients with Type 2 Diabetes.

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Protocol I8F-MC-GPGH
Phase 3

Statistical Analysis Plan Version 1 electronically signed and approved by Lilly: 30 March 2019
Statistical Analysis Plan electronically signed and approved by Lilly on date provided below.

Approval Date: 27-Jan-2021 GMT

2. Table of Contents

Section **Page**

Table of Contents

Table **Page**

Table of Contents

Figure **Page**

3. Revision History

Statistical analysis plan (SAP) Version 1 was approved prior to the first patient receiving study drug.

The second version was approved prior to the final database lock. The following represent major changes made for the second version:

1. Update definition of analysis set: Per agreement with the US FDA, patients that discontinue study drug due to inadvertent enrollment are excluded from efficacy analysis.
2. Update baseline definition for selected measures based on scientific consideration and to minimize missing values at baseline.
3. Update primary endpoint: Missing data imputation. Corrected the definition of "retrieved dropout." Use of local lab data for glycemic control measures when central lab data are not available.
4. Update the language to handle lack of convergence in longitudinal logistic regression analysis due to low number of events for hemoglobin A1c (HbA1c) and weight loss target analyses.
5. Update the language in section on adverse events of special interest and renamed section to "Special Safety Topics" to be consistent with the program SAP (see Section 6.13.2).
6. Add section to assess SARS-CoV-2 (COVID-19) impact (see Section 6.20).
7. Add an explanation of the unblinding plan that was created for Study GPGH after the initial approval of this analysis plan.

4. Study Objectives

4.1. Primary Objective

The primary objective of the study is to demonstrate that once-weekly (QW) LY3298176 (tirzepatide) 10 mg and/or 15 mg are noninferior to insulin degludec for change from baseline in hemoglobin (HbA1c) at 52 weeks.

4.2. Key Secondary Objectives Subject to Strong Type 1 Error Rate Control

Together with the primary objective, the following secondary objectives are subjected to strong control of the type 1 error rate (see Section 6.12.3).

- To demonstrate that QW tirzepatide 5 mg is noninferior to insulin degludec for change from baseline in HbA1c at 52 weeks.
- To demonstrate that QW tirzepatide 5 mg, 10 mg, and/or 15 mg are superior to insulin degludec for change from baseline in weight at 52 weeks.
- To demonstrate that QW tirzepatide 5 mg, 10 mg, and/or 15 mg are superior to insulin degludec for change from baseline in HbA1c at 52 weeks.
- To demonstrate that QW tirzepatide 5 mg, 10 mg, and/or 15 mg are superior to insulin degludec for the proportion of patients with HbA1c target values of <7.0% (53 mmol/mol) at 52 weeks.

4.3. Other Secondary and Exploratory Objectives Not Subject to Type 1 Error Rate Control

The following secondary and exploratory objectives are not subjected to strong control of the type 1 error rate.

Secondary Objectives:

- to demonstrate that QW tirzepatide 5 mg, 10 mg, and/or 15 mg are superior to insulin degludec at 52 weeks for:
    - mean change in fasting serum glucose (central lab) from baseline
    - proportion of patients achieving an HbA1c target values of ≤6.5% (48 mmol/mol) and <5.7% (39 mmol/mol)
    - mean change in 7-point self-monitored blood glucose (SMBG) profiles from baseline
    - proportion of patients who achieved weight loss ≥5%, ≥10%, and ≥15% from baseline
- to compare tirzepatide 5 mg, 10 mg, and 15mg to insulin degludec at 52 weeks for:

- patient-reported outcomes:
  - Diabetes Treatment Satisfaction Questionnaire status/Diabetes Treatment Satisfaction Questionnaire change
  - Impact of Weight on Self-Perception (IW-SP)
  - Ability to Perform Physical Activities of Daily Living (APPADL)

Exploratory objectives:

- to compare tirzepatide 5 mg, 10 mg, and 15 mg with insulin degludec for the following:
  - change in lipids (total cholesterol, high-density lipoprotein [HDL], very low density lipoprotein [VLDL], and triglycerides [TG])
  - change from baseline in mean body mass index (BMI)
  - mean change in waist circumference
  - biomarkers
  - European Quality of Life dimension (5-level EuroQol-5D [EQ-5D]) scores

4.4. Safety Objectives

To compare the safety of tirzepatide 5 mg, 10 mg, and 15 mg to insulin degludec for:

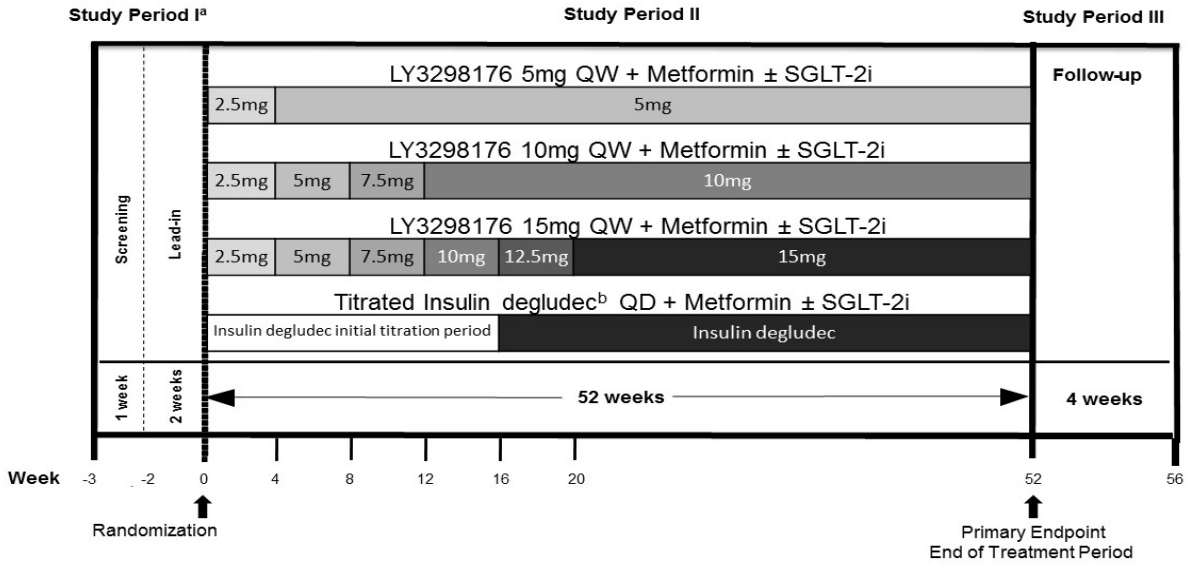
- treatment-emergent adverse events (TEAEs)
- early discontinuation of study drug due to adverse events (AEs)
- adjudicated pancreatic AEs
- serum calcitonin
- incidence of allergic and hypersensitivity reactions
- incidence of treatment-emergent anti-drug antibodies to tirzepatide
- mean change in systolic and diastolic blood pressure and heart rate from baseline
- occurrence of hypoglycemic episodes
- incidence of initiation of rescue therapy for severe, persistent hyperglycemia

5. Study Design

5.1. Summary of Study Design

Study GPGH is a Phase 3, multicenter, randomized, open-label, parallel-group study that will investigate the effects of treatment with tirzepatide 5 mg, 10 mg, and 15 mg QW compared with titrated insulin degludec in patients with type 2 diabetes mellitus (T2DM) who have inadequate glycemic control on stable doses of metformin, with or without a sodium/glucose cotransporter-2 inhibitor (SGLT-2i). The primary endpoint will be the mean change in HbA1c from baseline to 52 weeks.

Figure GPGH.5.1 illustrates the study design.



Abbreviations: FBG = fasting blood glucose; QD = once daily; QW = once weekly; SGLT-2i = sodium-glucose cotransporter-2 inhibitor; TTT = treat to target.

a Stable doses of metformin (≥1500 mg/day) ± SGLT-2i for ≥ 3 months prior to Visit 1 and during the screening/lead-in period.

b The starting dose of insulin degludec will be 10 IU/day ideally at bedtime, titrated to a FBG <90 mg/dL, following a TTT algorithm.

Figure GPGH.5.1. Illustration of study design for Clinical Protocol I8F-MC-GPGH.

Study Period I (Screening and Lead-in)

Screening (Visit 1)

The purpose of screening procedures at Visit 1 is to establish initial eligibility and to obtain blood samples for laboratory assessments needed to confirm eligibility at Visit 2. Patients who meet all applicable inclusion criteria and none of the applicable exclusion criteria at Visit 1 will continue on their prestudy therapy between Visits 1 and 2.

Lead-in (Visit 2 to Visit 3)

At Visit 2, the screening laboratory results will be reviewed and patient eligibility will be established with the exception of retinopathy status. A dilated fundoscopic exam will be performed between Visit 2 and Visit 3 as results are required to confirm eligibility.

Study Period II (52-Week Treatment Period)

Randomization (Visit 3)

At Visit 3, eligible patients will perform all required baseline study procedures (including the collection of all baseline laboratory measures) prior to randomization and prior to taking the first dose of study drug.

Postrandomization period (end of Visit 3 to Visit 24)

The starting dose of tirzepatide will be 2.5 mg QW for 4 weeks, followed by an increase to 5 mg QW, for the duration of the study in the 5 mg group. For the 10 mg group, the starting dose of tirzepatide will be 2.5 mg QW for 4 weeks, then the dose will be increased by 2.5 mg every 4 weeks (2.5 to 5 to 7.5 to 10 mg) until the 10 mg dose is reached and maintained for the duration of the study. For the 15 mg group, the starting dose of tirzepatide will be 2.5 mg QW for 4 weeks, then the dose will be increased by 2.5 mg every 4 weeks (2.5 to 5 to 7.5 to 10 to 12.5 to 15 mg) until the 15 mg dose is reached and maintained for the duration of the study.

The initial dose of insulin degludec will be 10 IU/day ideally at bedtime, titrated to a fasting blood glucose (FBG) <90 mg/dL, following a treat-to-target (TTT) algorithm.

Study Period III (Safety Follow-up Period)

Safety follow-up (Visit 801) visits

All patients who complete the treatment period are required to complete Visit 801, a safety follow-up visit, approximately 4 weeks after their last visit. Patients discontinuing the study early and performing an early termination (ET) visit will also be asked to perform the safety follow-up visit, so that the safety follow-up visit will be their final visit. During the safety follow-up period, patients will not receive study drug. Patients will be treated with another glucose-lowering intervention decided upon by the investigator. Initiation of new antihyperglycemic therapy for the safety follow-up period will not be classified as "rescue therapy."

5.2. Determination of Sample Size

Approximately 1420 patients (355 per group) may be randomly assigned in a 1:1:1:1 ratio to tirzepatide 5, 10, 15 mg QW, or insulin degludec. Patient randomization will be stratified based on country, baseline concomitant oral medication (metformin alone or metformin plus SGLT-2i), and baseline HbA1c (≤8.5% or >8.5% [69 mmol/mol]).

Although the primary objective of the trial is to demonstrate that once-weekly 10 and/or 15 mg tirzepatide doses are noninferior to titrated insulin degludec relative to mean change in HbA1c

from baseline (using a 0.3% noninferiority boundary), the study is powered to assess superiority of tirzepatide 10 and/or tirzepatide 15 mg QW to insulin degludec, relative to the primary endpoint: mean change in HbA1c from baseline to 52 weeks.

The power is assessed based on the following assumptions:

- each of the 10 and 15 mg tirzepatide QW treatment groups will be tested in parallel against insulin degludec at a 2-sided 0.025 significance level.
- use of a 2-sample t test utilizing HbA1c data collected before initiation of any rescue medication or premature treatment discontinuation with no more than 28% of patients initiating any rescue medication or prematurely discontinuing treatment in each treatment group
- a 0.35% greater mean reduction in HbA1c from baseline for 10 and 15 mg of tirzepatide compared with insulin degludec
- a superiority margin of 0.05%
- a common standard deviation of 1.1%

On the basis of these assumptions, randomly assigning approximately 1420 patients to the 4 treatments using a 1:1:1:1 ratio is required to ensure at least 90% power to demonstrate superiority of tirzepatide 10 mg and/or 15 mg QW doses to insulin degludec, relative to the primary endpoint for the "efficacy" estimand. Furthermore, this sample size will ensure 90% power to demonstrate superiority for the "treatment-regimen" estimand conducted using an analysis of covariance (ANCOVA) utilizing all available HbA1c data at 52 weeks, with missing data imputed with a conservative multiple imputation method (as described in Section 6.12.1.3) if the SD were to increase up to 1.3% due to the inclusion of data on rescue medications, inclusion of data after premature treatment discontinuation, and imputation of the missing data.

5.3. Method of Assignment to Treatment

Approximately 1420 patients who meet all criteria for enrollment will be randomized to 1 of the study treatment arms at Visit 3. Assignment to treatment arms will be determined by a computer-generated random sequence using an interactive web response system (IWRS). Patients will be randomized in a 1:1:1:1 ratio to receive 5 mg tirzepatide, 10 mg tirzepatide, 15 mg tirzepatide, or insulin degludec titrated to achieve a FPG <90 mg/dL. The randomization will be stratified by country, baseline HbA1c concentration (≤8.5%, >8.5% [≤69, >69 mmol/mol]), and use of concomitant oral antidiabetic medications (OAMs) (metformin alone, metformin plus an SGLT-2i).

6. A Priori Statistical Methods

6.1. Populations for Analyses

For purposes of analyses, Table GPGH.6.1 defines the following analysis sets:

Table GPGH.6.1. Description of Analysis Datasets

Analysis Set	Description
Screened population	All participants who sign informed consent
Randomized population	All patients who are randomly assigned a treatment arm
modified intention-to-treat (mITT) population	All randomly assigned participants who are exposed to at least 1 dose of study drug. In the event of a treatment error, participants will be analyzed according to the treatment they were randomized.
Efficacy analysis set (EAS)	Data obtained during Study Period II from the mITT population, excluding patients discontinuing study drug due to inadvertent enrollment and data after initiating rescue antihyperglycemic medication or stopping study drug (last dose date +7 days).
Full analysis set (FAS)	All available data obtained during Study Period II from mITT, excluding patients discontinuing study drug due to inadvertent enrollment, regardless of adherence to study drug or initiation of rescue antihyperglycemic medication.
Safety analysis set (SS)	Data obtained during Study Periods II or III from the mITT population, regardless of adherence to study drug or initiation of rescue antihyperglycemic medication.

6.2. General Considerations

Statistical analysis of this study will be the responsibility of Eli Lilly and Company (Lilly) or its designee. All statistical analyses will be conducted with SAS Version 9.4 or higher unless otherwise stated. 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 changes to the data analysis methods described in the protocol, and the justification for making the change, will be described in the SAP or the clinical study report (CSR). Some analyses and summaries described in this analysis plan may not be conducted if not warranted by data (e.g., few events to justify conducting an analysis). Listings of events will be provided in such situations. Additional analyses of the data may be conducted as deemed appropriate without further changes made to the protocol or SAP, even after the primary or final database locks (DBL).

Patients inadvertently receiving incorrect study treatment are expected to switch to their randomized treatment arm as soon as possible. Patients assigned to tirzepatide may not be able to tolerate the maximum dose of the randomized treatment arm and may continue study participation on a reduced maintenance dose. Continuing on a reduced maintenance dose will neither be considered as discontinuation of randomized treatment nor will be considered as non-compliant to randomized treatment. Additionally, to avoid potential selection biases, unless stated otherwise, statistical summaries and analyses will be conducted based on randomized maintenance dose regardless of the actual treatment received by the patient.

Unless specified otherwise, the last measurement during Visit 1 to Visit 3 (including unscheduled visits) collected prior to or on the first dose day will serve as baseline. For immunogenicity, data collected up to the first dose time will serve as baseline. For lab and electrocardiogram (ECG), baseline needs to be prior to or within 1 hour after the first dose time. For patient-reported outcomes, data collected at Visit 3, regardless of the timing relative to first dose, will serve as baseline.

There will be 2 estimands of interest in evaluating primary and secondary efficacy objectives. The first estimand, the "efficacy" estimand, represents efficacy prior to discontinuation of study drug without confounding effects of rescue therapy for persistent severe hyperglycemia. Analysis relative to the "efficacy" estimand will be conducted using the efficacy analysis set (EAS). The second estimand, the "treatment-regimen" estimand (including all data), represents the efficacy irrespective of adherence to investigational product or introduction of rescue therapy for persistent severe hyperglycemia. Analyses relative to the "treatment-regimen" estimand will be conducted using the full analysis set (FAS).

Unless specified otherwise, safety analyses will be conducted relative to the "treatment-regimen" estimand using the safety analysis set (SS).

End of study participation for a patient will be the earliest of date of death, date of withdrawal from further participation in the study, or date of safety follow-up visit (Visit 801). For patients considered to be lost-to-follow-up, end of study participation will be the date of lost-to-follow-up reported by the investigator. Patient data included in the database after the last date of study participation (date of early termination or date of safety follow-up) will be excluded from statistical analysis. Listings of such data will be provided.

Summary statistics for categorical measures (including categorized continuous measures) will include sample size, frequency, and percentages. Summary statistics for continuous measures will include sample size, mean, standard deviation (SD), median, minimum, and maximum. The summary statistics will be presented by nominal visit.

Statistical treatment comparisons will only be performed between tirzepatide doses and insulin degludec. Since the trial is not adequately powered to detect difference among tirzepatide doses, comparisons among tirzepatide arms will not be performed unless otherwise specified.

Statistical summaries and results of statistical analyses will be displayed in the following treatment order: 5 mg tirzepatide, 10 mg tirzepatide, 15 mg tirzepatide, insulin degludec.

6.3. Adjustments for Covariates

The study is stratified by country, baseline HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]), and use of concomitant oral antidiabetic medications (metformin alone, metformin plus an SGLT-2i). Where necessary to be included as a stratification factor, countries with fewer than 10 randomized patients will be pooled into 1 category (pooled country). For HbA1c related analyses, country/pooled country and concomitant oral antidiabetic medication use will be used as stratification factors and baseline HbA1c as a covariate. For other efficacy analyses, country/pooled country, baseline HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]), and use of

concomitant oral antidiabetic medications (metformin alone, metformin plus an SGLT-2i) will be used as stratification factors and the respective baseline value as a covariate. Stratification factors will be based on the information collected in the clinical database.

6.4. Handling of Dropouts or Missing Data

For the primary and key secondary efficacy endpoint analyses subject to type 1 error rate control, data for patients with missing values at the 52-week visit will be imputed based on the method described in Section 6.12.1.3. Unless specified otherwise, imputation of missing data will be limited to primary and key secondary efficacy endpoint analyses. Missing other secondary or exploratory efficacy parameter values and missing safety laboratory values will not be explicitly imputed.

6.5. Multicenter Studies

To investigate potential regional influence on efficacy, country/pooled country will be used as a stratification factor in primary and secondary efficacy analysis.

6.6. Multiple Comparisons/Multiplicity

The type 1 error rate control strategy for primary and key secondary efficacy objectives is illustrated in Section 6.12.3. As they are intended for different purposes, no multiplicity adjustments will be made for conducting separate analyses relative to the "efficacy" and "treatment-regimen" estimands. In addition, no multiplicity adjustments will be made for evaluating other secondary and exploratory efficacy objectives and safety assessments.

6.7. Patient Disposition

Reasons for screen failure as reported by investigators will be summarized.

A listing of final study disposition and a listing of randomized treatment assignment (planned treatment) for all randomized patients will be provided. Final study disposition and study treatment disposition for all randomized patients will be summarized by planned study treatment.

6.8. Patient Characteristics

A listing of patient demographics will be provided for all randomized patients. All demographic and baseline clinical characteristics will be summarized by study treatment for all randomized patients. Baseline demographic and clinical characteristics of special interest including but not limited to: age, gender, race, ethnicity, BMI, weight, country of enrollment, HbA1c, fasting serum glucose, duration of T2DM, baseline antihyperglycemic medication, cardiovascular (CV) history, renal function, history of retinopathy, results of the fundoscopic exam, and history of gallbladder disease.

6.9. Concomitant Therapy

Prespecified concomitant medications of interest will be summarized by treatment at randomization. Additionally, medications of interest initiated after randomization will be summarized. Concomitant therapies will be mapped using the World Health Organization

(WHO) Drug Dictionary in the clinical trial database and will be further classified using Anatomic-Therapeutic-Chemical (ATC) codes for reporting purposes.

Concomitant medications of interest include the following groups of medication:

- baseline antihyperglycemic therapy
    - prior glucagon-like peptide-1 receptor agonist (GLP-1 RA) use by type
    - metformin
    - sodium/glucose cotransporter-2 -inhibitors, by type
- baseline antihypertensive therapy, by type
- baseline lipid lowering therapy, by type
- changes to baseline medication in Study Period II:
    - antihypertensive therapy
    - lipid lowering therapy
- changes to baseline medication during Study Periods II and III:
    - antihyperglycemic therapy
- rescue therapy due to severe persistent hyperglycemia
- utilization of:
    - antidiarrheal medication
    - antiemetic medication

6.10. Treatment Exposure and Compliance

A listing of patients randomized but not receiving study treatment will be provided, if applicable. The listing will include patient identification, randomized treatment arm, and the reason for not receiving study treatment.

A listing of randomized patients who inadvertently received incorrect study treatment anytime during the study will be provided, if applicable. The listing will include patient identification, randomized treatment arm, and information related to the treatment incorrectly received: incorrectly received study treatment, start and stop dates during which the incorrect treatment was received.

Summary of duration of follow-up (defined as time in days from date of randomization to date of the safety follow-up visit) and duration on study treatment (defined as time in days from date of first dose of study treatment to date of last dose of study treatment plus 7 days) will be provided by therapy.

6.10.1. Exposure and Compliance to Tirzepatide

The number of patients prematurely discontinuing study treatment prior to the 52-week visit will be provided by study treatment. Reasons for prematurely discontinuing study treatment prior to

the 52-week visit will also be provided by study treatment. A time-to-event analysis of premature study treatment discontinuation will be conducted.

The proportion of patients with missing dosing information, receiving no LY dose, receiving 2.5 mg, 5 mg, 7.5 mg, 10 mg, 12.5 mg, or 15 mg will be presented by randomized treatment and week from first dose.

A listing and summary of patients continuing on a reduced maintenance dose of tirzepatide compared to the randomized dose may be provided.

Compliance will be defined as taking at least 75% of the scheduled tirzepatide doses. Compliance will be calculated by taking the number of doses administered (regardless of the actual dose administered) divided by the total number of doses expected to be administered ×100. Treatment compliance will be summarized descriptively over the entire study period by treatment using the mITT population.

6.10.2. Exposure and Compliance to Insulin Degludec Treat-to-Target Algorithm

Summary information on total daily dose of insulin degludec will be reported by visit. Information related to compliance to treat-to-target therapy including reasons for non-compliance will be summarized.

Compliance will be defined as taking at least 75% of the scheduled insulin doses. Compliance will be calculated by taking the number of doses administered (regardless of the actual dose administered) divided by the total number of doses expected to be administered × 100. Treatment compliance will be summarized descriptively over the entire study period by treatment using the mITT population

Additionally, summaries will be produced for the percent of patients to reach fasting serum glucose goals of <90 mg/dL (5 mmol/L) and <126 mg/dL (7 mmol/L) over time at weeks 16, 24, 40, and 52.

6.11. Important Protocol Deviations

Important protocol deviations are identified in the Trial Issues Management Plan (TIMP). A listing and a summary of important protocol deviations by treatment will be provided.

6.12. Efficacy Analyses

For the Food and Drug Administration (FDA) and potentially for other regulatory agencies, all efficacy assessments will be guided by the "treatment-regimen" estimand conducted using the FAS. Assessment of the primary and secondary efficacy objectives subject to type 1 error rate control (key secondary) will be conducted with multiple imputation of missing data (see Section 6.12.1.3) at 52 weeks. Assessment of other efficacy objectives will be conducted without imputation of missing data. For publications and other purposes, the assessment of efficacy objectives will be guided by the "efficacy" estimand using the EAS without imputation of missing data.

6.12.1. Primary Efficacy Analysis

The primary efficacy measure will be change in HbA1c from baseline (postbaseline – baseline). Both HbA1c values as well as change from baseline in HbA1c will be summarized by treatment and nominal visit (week). When applicable, HbA1c data from a local lab will be used when central lab data are not available. If scheduled HbA1c data at the primary endpoint visit are not available, unscheduled HbA1c data collected for the primary endpoint visit will be included in the analysis.

6.12.1.1. The Analysis Relative to the Efficacy Estimand

The analysis will be conducted utilizing HbA1c data in the EAS from baseline through the 52-week visit with the aid of a mixed model for repeated measures (MMRM). Restricted maximum likelihood (REML) will be used to obtain model parameter estimates and the Kenward-Roger option will be used to estimate the denominator degrees of freedom. The response variable of the MMRM model will be the primary measure and model terms will include treatment, visit, treatment-by-visit interaction, country/pooled country, and baseline concomitant oral antihyperglycemic medication use (metformin alone versus metformin plus an SGLT-2i) as fixed effects and baseline HbA1c as a covariate. An unstructured covariance structure will be used to model the within-patient errors. If this model fails to converge, the following covariance structures will be tested in order:

- heterogeneous Toeplitz
- heterogeneous first order autoregressive
- heterogeneous compound symmetry
- Toeplitz
- first order autoregressive
- compound symmetry

The first covariance structure that converges will be used. The resulting least squares mean (LSM) estimates of mean change from baseline in HbA1c will be plotted by visit and by study treatment.

With the aid of the MMRM analysis, 2-sided 97.5% confidence intervals (CI) for mean change in HbA1c from baseline to the 52-week visit for (10 mg tirzepatide – insulin degludec) as well as for (15 mg tirzepatide – insulin degludec) will be derived. If the upper limit of the CI is ≤0.3%, then the respective dose of tirzepatide (10 mg and/or 15 mg) will be declared noninferior to insulin degludec relative to change in HbA1c from baseline.

6.12.1.2. The Analysis Relative to the Treatment Regimen Estimand

The analysis will be conducted utilizing HbA1c data in the FAS at baseline and at the 52-week visit with the aid of an ANCOVA. The response variable will be the primary measure and model terms will include treatment, country/pooled country, and baseline concomitant oral antihyperglycemic medication use (metformin alone versus metformin plus an SGLT-2i) as fixed

effects and baseline HbA1c as a covariate. The ANCOVA analysis will be conducted with multiple imputation of missing primary measures (see Section 6.12.1.3 for details) and statistical inference over multiple imputation of missing data guided by Rubin (1987).

With the aid of the ANCOVA analysis 2-sided 97.5% CI for mean change in HbA1c from baseline to the 52-week visit for (10 mg tirzepatide – insulin degludec) as well as for (15 mg tirzepatide – insulin degludec) will be derived. If the upper limits of the CI are ≤0.3%, then the respective dose of tirzepatide (10 mg and/or 15 mg) will be declared noninferior to insulin degludec relative to change in HbA1c from baseline.

6.12.1.3. Methods for Multiple Imputations

For efficacy analyses relative to the "treatment-regimen" estimand, missing HbA1c data at the 52-week visit will be imputed based on "retrieved dropouts," defined as patients who had their HbA1c value measured at the 52-week visit in the same treatment arm who prematurely discontinued study drug. If the imputed value of HbA1c change from baseline is <-6.0% or >6.0%, then that value will be set to -6.0% or 6.0%, respectively, to avoid unrealistic imputed values.

6.12.1.4. Additional Analyses of the Primary Outcome

Upon successfully establishing noninferiority of tirzepatide compared to insulin degludec, superiority of tirzepatide compared to insulin degludec relative to change in HbA1c from baseline will be evaluated (see Section 6.12.3).

6.12.2. Secondary Efficacy Analyses Subject to Type 1 Error Rate Control

6.12.2.1. Mean Change in HbA1c from Baseline at the 52 Week Visit

Noninferiority of 5 mg LY3198176 to insulin degludec will be conducted in a manner similar to Section 6.12.1. Assessment of Superiority of tirzepatide doses compared to insulin degludec will be conducted using the same statistical models as those used for evaluating the primary objective in Section 6.12.1. Decisions will be guided by the 2-sided p-values for mean comparisons between tirzepatide doses and insulin degludec (see details in Section 6.12.3).

6.12.2.2. Mean Change in Body Weight from Baseline at the 52 Week Visit

The analysis for change in body weight from baseline (postbaseline – baseline) will be conducted in a manner similar to the primary analysis in Section 6.12.1. Baseline HbA1c concentration (≤8.5%, >8.5% [≤69, >69 mmol/mol]) will be used as a fixed factor in place of baseline HbA1c as a covariate and baseline body weight will be used as an additional covariate in the statistical model. Least squares mean estimates of mean change in body weight from baseline will be plotted by nominal visit and by study treatment. For multiple imputation of missing values, if the imputed value of weight change from baseline is <-50 kg or >50 kg, then that value will be set to -50 kg or 50 kg, respectively, to avoid unrealistic imputed values.

6.12.2.3. Proportion of Patients Achieving HbA1c <7% at the 52 Week Visit

Analyses relative to the "efficacy" estimand for the endpoint at 52 weeks will be performed using the EAS with missing values imputed from an MMRM model and then dichotomized. The MMRM model includes treatment, country/pooled country, baseline concomitant oral antihyperglycemic medication use, visit, and treatment-by-visit interaction as fixed effects, and baseline HbA1c as a covariate. After dichotomizing continuous HbA1c, the data are analyzed using a logistic regression model with treatment, country/pooled country, and baseline concomitant oral antihyperglycemic medication use as fixed effects, and baseline HbA1c as a covariate. In addition, an analysis will be conducted utilizing data using the EAS from baseline through the 52-week visit with the aid of a longitudinal logistic regression with repeated measurements with country/pooled country, baseline concomitant oral antihyperglycemic medication use, treatment, visit, and treatment-by-visit interaction as fixed effects, and baseline HbA1c as a covariate. If the longitudinal logistic model does not converge due to a small number of events, logistic regression will be utilized to analyze the proportion of patients achieving HbA1c <7% at nominal visits.

Analyses relative to the "treatment-regimen" estimand will be conducted utilizing HbA1c data in the FAS at baseline and at the 52-week visit with the aid of a logistic regression with multiple imputation of missing HbA1c data at the 52-week visit (see Section 6.12.1.3 for details). Model terms will include treatment, country/pooled country, and baseline concomitant oral antihyperglycemic medication use as fixed effects, and baseline HbA1c as a covariate. Statistical inference over multiple imputations will be guided by Rubin (1987).

6.12.3. Type 1 Error Rate Control Strategy for Primary and Key Secondary Efficacy Analyses

Since they are intended for different purposes, no type 1 error rate adjustments will be made for conducting analyses relative to "efficacy" and "treatment-regimen" estimands. For analyses within each estimand, the type 1 error control strategy for evaluation of the primary and key secondary objectives is illustrated in Figure GPGH.6.1.

1. $H_{15,1}$, $H_{15,2}$, and $H_{15,3}$ are evaluated hierarchically each at a 2-sided 0.025 significance level conditioned on successfully achieving the preceding objective. In parallel,
2. $H_{10,1}$, $H_{10,2}$, and $H_{10,3}$ are evaluated hierarchically each at a 2-sided 0.025 significance level conditioned on successfully achieving the preceding objective.
3.
    a. If all objectives in #1 and #2 above are successfully established, $H_{5,1}$, $H_{5,2}$, and $H_{5,3}$ are evaluated hierarchically, each at a 2-sided 0.05 significance level.
    b. If all objectives in only #1 or only #2 above are successfully established, $H_{5,1}$, $H_{5,2}$, and $H_{5,3}$ are evaluated hierarchically, each at a 2-sided 0.025 significance level.
4. If all objectives: $H_{5,1}$, $H_{5,2}$, and $H_{5,3}$ are successfully established and

a. If all objectives in #1 and #2 above are successfully established, then $H_{10,4}$, $H_{15,4}$, and $H_{5,4}$ will be evaluated hierarchically each at a 2-sided 0.05 significance level conditioned on successfully achieving the preceding objective.

b. If all objectives in only #1 or only #2 above are successfully established, then $H_{10,4}$, $H_{15,4}$, and $H_{5,4}$ will be evaluated hierarchically each at a 2-sided 0.025 significance level conditioned on successfully achieving the preceding objective.

5. If all objectives in #3 and #4 above are successfully established, and at least 1 objective from #1 or #2 above is not successfully established, recycle 100% of the unused alpha back to #1 or #2 above.

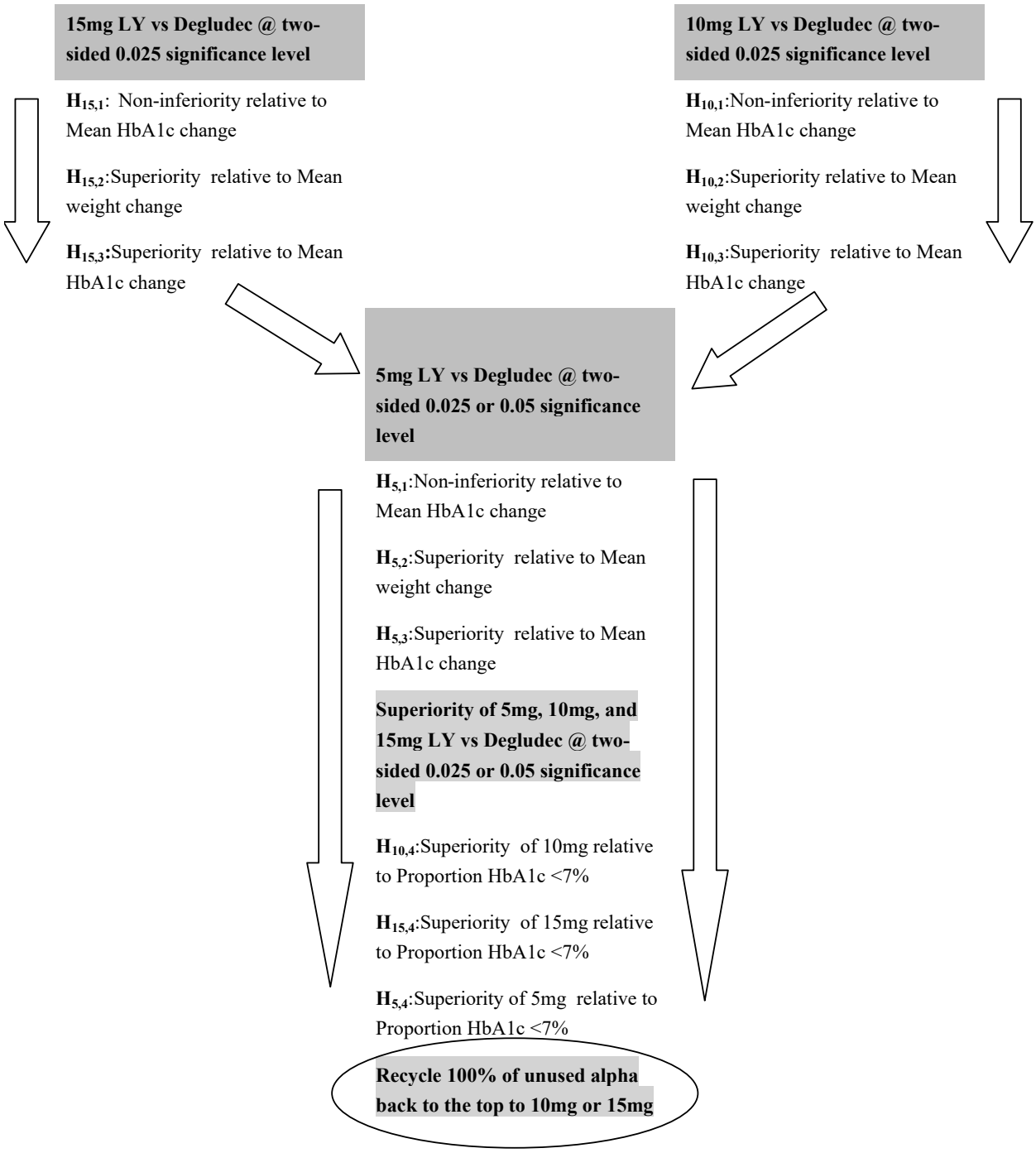


Figure GPGH.6.1. Type 1 error control strategy for primary and key secondary efficacy endpoints.

6.12.4. Other Secondary and Exploratory Efficacy Analyses

Other secondary and exploratory efficacy measures will be summarized by treatment and nominal visit. Statistical analyses will be conducted in a manner similar to Sections 6.12.1 and 6.12.2, however missing data will not be imputed and assessments are not subject to type 1 error rate control.

Table GPGH.6.2. Secondary and Exploratory Efficacy Measures Not Controlled for Type 1 Error

Objective	Relative to the efficacy measure:	Analysis Conducted in a manner similar to:	Additional Information
Secondary Analyses			
QW tirzepatide 5 mg, 10 mg, and/or 15 mg is superior to insulin degludec at 52 weeks	Change from baseline in fasting serum glucose (central lab)	6.12.1	Use baseline HbA1c strata as a fixed effect in place of baseline HbA1c as a covariate. LSM estimates will be plotted by treatment and visit.
	Proportion of patients achieving an HbA1c target value of ≤6.5% (48 mmol/mol)	6.12.2.3	None
	Proportion of patients achieving an HbA1c target value of <5.7% (39 mmol/mol).	6.12.2.3	None
	Change from baseline in 7-point SMBG profiles	6.12.1	Use baseline HbA1c strata as a fixed effect in place of baseline HbA1c as a covariate. Use baseline SMBG as a covariate. LSM estimates at 52-weeks will be plotted by treatment and 7-points. In addition to the analyses on each of the 7-points, similar analyses will be done for the 2-hour morning, midday, and evening meal excursions, the mean of all meals 2-hour excursion, the mean of all 7-point measurements, the mean of all pre-meal measurements, and the mean of all 2-hour postprandial measurements.
	Proportion of patients achieving weight loss from baseline of ≥5%.	6.12.2.3	Use baseline HbA1c strata as a fixed effect in place of baseline HbA1c as a covariate. Use baseline weight as a covariate.
	Proportion of patients achieving weight loss from baseline of ≥10%.	6.12.2.3	Use baseline HbA1c strata as a fixed effect in place of baseline HbA1c as a covariate. Use baseline weight as a covariate.

Secondary and Exploratory Efficacy Measures Not Controlled for Type 1 Error

Objective	Relative to the efficacy measure:	Analysis Conducted in a manner similar to:	Additional Information
	Proportion of patients achieving weight loss from baseline of ≥15%.	6.12.2.3	Use baseline HbA1c strata as a fixed effect in place of baseline HbA1c as a covariate. Use baseline weight as a covariate.
To compare tirzepatide 5 mg, 10 mg, and 15 mg to insulin degludec at 52 weeks	The various Diabetes Treatment Satisfaction Questionnaire change (DTSQc) scores (see Section 6.17.4 for more details).	6.12.1.2	Use baseline HbA1c strata as a fixed effect in place of baseline HbA1c as a covariate. Use baseline DTSQs score as a covariate.
	Changes from baseline in the Impact of Weight on Self-Perceptions Questionnaire (IW-SP) scores.	6.12.1.2	Use baseline HbA1c strata as a fixed effect in place of baseline HbA1c as a covariate. Use baseline IW-SP score as a covariate.
	Changes from baseline in the Ability to Perform Physical Activities of Daily Living (APPADL) scores.	6.12.1.2	Use baseline HbA1c strata as a fixed effect in place of baseline HbA1c as a covariate. Use baseline APPADL score as a covariate.
Exploratory Analyses			
To compare tirzepatide 5 mg, 10 mg, and 15 mg with insulin degludec for the following:	Change from baseline in lipid parameters (total cholesterol, HDL, VLDL, TG).	6.12.1	Use baseline HbA1c strata as a fixed effect in place of baseline HbA1c as a covariate. Use corresponding baseline lipid parameter as a covariate.
	Change from baseline in BMI.	6.12.1	Use baseline HbA1c strata as a fixed effect in place of baseline HbA1c as a covariate. Use corresponding baseline BMI as a covariate.
	Change from baseline in waist circumference.	6.12.1	Use baseline HbA1c strata as a fixed effect in place of baseline HbA1c as a covariate. Use corresponding baseline waist circumference as a covariate.
	Changes from baseline in the European Quality of Life – 5 dimensions (EQ-5D) index and visual analog scale (VAS) scores.	6.12.1.2	Use baseline HbA1c strata as a fixed effect in place of baseline HbA1c as a covariate. Use baseline EQ-5D score as a covariate.
To compare tirzepatide 5 mg, 10 mg, and 15 mg with insulin degludec for the following:	Proportion of patients achieving an HbA1c target <6.5%, without weight gain (<0.1 kg), and without documented symptomatic hypoglycemia or severe hypoglycemia	6.12.2.3	Include the rate of hypoglycemic events at baseline and baseline body weight as additional covariates

Secondary and Exploratory Efficacy Measures Not Controlled for Type 1 Error

Objective	Relative to the efficacy measure:	Analysis Conducted in a manner similar to:	Additional Information
	Proportion of patients achieving an HbA1c target <7.0%, without weight gain (<0.1 kg), and without documented symptomatic hypoglycemia or severe hypoglycemia	6.12.2.3	Include the rate of hypoglycemic events at baseline and baseline body weight as additional covariates
	Proportion of patients achieving an HbA1c target <6.5%, without weight gain (<0.1 kg), and without clinically significant documented symptomatic hypoglycemia or severe hypoglycemia	6.12.2.3	Include the rate of hypoglycemic events at baseline and baseline body weight as additional covariates
	Proportion of patients achieving an HbA1c target <7.0%, without weight gain (<0.1 kg), and without clinically significant documented symptomatic hypoglycemia or severe hypoglycemia	6.12.2.3	Include the rate of hypoglycemic events at baseline and baseline body weight as additional covariates

Abbreviations: HbA1c = hemoglobin A1c; QW = once-weekly.

6.13. Safety Analyses

Unless specified otherwise, safety assessments will be based on the SS (Table GPGH.6.1). All events that occur between the date of first dose of study drug to the date of patient's safety follow-up visit or patient's end of study participation will be included, regardless of the adherence to study drug or initiation of rescue therapy. For assessing the benefit and risk profile through 52 weeks, selected safety analyses will be conducted by utilizing safety data from first dose through the date of the 52-week visit. Some safety analyses may be conducted after excluding data after the initiation of new antihyperglycemic therapy.

The statistical assessment of homogeneity of the distribution of categorical safety responses among treatment arms will be conducted using Fisher's exact test, unless specified otherwise.

The mean change from baseline differences among treatments at all scheduled visits will be assessed via an MMRM using REML, unless specified otherwise. The model will include country/pooled country, baseline HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]), baseline concomitant oral antihyperglycemic medication, treatment group, visit and treatment-by-visit

interaction as fixed effects, and baseline value of the safety parameter as a covariate. To model the covariance structure within patients, the unstructured covariance matrix will be used. If this model fails to converge, the covariance structures specified in Section6.12.1.1 will be tested in order.

For selected safety parameters, time-to-first-event analysis via the cox-proportional hazards model may be conducted. Patients without the event will be censored at the end of study participation. For patients experiencing the event, the "time-to-first-event" will be the time (in days) from first dose to first occurrence of the event.

Where necessary, the rate of events will be analyzed using a generalized linear mixed-effects model assuming the number of events follow a negative binomial distribution and with treatment as a fixed effect. The logarithm of days during the active treatment period will be adjusted as an offset to account for possible unequal treatment duration of follow-up between patients.

6.13.1. Adverse Events

A listing of AEs occurring either before first dose or after the patient's last date of study participation will be provided. The listing will include patient identification including the treatment, site number, event information: AE group ID, event start date, Medical Dictionary for Regulatory Activities (MedDRA) System Organ Class (SOC), and Preferred Term (PT), seriousness, severity, outcome, relationship to study drug, time from first dose of study drug to the event, time from last dose of study drug to the event, and time from end of study participation to the event.

A TEAE is defined as an event that first occurred or worsened in severity after baseline. The MedDRA Lowest Level Term (LLT) will be used in the treatment-emergent derivation. The maximum severity for each LLT during the baseline period including ongoing medical history will be used as baseline severity. For events with a missing severity during the baseline period, it will be treated as "mild" in severity for determining treatment-emergence. Events with a missing severity during the postbaseline period will be treated as "severe" and treatment-emergence will be determined by comparing to baseline severity.

The percentages of patients with TEAEs will be summarized by treatment using MedDRA PT nested within SOC. Statistical comparisons will be applied at both the SOC and PT levels. Events will be ordered by decreasing frequency within SOC. For events that are sex-specific, the denominator and computation of the percentage will include only patients from the given sex.

An overview of the number and percentage of patients who experienced a TEAE, serious adverse event (SAE), died due to an AE, discontinued from study treatment or study due to an AE, relationship to study drug, and outcome of the AE will be summarized by treatment.

The percentages of patients with TEAEs, overall and common (common TEAEs occurred in ≥5% of patients before rounding), will be summarized by treatment using MedDRA PT. Events will be ordered by decreasing frequency.

The percentages of patients with TEAEs by maximum severity will be summarized by treatment using MedDRA PT. For each patient and TEAE, the maximum severity for the MedDRA PT is the maximum postbaseline severity observed from all associated LLTs mapping to the MedDRA PT. The maximum severity will be determined based on the non-missing severities. If all severities are missing for the defined postbaseline period of interest, it will show as missing in the table. Only counts and percentages will be included for the TEAEs by maximum security.

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

- death
- serious adverse event
- pregnancy
- permanent discontinuation of study treatment due to AE
- severe adverse events of special interest

6.13.1.1. Deaths

A listing of all deaths will be provided. The listing will include patient identification including the treatment, site number, date of death, age at the time of enrollment, sex, 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), investigator reported cause of death, and CEC adjudicated cause of death.

6.13.1.2. Other Serious Adverse Events

The number and percentage of patients who experienced an SAE (including deaths and SAEs temporally associated or preceding deaths) during the study follow-up will be summarized by treatment using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency within SOC.

A listing of all SAEs will be provided. The listing will include, but not be limited to, treatment, patient identification including the site number, date of event, age at the time of enrollment, sex, MedDRA SOC and PT, severity, action taken, outcome, relationship to study drug, time from first dose of study drug to the event, and event duration.

6.13.1.3. Discontinuation from Study Due to Adverse Event

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

6.13.1.4. Discontinuation from Study Drug Due to Adverse Event

The number and percentage of patients who prematurely discontinue study drug due to an AE will be summarized by treatment using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency within SOC. Time-to-event analyses will be conducted by treatment on time to study drug discontinuation as well as on time to study drug discontinuation due to an AE.

6.13.2. Special Safety Topics

6.13.2.1. Hypoglycemic Events

Definitions of different categories of hypoglycemic events are included in Table GPGH.6.3.

Table GPGH.6.3. Definitions of Hypoglycemic Event Categories

	Symptoms and/or Signs of Hypoglycemia	Blood Glucose Level
Documented symptomatic hypoglycemia	Yes	≤70 mg/dL (3.9 mmol/L)
Documented asymptomatic hypoglycemia	No	≤70 mg/dL (3.9 mmol/L)
Documented unspecified hypoglycemia	Unknown	≤70 mg/dL (3.9 mmol/L)
Clinically significant documented symptomatic hypoglycemia	Yes	<54 mg/dL (3.0 mmol/L)
Clinically significant documented asymptomatic hypoglycemia	No	<54 mg/dL (3.0 mmol/L)
Clinically significant documented unspecified hypoglycemia	Unknown	<54 mg/dL (3.0 mmol/L)

Severe hypoglycemia: Defined as an episode with severe cognitive impairment requiring the assistance of another person to actively administer carbohydrate, glucagon, or other resuscitative actions. Severe hypoglycemia will be reported as an SAE. Severe hypoglycemia will be considered an adverse event of special interest (AESI).

Nocturnal hypoglycemia: Defined as any hypoglycemic event that occurs between bedtime and waking.

To avoid duplicate reporting, all consecutive BG values ≤70 mg/dL (3.9 mmol/L) occurring within a 1-hour period may be considered to be a single hypoglycemic event.

Statistical summaries and analyses will exclude hypoglycemic events occurring while on rescue therapy. For each of the aforementioned hypoglycemia categories, incidence as well as rate per patient year of exposure will be provided by treatment. A listing of hypoglycemic events occurring while on rescue therapy will also be provided.

For each of the hypoglycemia categories mentioned above, the incidence of any hypoglycemic event (with blood glucose ≤70 mg/dL [3.9 mmol/L] or <54 mg/dL [3.0mmol/L]) will be analyzed using logistic regression with treatment, country/pooled country, and baseline concomitant oral antihyperglycemic medication as fixed effects and baseline HbA1c and rate of hypoglycemic events at baseline as covariates. The rate of hypoglycemic episodes will be analyzed using a generalized linear mixed-effects model assuming the number of hypoglycemic episodes follows a negative binomial distribution with the mean modeled using country/pooled country, baseline concomitant oral antihyperglycemic medication and treatment as a fixed effects and baseline HbA1c as a covariate. The logarithm of days during the active treatment period will be adjusted as an offset to account for possible unequal treatment duration of follow-up between patients.

6.13.2.2. Severe Persistent Hyperglycemia

A summary of initiation of rescue therapy in response to severe, persistent hyperglycemia will be provided by treatment. If there are a sufficient number of episodes, time-to-first-event analyses for the initiation of rescue therapy will be conducted by treatment using a cox proportional regression model. A listing of patients who initiated rescue therapy will be provided.

6.13.2.3. Pancreatitis

Summaries of adjudicated and investigator-reported pancreatic events will be provided by treatment. Determination of investigator-reported events will be through the predefined Standardized MedDRA Queries (SMQ) search for acute pancreatitis and MedDRA PT of pancreatitis chronic. Detailed searching criteria can be found in Appendix 1. Treatment-emergent adjudication-confirmed pancreatitis will be considered an AESI.

6.13.2.3.1. Pancreatic Enzyme Assessment

Observed pancreatic enzyme data (p-amylase and lipase) will be summarized by treatment and nominal visit. The number and proportion of patients with pancreatic enzyme values exceeding the following thresholds will be provided by treatment, baseline pancreatic enzyme value (≤Upper Limit of Normal [ULN], >ULN), and nominal visit: (>1 to ≤3) × ULN, (>3 to ≤5) × ULN, (>5 to ≤10) × ULN, >10 × ULN.

Additionally, the number and proportion of patients with maximum postbaseline pancreatic enzyme values exceeding the following thresholds will be provided by baseline pancreatic enzyme value (≤ULN, >ULN), and treatment: (>1 to ≤3) × ULN, (>3 to ≤5) × ULN, (>5 to ≤10) × ULN, >10 × ULN.

An MMRM analysis will be used to analyze each pancreatic enzyme with a log transformed (postbaseline measure/baseline measure) response variable and stratification factors, treatment, nominal visit, and treatment-by-nominal visit interaction as fixed effects, and baseline value as a covariate.

6.13.2.4. Thyroid Malignancies and C-Cell Hyperplasia

Treatment-emergent thyroid malignancies and C-cell hyperplasia will be identified using predefined MedDRA high level terms (HLTs) of thyroid neoplasms malignant and PT of thyroid C-cell hyperplasia. Detailed searching criteria can be found in Appendix 1. A summary by treatment and PT, PT within SMQ, and HLT will be provided. Thyroid malignancies and C-cell hyperplasia will be considered AESIs.

6.13.2.5. Malignancies

The AE database will be searched using predefined SMQs to identify events consistent with malignancy. Detailed searching criteria can be found in Appendix 1. A summary by treatment and PT within SMQ and a listing of TEAEs will be provided. Malignancy will be considered an AESI.

6.13.2.5.1. Calcitonin

Observed calcitonin data will be summarized by treatment and nominal visit. Additionally, the number and proportion of patients with a calcitonin value exceeding the following thresholds will be provided by treatment, baseline calcitonin value (≤20 ng/L, >20 ng/L), nominal postbaseline visit, and postbaseline calcitonin categories: ≤20 ng/L, >20 ng/L to ≤35 ng/L, >35 ng/L to ≤50 ng/L, >50 ng/L to ≤100 ng/L, >100 ng/L.

Additionally, the number and proportion of patients with a maximum postbaseline calcitonin value exceeding the following thresholds will be provided by treatment and baseline calcitonin value (≤20 ng/L, >20 ng/L): ≤20 ng/L, >20 ng/L to ≤50 ng/L, >50 ng/L to ≤100 ng/L > 100 ng/L.

6.13.2.6. Major Adverse Cardiovascular Events (MACE)

Major adverse cardiovascular events reported by investigators are adjudicated by an independent clinical endpoint committee (CEC) in a blinded fashion. The major adverse cardiovascular events (MACE) events of special interest are: deaths due to CV cause, myocardial infarction, hospitalization for unstable angina, hospitalization for heart failure, coronary interventions (such as coronary artery bypass graft [CABG] or percutaneous coronary intervention [PCI]); and cerebrovascular events, including cerebrovascular accident (stroke) and transient ischemic attack (TIA).

A listing of patients reporting MACE events, either reported by investigator or identified by the CEC, will be provided. The listing will include treatment, patient identification including the site number, date of event, type of event as reported by the investigator, type of event as adjudicated by the CEC, time from first dose of study drug to the event, and time from last dose to the event (if patient has discontinued study drug prior to the event). Only adjudication-confirmed MACE will be considered an AESI.

6.13.2.7. Supraventricular Arrhythmias and Cardiac Conduction Disorders

The AE database will be searched using predefined SMQs or MedDRA HLTs to identify events consistent with supraventricular arrhythmias and cardiac conduction disorders. Detailed searching criteria can be found in Appendix 1. Incidence of the resulting TEAEs will be summarized by treatment and PT within SMQ and HLT. Treatment-emergent severe/serious supraventricular arrhythmias and cardiac conduction disorders will be considered AESIs.

6.13.2.8. Hypersensitivity Events

Hypersensitivity reactions and related information reported via the "Hypersensitivity and Anaphylactic Reactions" eCRF will be summarized by treatment. Two main analyses are performed:

- **Potential Immediate Hypersensitivity:** Analysis of TEAEs occurring from the start of study drug administration up to 24 hours after the end of study drug administration. For events without the hypersensitivity CRF, only date (no time) information is collected. Events occurring on the same date as the study drug injection date will be included.

- **Potential Non-Immediate Hypersensitivity:** Analysis of TEAEs occurring more than 24 hours after the end of study drug administration, but prior to subsequent study drug administration.

Summaries of all potential hypersensitivity reactions will be generated by PT with decreasing frequency by treatment. The AE database will be searched using predefined SMQs to identify events consistent with hypersensitivity events. Detailed searching criteria for hypersensitivity events can be found in Appendix 1. Severe/serious hypersensitivity events identified by predefined SMQ searches will be considered AESIs.

6.13.2.9. Injection Site Reactions

Injection site reactions, incidence, and related information reported via the "Injection Site Reactions" eCRF will be summarized by treatment. Information to be summarized includes the location of the reaction, timing of the reaction relative to study drug administration, and characteristics of the injection site reaction: erythema, induration, pain, puritus, and edema.

Additionally, potential injection site reactions will be reported by PT within MedDRA HLTs of injection site reactions, administration site reactions, and infusion-related reactions. Detailed searching criteria for injection site reaction events can be found in Appendix 1. The PT will be used for summary within each HLT category. Only severe/serious injection site reactions will be considered AESIs.

6.13.2.10. Immunogenicity

Treatment-emergent anti-drug antibodies (TE ADA) are defined as those with a titer 2-fold (1 dilution) greater than the minimum required dilution if no ADAs were detected at baseline (treatment-induced ADA) or those with a 4-fold (2 dilutions) increase in titer compared to baseline if ADAs were detected at baseline (treatment boosted ADA). A patient is evaluable for TE ADA if the patient has a non-missing baseline ADA result, and at least 1 non-missing postbaseline ADA result.

The frequency and percentage of patients with preexisting ADA, with TE ADA, with cross-reactive antibodies, and with neutralizing antibodies will be tabulated by dose, where proportions are relative to the number of patients who are TE ADA evaluable. The frequency and percentage of patients with hypersensitivity and injection site reactions by TE ADA status will be tabulated if warranted by the data.

A listing may be provided of all immunogenicity assessments for those patients who at any time had TE ADA present. This includes the tirzepatide concentration from a simultaneous pharmacokinetic (PK) sample, and the clinical interpretation result.

A listing may be provided for all participants who had ADAs present at any time (including baseline) or had any hypersensitivity or injection site reaction TEAEs.

Depending on the number of patients with TE ADA, selected efficacy and safety subgroup analyses by TE ADA categories may be performed if deemed necessary.

Treatment-emergent ADAs that are associated with AEs of either severe/serious hypersensitivity or severe/serious injection site reactions will be classified as AESIs.

6.13.2.11. Diabetic Retinopathy Complications

Results of the baseline dilated fundoscopic exam will be summarized by treatment. Any TEAE suspected of worsening retinopathy triggers a follow-up dilated fundoscopic exam. A summary of TEAEs suspected of worsening retinopathy and a summary of the results of the follow-up dilated fundoscopic exam will be summarized by treatment and PT. The cases with repeated fundoscopy during the course of the trial, based on clinical suspicion of worsening retinopathy that have either findings of de novo retinopathy or progression of retinopathy, and severe/serious adverse events from PTs defined in searching criteria in Appendix 1 will be considered AESIs and summarized.

6.13.2.12. Hepatobiliary Safety

6.13.2.12.1. Hepatobiliary Disorders

The AE database will be searched using SMQs to identify events consistent with hepatobiliary disorders. Detailed searching criteria can be found in Appendix 1. A summary by treatment and PT within SMQ will be provided. Severe/serious hepatobiliary disorders will be considered AESIs.

6.13.2.12.2. Acute Gallbladder Disease

The AE database will be searched using predefined SMQs to identify events consistent with acute gallbladder diseases. Detailed searching criteria for these AEs can be found in Appendix 1. A summary by treatment and PT within SMQ will be provided. Severe/serious acute gallbladder diseases will be considered AESIs.

6.13.2.12.3. Liver Enzymes

Analyses for laboratory analyte measurements are described in Section 6.16. This section describes additional analyses of liver enzymes. In addition, the following will be provided by treatment group:

- A shift table of maximum to maximum alanine aminotransferase (ALT) measurement from baseline to postbaseline with the following categories: ≤ upper limit of normal (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 to postbaseline with the following categories: ≤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: ≤ULN, >1 to <2 × ULN, ≥2 × ULN.
- Shift tables of serum alkaline phosphatase (ALP) from baseline to postbaseline with the following categories: ≤ULN, >1 to <2 × ULN, ≥2 × ULN.

Maximum baseline will be the maximum non-missing observation in the baseline period. The maximum value will be the maximum non-missing value from the postbaseline period. Planned and unplanned measurements will be included.

6.13.2.13. Gastrointestinal Safety

The time courses of prevalence and incidence (newly-occurring episodes) of nausea, vomiting, diarrhea, and combined will be plotted by treatment 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 by treatment.

The PTs in the gastrointestinal SOC will be used to identify gastrointestinal AEs. The incidence of the resulting TEAEs will be summarized by treatment and PT. Preferred Terms with severe/serious cases in the gastrointestinal SOC will be considered AESIs.

6.13.2.14. Acute Renal Events

Laboratory measures related to renal safety will be analyzed as specified for laboratory measurements in Section 6.16.

Two shift tables examining renal function will be created. A min-to-min shift table of estimated glomerular filtration rate (eGFR) estimated by the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation with units ml/min/1.73m^2, using categories (<45, ≥45 to <60, ≥60 to <90, and ≥90 mL/min/1.73m^2). A max-to-max shift table of urine albumin-to-creatinine ratio (UACR), using the categories UACR <30 mg/g, 30 mg/g ≤ UACR ≤300 mg/g, UACR >300 mg/g (respectively, these represent normal, microalbuminuria, and macroalbuminuria).

The AE database will be searched using SMQs of acute renal failure and chronic kidney disease to identify events consistent with acute renal events. The incidence of the resulting TEAEs will be summarized by treatment and PT. Detailed searching criteria can be found in Appendix 1. Severe/serious acute renal events will be considered AESIs.

6.13.2.15. Dehydration

The AE database will be searched using an SMQ of dehydration to identify events consistent with dehydration. Detailed searching criteria can be found in Appendix 1. Severe/serious dehydration events will be considered AESIs.

6.13.2.16. Metabolic Acidosis, Including Diabetic Ketoacidosis

The AE database will be searched using MedDRA PT to identify events consistent with metabolic acidosis, including diabetic ketoacidosis. Detailed searching criteria can be found in Appendix 1. The incidence of the resulting TEAEs will be summarized by treatment and PT. Severe/serious metabolic acidosis, including diabetic ketoacidosis, will be considered an AESI.

6.13.2.17. Amputation/Peripheral Revascularization

The AE database will be searched using MedDRA PT to identify events consistent with amputation or peripheral revascularization. The incidence of the resulting TEAEs will be

summarized by treatment and PT. Amputation/peripheral revascularization will be considered an AESI.

6.13.2.18. Major Depressive Disorder/Suicidal Ideation

The AE database will be searched using SMQs to identify events consistent with major depressive disorder or suicidal ideation. Detailed searching criteria can be found in Appendix 1. The incidence of the resulting TEAEs will be summarized by treatment and PT. Severe/serious major depressive disorder/suicidal ideation or behavior will be considered an AESI.

6.13.2.19. Treatment of Overdose

A listing of patients reporting overdosing of tirzepatide will be provided as a protocol deviation.

6.14. Vital Signs

Descriptive summaries by treatment and 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.

A MMRM using REML will be used to fit the changes from baseline in vital signs at all scheduled postbaseline visits. The model will include country/pooled country, baseline HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]), baseline concomitant oral antihyperglycemic medication (metformin alone vs metformin plus an SGLT-2i), treatment group, visit, and treatment-by-visit interaction as fixed effects, and baseline value of the dependent variable as a covariate. To model the covariance structure within patients, the unstructured covariance matrix will be used.

Counts and percentages of patients with abnormal sitting systolic blood pressure (BP), sitting diastolic BP, and pulse will be presented by treatment. The criteria for identifying patients with treatment-emergent vital sign abnormalities are stated in Table GPGH.6.4.

Table GPGH.6.4. Categorical Criteria for Abnormal Blood Pressure and Pulse Measurements

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

6.15. Electrocardiograms

Summary statistics by treatment and by nominal visit will be provided for ECG parameters (heart rate, PR, QRS, QT, and QT corrected using Fredericia's correction factor [QTcF]). When the QRS is prolonged (for example, a complete bundle branch block), QT and QTc should not be

used to assess ventricular repolarization. Thus, for a particular ECG, the following will be set to missing (for analysis purposes) when QRS is ≥120 msec: QT and QTcF.

The criteria for identifying patients with treatment-emergent quantitative ECG abnormalities is based on Table GPGH.6.5.

In addition, the percentage of patients with QT greater than 500 msec will be summarized, and the percentage of patients with QTcF greater than 500 msec will be summarized.

The percentage of patients who experienced a treatment-emergent increase from baseline in QTcF interval of greater than 30 msec, 60 msec, or 75 msec at any time will be summarized. The maximum value during the study follow up will be analyzed. Planned and unplanned measurements will be included.

Treatment-emergent qualitative ECG abnormalities are defined as qualitative abnormalities that first occurred after baseline. A listing of abnormal qualitative ECGs will be created.

Table GPGH.6.5. Selected Categorical Limits for ECG Data

	Low		High	
Parameter	**Males**	**Females**	**Males**	**Females**
Heart Rate (bpm)	<50 and decrease ≥15	<50 and decrease ≥15	>100 and increase ≥15	>100 and increase ≥15
PR Interval (msec)	<120	<120	≥220	≥220
QRS Interval (msec)	<60	<60	≥120	≥120
QTcF (msec)	<330	<340	>450	>470

6.16. Clinical Laboratory Evaluation

All laboratory data will be reported in the International System of Units. Selected laboratory measures will also be reported using conventional units. Values that are outside of reference ranges will be flagged as high (H) or low (L) in the listings. Descriptive summaries by treatment and by nominal visit will be provided for the baseline and postbaseline values as well as the change from baseline values.

Observed and change from baseline values for each visit will be displayed in box plots for patients who have both a baseline and a postbaseline planned measurement. Unplanned measurements will be excluded from box plots.

A shift table will be provided including unplanned measurements. The shift table will include the number and percentage of patients within each baseline category (low, normal, high, or missing) versus each postbaseline category (low, normal, high, or missing) by treatment. The proportion of patients shifted will be compared between treatments using Fisher's exact test.

For qualitative laboratory analytes, the number and percentage of patients with normal and abnormal values will be summarized by treatment.

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

6.17. Health Outcomes

The patient-reported outcome questionnaires will be completed by the patients at baseline and at 52 weeks (or early termination visit prior to 52 weeks). These include use of the mITT population on the EAS, and use of a 2-sided alpha level of 0.05 and a 2-sided 95% CI for pairwise comparisons.

No multiplicity adjustment will be made in the evaluation of health outcome measures. Item-level missingness is dealt with as per the instrument developers' instruction.

6.17.1. EQ-5D

Each item will be summarized descriptively by treatment at each scheduled visit at which the EQ-5D is administered. The changes from baseline to week 52 (last observation carried forward [LOCF]) in the index and Visual Analog Scale (VAS) scores will be analyzed using an ANCOVA model with model terms of treatment, country/pooled country, baseline HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]), and baseline concomitant oral antihyperglycemic medication use (metformin alone vs metformin plus an SGLT-2i) as fixed effects, and baseline EQ-5D score as a covariate.

6.17.2. Impact of Weight on Self-Perceptions Questionnaire

Descriptive summaries by treatment at each scheduled visit at which the IW-SP is administered will be presented for each item. Treatment comparisons of the raw and transformed overall IW-SP score changes from baseline to week 52 (LOCF) will be analyzed using an ANCOVA model with model terms of treatment, country/pooled country, baseline HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]), and baseline concomitant oral antihyperglycemic medication use (metformin alone vs metformin plus an SGLT-2i) as fixed effects, and baseline IW-SP score as a covariate.

6.17.3. Ability to Perform Physical Activities of Daily Living

Descriptive summaries by treatment at each scheduled visit at which the APPADL is administered will be presented for each item. Treatment comparisons of the raw and transformed overall APPADL score changes from baseline to week 52 will be analyzed using an ANCOVA model with model terms of treatment, country/pooled country, baseline HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]), and baseline concomitant oral antihyperglycemic medication use (metformin alone vs metformin plus an SGLT-2i) as fixed effects, and baseline APPADL score as a covariate.

6.17.4. Diabetes Treatment Satisfaction Questionnaire

The Diabetes Treatment Satisfaction Questionnaire (DTSQ) contains 8 items (conceptually the same items in the status [DTSQs] and change [DTSQc] versions). Six items (1, and 4 through 8)

are summed to produce a measure of treatment satisfaction and the 2 remaining items (2 and 3) are treated individually to assess, respectively, the perceived frequency of hyperglycemia and hypoglycemia. The DTSQs is used to assess treatment satisfaction at baseline and the DTSQc is used to assess relative change in satisfaction from baseline at week 52 or early termination.

Descriptive summaries will be provided at baseline (DTSQs only) and at 52 weeks (DTSQc only) for the perceived hyperglycemia item, perceived hypoglycemia item, and 6-item overall satisfaction score.

Treatment comparisons in the DTSQc at Week 52 will be analyzed using an ANCOVA model with model terms of treatment, country/pooled country, baseline HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]), and baseline concomitant oral antihyperglycemic medication use (metformin alone versus metformin plus an SGLT-2i) as fixed effects, and baseline DTSQs score as a covariate. The analyses will be conducted for the perceived hyperglycemia item, perceived hypoglycemia item, and 6-item overall satisfaction score.

6.18. Subgroup Analyses

Efficacy subgroup analyses will be guided by the efficacy estimand. Subgroup analyses for HbA1c will also be conducted for the treatment-regimen estimand.

Subgroup analyses may be done by country to support local regulatory registrations. Subgroups with few subjects may be excluded from subgroup analyses when appropriate.

6.18.1. Subgroup Analysis of HbA1c Change at 52 Weeks

Subgroup analyses by the following baseline characteristics will be provided: age group (<65, ≥65 years), age group (<75, ≥75 years), race, gender, ethnicity, region of enrollment (US, OUS), duration of diabetes (<median, ≥median), HbA1c (≤8.5%, >8.5%), type of antihyperglycemic medication use (metformin alone, metformin plus SGLT-2i), renal impairment (eGFR <60, ≥60), BMI group (≤27, >27), and BMI group (<30, ≥30 - <35, ≥35).

6.18.2. Subgroup Analysis of Weight Change at 52 Weeks

Subgroup analyses by the following baseline characteristics will be provided: age group (<65, ≥65 years), age group (<75, ≥75 years), race, gender, ethnicity, region of enrollment (US, OUS), duration of diabetes (<median, ≥median), HbA1c (≤8.5%, >8.5%), type of antihyperglycemic medication use (metformin alone, metformin plus SGLT-2i), renal impairment (eGFR <60, ≥60), BMI group (≤27, >27), and BMI group (<30, ≥30 - <35, ≥35).

6.18.3. Subgroup Analysis of TEAE Through Safety Follow-up

Subgroup analyses by the following baseline characteristics will be provided: age group (<65, ≥65 years), age group (<75, ≥75 years), race, gender, ethnicity, renal impairment (eGFR <60, ≥60), BMI group (≤27, >27), and BMI group (<30, ≥30 - <35, ≥35).

Other exploratory subgroup analyses may be performed as deemed appropriate.

6.19. Interim Analyses and Data Monitoring

A Data Monitoring Committee (DMC) will have the responsibility for periodic review of unblinded interim analysis results in order to monitor the safety of the patients in the study. A sponsor statistical analysis group external to the study team will perform the data analysis for the DMC. No interim analyses of efficacy will be conducted with a view of early study termination or study modification. Thus, interim analyses will have no bearing on type I error controls associated with final efficacy analyses. Detailed information regarding safety reviews including the statistical reports to be reviewed by the DMC will be described in the program DMC Charter.

6.20. COVID-19 Impact Assessment

This section lists the potential statistical analyses that may be performed to assess the impact of the COVID-19 pandemic when appropriate.

6.20.1. Patients Impacted by COVID-19

A listing of patients impacted by COVID-19 will be provided. The listing will include patient identification, treatment, date of impact, and description of impact.

6.20.2. Adverse Events

A summary table for patients with AEs related to COVID-19, including death due to COVID-19, COVID-19 SAEs, and COVID-19 AEs, will be provided by study treatment.

6.20.3. Patient Disposition

Patient disposition with reasons related to COVID-19 (such as, COVID-19 AE, patient decision, etc.) will be summarized for study and study treatment discontinuation by treatment group.

6.20.4. Study Visits

A summary of patients with study visits impacted by COVID-19 will be provided by treatment group. In this table, the number and proportion of patients missing study visits, including the primary endpoint visit, and having home health visits and/or virtual visits will be summarized.

6.20.5. Mitigation Summary

A summary table for patients having protocol deviations and mitigations due to COVID-19 (such as, missing study visit, having home health visit, etc.) will be provided by treatment group. An additional summary may be provided by country of enrollment and treatment group.

6.20.6. Measures Related to Primary and Key Secondary Objectives

Patients missing measures (HbA1c and body weight) related to primary and key secondary objectives will be summarized by visit and treatment group. In addition, the number of patients utilizing alternative options to in-person visits (such as local lab, home health visits, etc.) to collect primary and key secondary measures may be summarized by visit and treatment group.

7. Unblinding Plan

Study GPGH is an open-label trial. A joint unblinding plan with Study I8F-MC-GPGM was developed after the first version of this analysis plan but prior to the first unblinded data transfer. The purpose was to detail the steps taken to maintain study team blinding for the sponsor prior to final study database lock.

8. References

Rubin DB. *Multiple imputation for nonresponse in surveys*. New York: John Wiley & Sons Inc.; 1987.

9. Appendices

Appendix 1. Searching Criteria for Adverse Events of Special Interest

The AESI analyses are detailed in Section 6.13.2. The search criteria for each AESI are stored in CLUWE: T:\prd\ly3298176\common\AESI_Lab\Search criteria AESIs_TZP.xlsx.

Leo Document ID = 4ca17e8b-c741-4b72-a02b-2f51394e2e05

Approver: PPD

Approval Date & Time: 27-Jan-2021 17:01:01 GMT

Signature meaning: Approved