Statistical Analysis Plan (V2): I8F-MC-GPGL

A Phase 3, Randomized, Open-Label Trial Comparing Efficacy and Safety of Tirzepatide Versus Semaglutide Once Weekly as Add-on Therapy to Metformin in Patients With Type 2 Diabetes (SURPASS-2)

NCT03987919

Approval Date: 18-Feb-2021

1. Statistical Analysis Plan I8F-MC-GPGL: A Phase 3, Randomized, Open-Label Trial Comparing Efficacy and Safety of Tirzepatide versus Semaglutide Once Weekly as Add-on Therapy to Metformin in Patients with Type 2 Diabetes (SURPASS-2)

Tirzepatide (LY3298176) for Type 2 Diabetes Mellitus

Phase-3 randomized, open-label trial comparing 3 doses of tirzepatide to semaglutide in patients with type 2 diabetes.

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Protocol I8F-MC-GPGL
Phase 3

Statistical Analysis Plan Version 1 electronically signed and approved by Lilly: 26-Jul-2019.
Statistical Analysis Plan Version 2 electronically signed and approved by Lilly on date provided below.

Approval Date: 18-Feb-2021 GMT

2. Table of Contents

Section **Page**

Table of Contents

Table **Page**

Table of Contents

Figure **Page**

Table of Contents

Appendix **Page**

3. Revision History

Statistical Analysis Plan (SAP) Version 1 was approved prior to the first patient visit.

SAP Version 2 was approved before the final database lock. The key changes are summarized below:

1. Updated the definition of analysis set. Per agreement with the United States (US) Food and Drug Administration (FDA), excluded patients who discontinued study drug due to inadvertent enrollment from efficacy analyses.
2. Updated baseline definition for selected measures.
3. Corrected the definition of "retrieved dropout."
4. Updated the language to handle lack of convergence in longitudinal logistic regression analysis due to possible low number of events for hemoglobin A1c (HbA1c) and weight loss target analyses.
5. Added section to assess SARS-Cov-2 (COVID-19) impact.
6. Added subgroup analyses for HbA1c (%) guided by treatment-regimen estimand (i.e., with multiple imputation).
7. Added achieving HbA1c<5.7% for tirzepatide 10 and/or 15 mg as a key secondary objective subject to type 1 error rate control and revision to the type 1 error rate control strategy

4. Study Objectives

4.1. Primary Objectives

Primary objectives of the study are to demonstrate that tirzepatide once weekly (QW) 10 mg and/or 15 mg are noninferior to semaglutide 1 mg in hemoglobin A1c (HbA1c) change from baseline to 40 weeks.

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

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

- To demonstrate that tirzepatide 5 mg is noninferior to semaglutide 1 mg for change from baseline in HbA1c at 40 weeks.
- To demonstrate superiority of QW tirzepatide 5 mg, and/or 10 mg, and/or 15 mg to semaglutide 1 mg for change from baseline in HbA1c at 40 weeks.
- To demonstrate superiority of QW tirzepatide 5 mg, and/or 10 mg, and/or 15 mg to semaglutide for change from baseline in body weight at 40 weeks.
- To demonstrate superiority of QW tirzepatide 5 mg, and/or 10 mg, and/or 15 mg to semaglutide for the proportion of patients with HbA1c target values of <7% (53 mmol/mol) at 40 weeks.
- To demonstrate superiority of QW Tirzepatide 10 and/or 15 mg to semaglutide for the proportions of patients with HbA1c target values of < 5.7% (39 mmol/mol) at 40 weeks.

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

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

To compare QW tirzepatide 5 mg, 10 mg, and 15 mg to semaglutide 1 mg at 40 weeks for:

- the proportion of patients achieving an HbA1c target value of ≤6.5% (48 mmol/mol)
- mean change in fasting serum glucose (central laboratory) from baseline
- mean change in daily average 7-point self-monitored blood glucose (SMBG) profiles from baseline
- proportion of patients who achieved weight loss of ≥5%, ≥10%, and ≥15% from baseline
- patient-reported outcomes
    - Diabetes Treatment Satisfaction Questionnaire Status (DTSQs) (baseline only); Diabetes Treatment Satisfaction Questionnaire Change (DTSQc) (end of treatment only)
    - impact of weight on self-perception (IW-SP) (baseline and end of treatment)

  - Ability to Perform Physical Activities of Daily Living (APPADL) (baseline and end of treatment)
  - European Quality of Life: 5 dimensions, 5 levels (EQ-5D-5L) (baseline and end of treatment)
  - Impact of Weight on Quality of Life Clinical Trials- Lite (IWQOL-Lite-CT) (baseline an end of treatment)
- changes in fasting glucagon, C-peptide, and insulin levels
- changes in lipids (total cholesterol, high-density lipoprotein [HDL], very low-density lipoprotein [VLDL], and triglycerides)
- changes from baseline in mean body mass index (BMI)
- mean change in waist circumference
- biomarkers

To compare QW tirzepatide 5 mg, to semaglutide 1 mg at 40 weeks for:

- the proportion of patients achieving an HbA1c target value of <5.7% (39 mmol/mol)

4.4. Safety Objectives

To compare the safety of tirzepatide 5 mg, 10 mg, and 15 mg to semaglutide 1 mg for:

- treatment-emergent adverse events (TEAEs)
- early discontinuations 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 (TE-ADAs) to tirzepatide
- mean change in systolic and diastolic blood pressure and heart rate from baseline
- occurrence of hypoglycemic episodes
- time to initiation of rescue therapy for severe persistent hyperglycemia

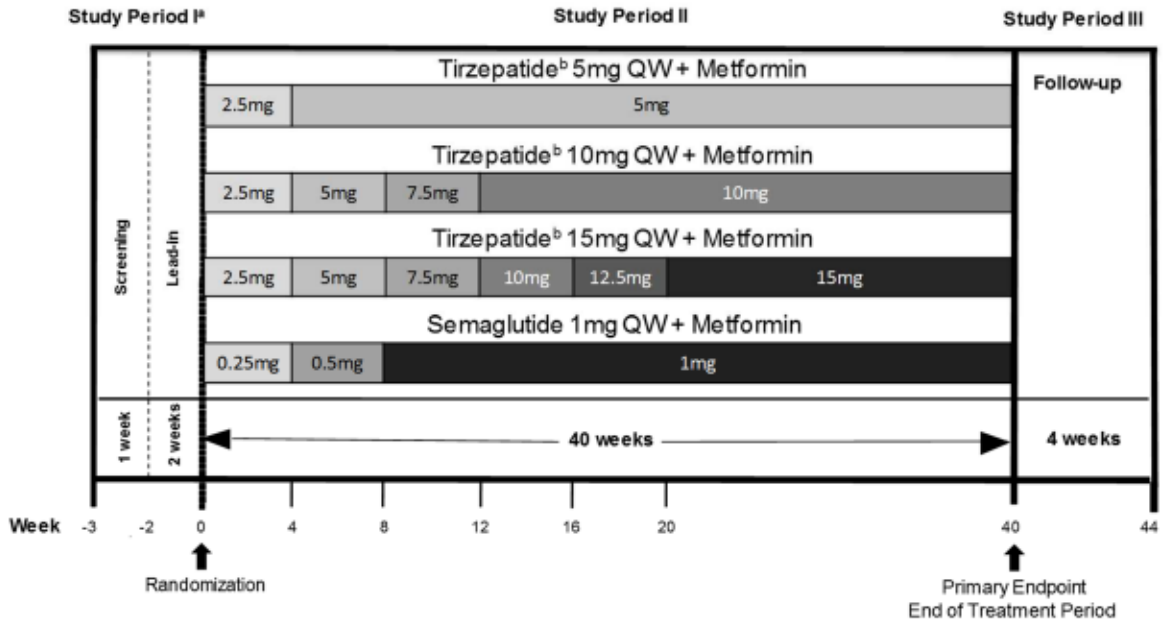
5. Study Design

5.1. Summary of Study Design

Study I8F-MC-GPGL (GPGL) is a Phase 3, open-label, randomized, active -controlled, multicenter, parallel-arm, study to compare the safety and efficacy of 3 doses of tirzepatide with semaglutide in patients with type 2 diabetes mellitus (T2DM) inadequately controlled with ≥1500 mg/day of metformin alone. The primary endpoint will be the mean change in HbA1c from baseline to 40 weeks.

The patient and investigator will know if the patient is going to receive tirzepatide or semaglutide treatment. However, the tirzepatide doses will be blinded.

Figure GPGL.5.1 illustrates the study design.



Abbreviation: QW = once-weekly.
[a] Stable doses of metformin ≥1500 mg/day for at least 3 months prior to Visit 1 and during the screening/lead-in period.
[b] All tirzepatide doses will be double-blinded.

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

Study Period I (screening and lead-in)

Screening (Visit 1 [Week -3])

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. The patient will sign the informed consent form before any study procedures are performed.

Lead-in (Visit 2 [Week -2] to Visit 3 [Week 0])

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 (40-week treatment period)

Randomization (Visit 3 [Week 0])

At Visit 3, 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 [Week 0] to Visit 11 [Week 40]):

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 for the 5-mg arm. For the 10-mg arm, 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 arm, 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.

For the semaglutide arm, the starting dose will be 0.25 mg QW for 4 weeks, then the dose will be increased to 0.5 mg QW for 4 weeks, followed by an increase to 1 mg QW for the duration of the study.

Study Period III (safety follow-up period)

Safety Follow-Up (Visit 801 [Week 44]) 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. Initiation of new antihyperglycemic therapy for the safety follow-up period will not be classified as “rescue therapy.”

5.2. Method of Assignment to Treatment

Approximately 1872 patients (468 in each treatment group) 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. Patients will be randomized in a 1:1:1:1 ratio to one of the treatment arms: 5 mg tirzepatide, 10 mg tirzepatide, 15 mg tirzepatide, or 1 mg semaglutide. The randomization will be stratified by country and baseline HbA1c concentration (≤8.5%, >8.5% [≤69, >69 mmol/mol]).

6. A Priori Statistical Methods

6.1. Populations for Analyses

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

Table GPGL.6.1. Analysis Populations/Data Sets

Population/Data Set	Description
Screened population	All participants who sign informed consent
Randomized population	All patients who are randomly assigned a treatment arm
Modified intent-to-treat (mITT) population	All randomly assigned participants who took 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 mITT, excluding patients discontinued study drug due to inadvertent enrollment and data after initiating rescue antihyperglycemic medication or prematurely stopping study drug (last dose date + 7 days).
Full analysis set (FAS)	All available data obtained during Study Period II from mITT, excluding patients discontinued study drug due to inadvertent enrollment, regardless of adherence to study drug or initiation of rescue antihyperglycemic medication.
Safety analysis set (SS)	All available data obtained during Study Period II or III from mITT population, regardless of adherence to study drug or initiation of new 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. 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.

Additionally, 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 one 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. First estimand, the "efficacy" estimand, represents efficacy prior to discontinuation of study drug or initiating rescue therapy for severe persistent hyperglycemia. Analysis relative to the "efficacy" estimand will be conducted using the efficacy analysis set (EAS). Second estimand, the "treatment-regimen" estimand, represents the efficacy irrespective of adherence to investigational product or introduction of rescue therapy for severe persistent hyperglycemia. Analysis relative to "treatment-regimen" estimand will be conducted using full analysis set (FAS).

Unless specified otherwise, safety analyses will be conducted relative to "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 death, date of ET or date of safety follow-up) will be excluded from statistical analysis. Listings of such data may 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, and standard deviation, median, minimum, and maximum. The summary statistics will be presented by the nominal visit.

Statistical treatment comparisons will only be performed between tirzepatide doses and semaglutide. Since the trials are not adequately powered to detect differences among tirzepatide doses, comparisons among tirzepatide arms will not be performed unless otherwise specified.

Statistical summaries and results of statistical analysis will be displayed in the following treatment order: 5 mg tirzepatide, 10 mg tirzepatide, 15 mg tirzepatide, and 1 mg semaglutide.

6.3. Adjustments for Covariates

The study is stratified by country and baseline HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]). 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, pooled country will be used as stratification factors and baseline HbA1c as a covariate. For other efficacy analyses, pooled country, and baseline HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]) will be used as stratification factors and respective baseline value as a covariate. Stratification factors will be based on the information collected in clinical database.

6.4. Handling of Dropouts or Missing Data

For the primary and secondary efficacy endpoint analyses subject to type 1 error rate control, data for patients with missing values at the 40-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 analysis. 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 Adjustments

Type 1 error rate control strategy for primary and key secondary efficacy objectives is illustrated in Section 6.12.3. No multiplicity adjustments will be made for conducting separate analyses relative to "efficacy" and "treatment-regimen" estimands, 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.

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

6.8. Patient Characteristics

A listing of patient demographics will be provided. All demographic and baseline clinical characteristics will be summarized by study treatment for the patients in the modified intent-to-treat (mITT) population. Baseline demographic and clinical characteristics of special interest include but not limited to: age, gender, race, weight, BMI, country of enrollment, HbA1c, fasting glucose, duration of T2DM, and estimated glomerular filtration rate (eGFR).

6.9. Concomitant Therapy

The prespecified concomitant medications of interest will be summarized by treatment at randomization. Additionally, medications of interest initiated after randomization and change to medications of interest used at randomization will be summarized. The concomitant therapies will be mapped using the World Health Organization Drug dictionary in the clinical trial database.

The concomitant medications of interest include the following groups of medication:

- baseline metformin total daily dose
- baseline antihypertensive therapy by type
- baseline lipid lowering therapy by type
- utilization of other antihyperglycemic therapy in Study Period II and Study Period III
- utilization of the following medications in Study Period II
    - antihypertensive therapy
    - lipid lowering therapy
- rescue therapy due to severe persistent hyperglycemia
- initiation of the following medications in Study Period II:
    - 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.

Summary of duration to the end of the study (defined as time in days from date of randomization to date of safety follow-up or date of early study discontinuation) 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 study treatment.

6.10.1. Exposure and Compliance to Study Treatment

The number of patients prematurely discontinuing study treatment prior to the 40-week visit will be provided by study treatment. Reasons for prematurely discontinuing study treatment prior to the 40-week visit will be provided by study treatment. Time-to-event analysis of premature study treatment discontinuation will be conducted.

If data warrant, the proportion of patients receiving 2.5 mg, 5 mg, 7.5 mg, 10 mg, 12.5 mg, or 15 mg for tirzepatide treatment groups, and 0.25 mg, 0.5 mg, and 1 mg for the semaglutide treatment group may be presented by randomized treatment and time intervals from first dose.

A listing of patients who missed ≥3 consecutive doses may be produced.

Overall treatment compliance will be defined as taking at least 75% of the scheduled tirzepatide/semaglutide 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 by treatment using the mITT population.

6.11. Important Protocol Deviations

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

6.12. Efficacy Analysis

For the FDA and potentially for other regulatory agencies, all efficacy assessments will be guided by the "treatment-regimen" estimand conducted using 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 40 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 data set without imputation of missing data. A listing of patients randomized but not included in efficacy analyses (i.e., not treated, discontinued treatment due to inadvertent enrollment) will be provided.

6.12.1. Primary Efficacy Analysis

The primary efficacy measure will be change in HbA1c from baseline (postbaseline - baseline) at 40 weeks. Values for HbA1c and change from baseline in HbA1c will be summarized by treatment and nominal visit (week). When applicable, HbA1c data from local lab will be used when central lab data is not available. If scheduled HbA1c data at the primary endpoint visit is not available, unscheduled HbA1c data collected for the primary endpoint visit will be included in analysis.

6.12.1.1. Primary Analysis Relative to the Efficacy Estimand

The analysis will be conducted utilizing HbA1c data in the EAS from baseline through the 40-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 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, pooled country 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 estimate of mean change from baseline in HbA1c will be summarized 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 40-week visit between 10 mg tirzepatide QW and 1 mg semaglutide QW, as well as between 15 mg tirzepatide QW and 1 mg semaglutide QW will be derived and summarized. If the upper limit of the CI is ≤0.3%, then the respective dose of tirzepatide will be declared noninferior to semaglutide relative to change in HbA1c from baseline.

6.12.1.2. Analysis Relative to the Treatment-Regimen Estimand

The analysis will be conducted utilizing HbA1c data in the FAS at baseline and at the 40-week visit with the aid of an analysis of covariance (ANCOVA). The response variable will be the primary measure and model terms will include treatment, pooled country as fixed effects, and baseline HbA1c as a covariate. The ANCOVA analysis will be conducted with multiple imputation of missing primary measure (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 40-week visit between 10 mg tirzepatide QW and 1 mg semaglutide QW, as well as between 15 mg tirzepatide QW and 1 mg semaglutide QW will be derived and summarized. If the upper limit of the CI is ≤0.3%, then the respective dose of tirzepatide (10 mg or 15 mg) will be declared noninferior to semaglutide relative to change in HbA1C from baseline.

6.12.1.3. Methods for Multiple Imputation

For efficacy analysis relative to "treatment-regimen estimand," missing HbA1c data at the 40-week visit will be imputed based on "retrieved dropouts", defined as patients who had their HbA1c value measured at the 40-week visit in the same treatment arm who prematurely discontinued study treatment. In cases when there are not enough retrieved dropouts to provide a reliable imputation model (i.e., the model implemented by the SAS program does not converge), an alternative multiple imputation method with reference to the baseline values (return-to-baseline multiple imputation) will be used. If value of the imputed HbA1c change from baseline is <-6.0% or >6.0%, that value will be set to -6.0% or 6.0%, respectively, to avoid unrealistic imputed values.

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 40-Week Visit

Noninferiority of 5 mg tirzepatide to 1 mg semaglutide will be conducted in a manner similar to the primary analysis in Section 6.12.1.

Assessment of superiority of tirzepatide doses compared with semaglutide 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 semaglutide (see details in Section 6.12.3).

6.12.2.2. Mean Change in Body Weight From Baseline at the 40-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 category (≤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 of the corresponding variable will be used as an additional covariate in the statistical model. Least squares mean estimate of mean change in body weight from baseline will be summarized by nominal visit and by study treatment. Treatment comparison between tirzepatide doses and semaglutide will be performed. For the multiple imputation of missing values, if value of the imputed weight change from baseline is <-50 kg or >50 kg, 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 40-Week Visit

The analysis relative to the "efficacy" estimand for the endpoint at 40 weeks will be performed using EAS with missing values imputed from an MMRM model and then dichotomized. The MMRM model includes treatment, stratification factors, visit, and treatment-by-visit interaction as fixed effects, and baseline HbA1c as a covariate. After dichotomizing continuous HbA1c, the data is analyzed using a logistic regression model with treatment, country/pooled country as fixed effects, and baseline HbA1c as a covariate. In addition, analysis will be conducted utilizing data from the EAS from baseline through the 40-week visit with the aid of a longitudinal logistic regression with repeated measurements with pooled country, treatment, visit, treatment-by-visit interaction as fixed effects, and baseline HbA1c as a covariate. In case the longitudinal logistic model does not converge due to small number of events, logistic regression will be utilized to analyze proportion of patients achieving HbA1c<7% at nominal visits.

Analysis relative to "treatment-regimen" estimand will be conducted utilizing HbA1c data from the FAS at baseline and at the 40-week visit with the aid of a logistic regression with multiple imputation of missing HbA1c data at the 40-week visit (see Section 6.12.1.3 for details). Model terms will include treatment, pooled country as fixed effects, and baseline HbA1c as a covariate and statistical inference over multiple imputations will be guided by Rubin (1987).

Similar analyses will be performed for proportion of patients achieving HbA1c <5.7%.

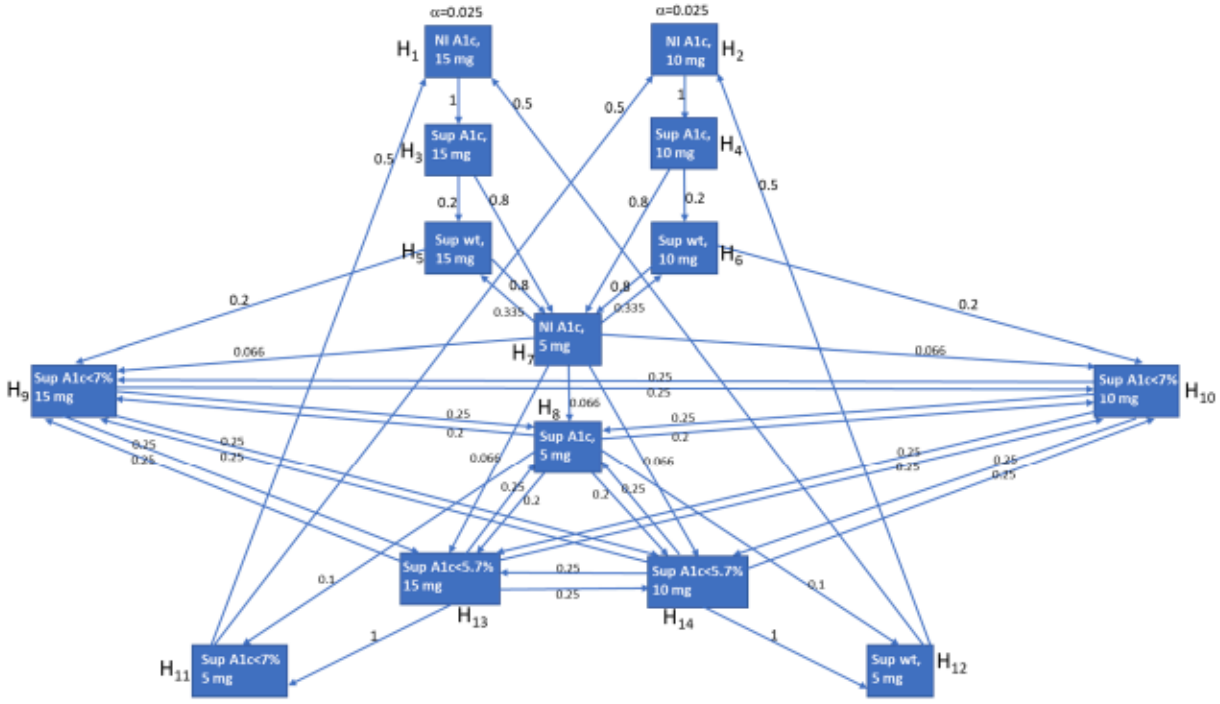
6.12.3. Type 1 Error Rate Control Strategy for Primary and Key Secondary Efficacy Analysis

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

The primary and key secondary objective hypotheses are as follows,

- H_1, and H_2: Noninferiority test of tirzepatide 10 mg and 15 mg versus semaglutide in HbA1c change from baseline at 40 weeks.
- H_3, and H_4: Superiority test of tirzepatide 10 mg, and 15 mg versus semaglutide in HbA1c change from baseline at 40 weeks.
- H_5, and H_6: Superiority test of tirzepatide 10 mg, and 15 mg versus semaglutide in body weight change from baseline at 40 weeks.
- H_7: Noninferiority test of tirzepatide 5 mg versus semaglutide in HbA1c change from baseline at 40 weeks.
- H_8: Superiority test of tirzepatide 5 mg versus semaglutide in HbA1c change from baseline at 40 weeks.
- H_9, H_{10}, and H_{11}: Superiority test of tirzepatide 10 mg, 15 mg, and 5 mg versus semaglutide in proportion of patients achieving HbA1c <7% at 40 weeks.
- H_{12}: Superiority test of tirzepatide 5 mg versus semaglutide in body weight change from baseline at 40 weeks.
- H_{13}, and H_{14}: Superiority test of tirzepatide 10 mg, and 15 mg versus semaglutide in proporation of patients achieving HbA1c < 5.7% at 40 weeks.

A graphical testing scheme (Bretz et al. 2009) presented in Figure GPGL.6.1 will be used to strongly control for type 1 error. H_1, and H_2 will be initially tested each at a 0.025 significance level.



Abbreviations: A1c =HbA1C; NI =noninferiority; Sup = superiority; wt =body weight.

Figure GPGL.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. Missing data will not be imputed and assessments are not subject to type 1 error rate control.

Table GPGL.6.2. Secondary and Exploratory Efficacy Analyses not Controlled for Type I Error

Objective	Relative to the efficacy measure:	Analysis conducted in a manner similar to:	Additional information
To compare tirzepatide 5 mg, 10 mg, and 15 mg to semaglutide 1 mg	Proportion of patients achieving an HbA1c target value of ≤6.5% (48 mmol/mol)	6.12.2.3	None
	Change from baseline in fasting serum glucose	6.12.1	Use baseline HbA1c strata as a fixed effect in place of baseline HbA1c as a covariate. Use baseline fasting serum glucose as a covariate.
	Change from baseline in 7-point self-monitored blood glucose (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. Least squares mean estimates at 40- weeks will be summarized 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 who achieved weight loss of ≥5%, from baseline	6.12.2.3	Use baseline HbA1c strata as a fixed effect in place of baseline HbA1c as a covariate and baseline weight as a covariate
	Proportion of patients who achieved weight loss of ≥10% from baseline	6.12.2.3	Use baseline HbA1c strata as a fixed effect in place of baseline HbA1c as a covariate and baseline weight as a covariate
	proportion of patients who achieved weight loss of ≥15% from baseline	6.12.2.3	Use baseline HbA1c strata as a fixed effect in place of baseline HbA1c as a covariate and baseline weight as a covariate

Secondary and Exploratory Efficacy Analyses not Controlled for Type I Error

Objective	Relative to the efficacy measures	Analysis conducted in a manner similar to:	Additional information
	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 baseline waist circumference 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 baseline BMI as a covariate.
	Change from baseline in fasting glucagon, c-peptide, insulin level, HOMA-B, HOMA-S, and HOMA-IR	6.12.1	Use baseline HbA1c strata as a fixed effect in place of baseline HbA1c as a covariate. Use corresponding baseline parameter as a covariate.
	Change from baseline in lipid parameters(total-cholesterol, HDL-C, VLDL-C, 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 patient reported outcomes: APPADL, IW-SP, EQ-5D-5L, IWQOL-Lite-CT	6.12.1.2	Use baseline HbA1c strata as a fixed effect in place of baseline HbA1c as a covariate. Use corresponding baseline patient outcome score as a covariate.
	DTSQc Score	6.12.1.2	Use baseline HbA1c strata as a fixed effect in place of baseline HbA1c as a covariate. Use corresponding baseline patient DTSQs score as a covariate.
	Proportion of patients achieving an HbA1c target ≤6.5%, ≥10% weight loss, and without hypoglycemia events with blood glucose level <54 mg/dL or severe hypoglycemia	6.12.2.3	Include baseline body weight as additional covariates.
To characterize efficacy (beyond the end of treatment) with tirzepatide 5 mg, 10 mg, and 15 mg	Change from baseline in HbA1c, body weight	6.12.1	The analysis will be conducted utilizing HbA1c (body weight) data from baseline through the end of follow-up (44 weeks).
To compare tirzepatide 5 mg, to semaglutide 1 mg	Proportion of patients achieving an HbA1c target value of <5.7% (39 mmol/mol)	6.12.2.3	None

Abbreviations: APPADL = Ability to Perform Physical Activities of Daily Living Questionnaire; DTSQC = Diabetes Treatment Satisfaction Questionnaire Change; DTSQs = Diabetes Treatment Satisfaction Questionnaire Change; BMI = body mass index; EQ-5D-5L = European Quality of Life: 5 dimensions, 5 levels; HbA1c = hemoglobin A1c; HDL-C = high-density lipoprotein cholesterol; HOMA-B = homeostatic model assessment for β-cell function; HOMA-IR = homeostatic model assessment for insulin resistance; HOMA-S = homeostatic model assessment for insulin sensitivity; IWQOL-Lite-CT = Impact of Weight on Quality of Life-Lite Clinical Trial Version; IW-SP = Impact of Weight on Self-Perceptions Questionnaire; VLDL-C – very-low-density lipoprotein cholesterol; TG = triglycerides.

6.13. Safety Analysis

Unless specified otherwise, safety assessments will be based on the SS (see Table GPGL.6.1). All events that occur between the date of the 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. For assessing benefit and risk profile through 40-weeks, selected safety analyses may be conducted by utilizing safety data from first dose through the date of 40-week visit. Some safety analysis may be conducted after excluding data after the initiation of new antihyperglycemic therapy.

Unless specified otherwise, comparisons of tirzepatide doses to semaglutide will be performed.

For selected safety parameters, the difference among treatment mean change from baseline in continuous safety parameters at all scheduled visits will be assessed via a MMRM using REML. The model will include country/pooled country, baseline HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]), 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 Section 6.12.1.1 will be tested in order.

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

Where specified, 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 years in specified time interval will be adjusted as an offset to account for possible unequal treatment duration in the specified time interval between patients.

6.13.1. Adverse Events

A TEAE is defined as an event that first occurred or worsened in severity after baseline. The Medical Dictionary for Regulatory Activities (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 preferred term (PT) nested within system organ class (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.

Overview of the number and percentage of patient who experienced a TEAE, serious adverse event (SAE), death, discontinued from study treatment, or study due to an AE, and relationship to study drug, , 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 severity.

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
- severe adverse events of special interest (severe AESI)

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, 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).

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 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. Listing will include but not limited to treatment, patient identification including the site number, treatment group, date of event, age at the time of enrollment, gender, 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 Treatment 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.

6.13.1.5. Treatment Overdose

A listing of patients reporting over-dosing of tirzepatide/semaglutide will be provided.

6.13.2. Special Safety Topics

6.13.2.1. Hypoglycemic Events

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

Table GPGL.6.3. Definitions of Hypoglycemic Event Categories

	Symptoms and/or Signs of Hypoglycemia	Blood Glucose Level
Glucose Alert Value:		
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 Hypoglycemia (Level 2):		
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 (Level 3)		

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, and considered as AESI.

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

To avoid duplicate reporting, all consecutive blood glucose 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 after initiation of a new antihyperglycemic therapy. For severe hypoglycemia and level 2 hypoglycemia, incidence as well as rate per patient year of exposure will be provided by treatment at specified time intervals. A listing of severe hypoglycemic events will also be provided. The incidence of hypoglycemic event will be analyzed using logistic regression with treatment, baseline HbA1C category, and country/pooled country as fixed effects. The rate of hypoglycemic episodes per patient year may 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 HbA1C category, and treatment as fixed effects. The logarithm of years 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.

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 sufficient number of episodes (>=10), time-to-first-event analyses for the initiation of rescue therapy will be conducted by treatment using a cox proportional regression model. For patients without event, "time-to-event" will be censored at end of treatment period. A listing of patients who initiated rescue therapy will be provided.

6.13.2.3. Pancreatitis

If data warrants, summaries of adjudicated and investigator-reported pancreatic events will be provided by treatment. Determination of investigator-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 Appendix 1. Treatment-emergent adjudication-confirmed pancreatitis will be considered as 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.

Additionally, the number and proportion of patients with maximum postbaseline pancreatic enzyme values exceeding the following thresholds will be provided by maximum baseline pancreatic enzyme value (≤1×upper limit of normal (ULN), >1×ULN), and treatment: ≤1×ULN, (>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 response variable and stratification factors, treatment, nominal visit, and treatment-by-nominal visit interaction as fixed effects, and log 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 pre-defined 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 and a listing will be provided. Thyroid malignancies and C-cell hyperplasia will be considered as AESI.

6.13.2.4.1. Calcitonin

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

6.13.2.5. Malignancy

The AE database will be searched using pre-defined 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 as AESI.

6.13.2.6. Major Adverse Cardiovascular Events

Major adverse cardiovascular events (MACE) reported by investigators are adjudicated by an independent CEC in a blinded fashion. The MACE events of special interest include: deaths due to cardiovascular 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). Deaths adjudicated as deaths due to undetermined cause by the CEC will be considered as deaths due to cardiovascular cause in statistical analysis.

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 the patient has discontinued study drug prior to the event). Only positively adjudicated MACE will be considered as AESI.

6.13.2.7. Supraventricular Arrhythmias and Cardiac Conduction Disorders

The AE database will be searched using pre-defined standardized MedDRA query (SMQ) or MeDRA HLT 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 as AESI.

6.13.2.8. Hypersensitivity Events

Hypersensitivity reactions and related information 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 eCRF, only date (no time) information are collected, the events occurred 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. Detailed searching criteria can be found in Appendix 1. Severe/serious hypersensitivity events identified by pre-defined SMQ search will be considered as 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 timing of the reaction relative to study drug administration, and characteristics of the injection site reaction: erythema, induration, pain, puritis, and edema. Patient based and event based summaries will be created.

Additionally, potential injection site reactions will be searched by pre-defined 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 by treatment within each HLT category. Only the severe/serious injection site reactions will be considered as AESI.

6.13.2.10. 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 the PTs defined in searching criteria in Appendix 1 will be considered as AESI and summarized.

6.13.2.11. Hepatic Safety

6.13.2.11.1. 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 as AESI.

6.13.2.11.2. Acute Gallbladder Disease

The AE database will be searched using pre-defined 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 as AESI.

6.13.2.11.3. Liver Enzymes

Analyses for laboratory analyte measurements are described in Section 6.13.5. 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 ($\leq 1 \times$ ULN, $> 1 \times$ ULN) to postbaseline with the following categories: $\leq 1 \times$ ULN, >1 to $<3 \times$ ULN, ≥ 3 to $<5 \times$ ULN, ≥ 5 to $<10 \times$ ULN, $\geq 10 \times$ ULN.
- A shift table of maximum to maximum aspartate transaminase (AST) measurement from baseline ($\leq 1 \times$ ULN, $> 1 \times$ ULN) to postbaseline with the following categories: $\leq 1 \times$ ULN, >1 to $<3 \times$ ULN, ≥ 3 to $<5 \times$ ULN, ≥ 5 to $<10 \times$ ULN, $\geq 10 \times$ ULN.
- Shift tables of maximum to maximum total bilirubin and direct bilirubin from baseline to postbaseline with the following categories: $\leq 1 \times$ ULN, > 1 to $<2 \times$ ULN, $\geq 2 \times$ ULN.
- Shift tables of serum alkaline phosphatase from baseline to postbaseline with the following categories: $\leq 1 \times$ ULN, >1 to $<2 \times$ ULN, $\geq 2 \times$ ULN.

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.

6.13.2.12. Gastrointestinal Safety

The time courses of prevalence and incidence (newly-occurring episodes)of nausea, vomiting, diarrhea, and combinedwill 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. The PTs with severe/serious cases in the gastrointestinal SOC will be considered as AESI.

6.13.2.13. Renal Safety

Laboratory measures related to renal safety will be analyzed as specified for laboratory measurements in Section 6.13.5.

Two shift tables examining renal function will be created. A minimum-to-minimum shift table of eGFR estimated by the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation with units mL/min/1.73m^2, using categories ((<30, ≥30 to <45, ≥45 to <60, ≥60 to <90, and ≥90 mL/min/1.73m^2). A maximum-to-maximum 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 for these AEs can be found in Appendix 1. Severe/serious acute renal events will be considered as AESI.

6.13.2.14. Dehydration

The AE database will be searched using 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 as AESI.

6.13.2.15. 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 as AESI.

6.13.2.16. Amputation/Peripheral Revascularization

The AE database will be searched using MedDRA PT to identify events for amputation or peripheral revascularization. The incidence of the resulting TEAEs will be summarized by treatment and PT. Amputation/Peripheral Revascularization will be considered as AESI.

6.13.2.17. 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 as AESI.

6.13.3. 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.

An 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]), treatment group, visit, and treatment-by-visit interaction 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 by treatment. The criteria for

identifying patients with treatment-emergent vital sign abnormalities are stated in Table GPGL.6.4.

Table GPGL.6.4. Categorical Criteria for Abnormal Treatment-Emergent 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

Abbreviations: BP = blood pressure; bpm = beats per minute.

6.13.4. 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 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 GPGL.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.

Table GPGL.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

Abbreviations: bpm = beats per minute; ECG = electrocardiogram; QTCF = Fredericia's corrected QT interval.

6.13.5. Clinical Laboratory Evaluation

All laboratory data will be reported in the International System of Units and 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 for selected measurements.

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

Shift tables will be produced for selected measurements. A shift table will include 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.

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

6.13.6. 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 (1:10) 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 reaction TEAEs by TE ADA status may be tabulated if data warrant.

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

A listing may be provided for all participant who had ADA present at any time (including baseline) or had any hypersensitivity or injection site reaction TEAE.

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 ADA that are associated with AEs of either severe/serious hypersensitivity or severe/serious injection site reaction will be classified as AESIs.

6.14. Health Outcomes

The patient-reported outcome questionnaires will be completed by the patients at baseline and at 40 weeks (or ET visit prior to 40 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 comparisons of tirzepatide doses to semaglutide.

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.

6.14.1. EQ-5D-5L

A descriptive frequency table of individual items in EQ-5D-5L questionnaire will present baseline, observed endpoint, and endpoint including last observation carried forward (LOCF, exclude baseline observation) values as separate summaries. Using EAS, the changes from baseline to Week 40,with and without 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]) as fixed effects, and baseline EQ-5D-5L score as a covariate. Analysis of covariance analyses on FAS will be performed as a sensitivity analysis.

6.14.2. Impact of Weight on Self-Perceptions Questionnaire

A descriptive frequency table of individual items in IW-SP questionnaire will present baseline, observed endpoint, and endpoint including last observation carried forward (LOCF, exclude baseline observation) values as separate summaries. Using EAS, treatment comparisons of the raw and transformed overall IW-SP score changes from baseline to Week 40, with and without LOCF, will be analyzed using an ANCOVA model with model terms of treatment, country, baseline HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]) as fixed effects, and baseline IW-SP score as a covariate. Analysis of covariance analyses on FAS will be performed as a sensitivity analysis.

6.14.3. Ability to Perform Physical Activities of Daily Living

A descriptive frequency table of individual items in APPADL questionnaire will present baseline, observed endpoint, and endpoint including last observation carried forward (LOCF, exclude baseline observation) values as separate summaries. Using EAS, treatment comparisons of the raw and transformed overall APPADL score changes from baseline to Week 40, with and without LOCF, will be analyzed using an ANCOVA model with model terms of treatment, country, baseline HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]) as fixed effects, and baseline APPADL score as a covariate. Analysis of covariance analyses on FAS will be performed as a sensitivity analysis.

6.14.4. Diabetes Treatment Satisfaction Questionnaire

The 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 40 or ET.

A descriptive frequency table of individual items in DTSQ questionnaire will present baseline (DTSQs only), observed endpoint (DTSQc only), and endpoint including last observation carried forward (LOCF, exclude baseline observation, DTSQc only) values as separate summaries.

Using EAS, treatment comparisons in the DTSQc at Week 40, with and without LOCF will be analyzed using an ANCOVA model with model terms of treatment, country, baseline HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]) 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. Analysis of covariance analyses on FAS will be performed as a sensitivity analysis.

6.14.5. Impact of Weight on Quality of Life-Lite Clinical Trials Version

A descriptive frequency table of individual items in IWQOL-Lite-CT questionnaire will present baseline, observed endpoint, and endpoint including last observation carried forward (LOCF, exclude baseline observation) values as separate summaries. Using EAS, treatment comparisons of the raw and transformed overall IWQOL-Lite-CT score changes from baseline to Week 40, with and withoutLOCF, will be analyzed using an ANCOVA model with model terms of treatment, country, baseline HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]) as fixed effects, and baseline IWQOL-Lite-CT score as a covariate. Analysis of covariance analyses on FAS will be performed as a sensitivity analysis.

6.15. Subgroup Analyses

Efficacy subgroup analyses will be guided by the efficacy estimand. Safety subgroup analyses will be guided by the treatment-regimen estimand using the SS dataset. Subgroup analysis of

HbA1c change at 40 weeks will also be guided by the treatment-regimen estimand in addition to analyses guided by efficacy estimand.

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

6.15.1. Subgroup Analysis of HbA1c Change at 40 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 and outside of US [OUS]), duration of diabetes (<median, ≥median), duration of diabetes (≤5, >5 to ≤10, >10 years), HbA1c (≤8.5%, >8.5%), renal impairment (eGFR <60, ≥60 mL/min/1.73m^2), BMI group (≤27, >27 kg/m^2) and BMI group (<30, ≥30 to <35, ≥35 kg/m^2).

6.15.2. Subgroup Analysis of Weight Change at 40 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 and OUS), duration of diabetes (<median, ≥median), HbA1c (≤8.5%, >8.5%), renal impairment (eGFR <60, ≥60 mL/min/1.73m^2), BMI group (≤27, >27 kg/m^2) and BMI group (<30, ≥30 - <35, ≥35 kg/m^2).

6.15.3. Subgroup Analysis of TEAE through Safety Follow-up

Subgroup analyses of TEAE occurring in >= 5% of subjects 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 mL/min/1.73m^2), BMI group (≤27, >27 kg/m^2) and BMI group (<30, ≥30 - <35, ≥35 kg/m^2).

Other exploratory subgroup analyses may be performed as deemed appropriate.

6.16. Interim Analyses and Data Monitoring Committee

No interim analyses is planned for this study.

6.17. COVID-19 Impact Assessment

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

6.17.1. Patients Impacted by COVID-19

Listings of patients with protocol deviation or mitigation due to COVID-19, patients with COVID-19 AEs or death, and patients dispositions with reasons related to COVID-19 will be provided.

6.17.2. Adverse Events

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

6.17.3. Patient Disposition

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

6.17.4. Study Visits

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

6.17.5. Mitigation Summary

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

6.17.6. Measures Related to Primary and Key Secondary Objectives

Patients missing measures (HbA1C, fasting glucose, 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

Unblinding plan will be in a separate document.

8. References

Bretz F, Maurer W, Brannath B, Posch M. A graphical approach to sequentially rejective multiple test procedures. *Stat Med*. 2009;28(4):586-604.

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 adverse events of special interest (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 = f01aec5c-f4ec-4490-9cf6-7b4687d05edd

Approver: PPD

Approval Date & Time: 18-Feb-2021 13:48:13 GMT

Signature meaning: Approved