Statistical Analysis Plan Version 2 I8F-JE-GPGO

A Phase 3 Study of Tirzepatide Monotherapy Compared to Dulaglutide 0.75 mg in Patients with Type 2 Diabetes Mellitus (SURPASS J-mono)

NCT03861052

Approval Date: 07-Apr-2021

1. Statistical Analysis Plan: I8F-JE-GPGO: A Phase 3 Study of Tirzepatide Monotherapy Compared to Dulaglutide 0.75 mg in Patients with Type 2 Diabetes Mellitus (SURPASS J-mono)

Tirzepatide for Type 2 Diabetes Mellitus

Phase 3, randomized, double-blind, parallel, active-controlled, 52-week study which will assess the safety and efficacy of tirzepatide (5, 10, and 15 mg), compared to dulaglutide 0.75 mg in patients with Type 2 Diabetes. Mellitus.

Eli Lilly Japan K.K
Kobe, Hyogo Japan

[Protocol I8F-JE-GPGO]
[Phase 3]

Statistical Analysis Plan electronically signed and approved by Lilly on date provided below.

Approval Date: 07-Apr-2021 GMT

2. Table of Contents

Section **Page**

1. Statistical Analysis Plan: I8F-JE-GPGO: A Phase 3 Study of Tirzepatide Monotherapy Compared to Dulaglutide 0.75 mg in Patients with Type 2 Diabetes Mellitus (SURPASS J-mono) 1
2. Table of Contents 2
3. Revision History 6
4. Study Objectives 7
4.1. Primary Objective 7
4.2. Secondary Objectives 7
5. Study Design 8
5.1. Summary of Study Design 8
5.2. Determination of Sample Size 9
5.3. Method of Assignment to Treatment 9
6. A Priori Statistical Methods 10
6.1. Populations for Analyses 10
6.2. General Considerations 10
6.3. Adjustments for Covariates 11
6.4. Handling of Dropouts or Missing Data 12
6.5. Multicenter Studies 12
6.6. Multiple Comparisons/Multiplicity 12
6.7. Patient Disposition 12
6.8. Patient Characteristics 12
6.9. Treatment Exposure and Compliance 12
6.10. Concomitant Therapy 13
6.11. Important Protocol Deviations 13
6.12. Efficacy Analysies 14
6.12.1. Primary Outcome and Methodology 14
6.12.2. Secondary Efficacy Analyses Subject to Type 1 Error Rate Control 15
6.12.2.1. Mean Change in Body Weight from Baseline at the 52-Week Visit 15
6.12.3. Type 1 Error Rate Control Strategy for Primary and Secondary Objective 15
6.12.4. Supplementary Analyses of the Primary and Secondary Outcomes 16
6.12.5. Secondary Efficacy Analyses 16

6.13. Safety Analyses 18
6.13.1. Adverse Events 18
6.13.1.1. Deaths 19
6.13.1.2. Other Serious Adverse Events 20
6.13.1.3. Discontinuation from Study Due to Adverse Event 20
6.13.1.4. Discontinuation from Study Drug Due to Adverse Event 20
6.13.1.5. Treatment of Overdose 20
6.13.2. Special Safety Topics 20
6.13.2.1. Hypoglycemic Events 20
6.13.2.2. Rescue Therapy for Severe Persistent Hyperglycemia 21
6.13.2.3. Pancreatitis 21
6.13.2.3.1. Pancreatic Enzyme Assessment 22
6.13.2.4. Thyroid Disease, Malignancies and C-Cell Hyperplasia 22
6.13.2.5. Malignancy 22
6.13.2.6. Calcitonin 22
6.13.2.7. Major Adverse Cardiovascular Events (MACE) 22
6.13.2.8. Supraventricular Arrhythmias and Cardiac Conduction Disorders 23
6.13.2.9. Hypersensitivity Events 23
6.13.2.10. Injection Site Reactions 23
6.13.2.11. Diabetic Retinopathy Complications 24
6.13.2.12. Hepatobiliary Safety 24
6.13.2.13. Hepatobiliary Disorders 24
6.13.2.13.1. Acute Gallbladder Disease 24
6.13.2.13.2. Liver Enzymes 24
6.13.2.14. Gastrointestinal Safety 25
6.13.2.15. Acute Renal Events 25
6.13.2.16. Metabolic Acidosis, Including Diabetic Ketoacidosis 25
6.13.2.17. Amputation/Peripheral Revascularization 26
6.13.2.18. Major Depressive Disorder/Suicidal Ideation 26
6.13.3. Vital Signs 26
6.13.4. Clinical Laboratory Evaluation 27
6.13.5. Immunogenicity 27
6.14. Health Outcomes 28
6.14.1. EQ-5D-5L 28
6.14.2. Diabetes Treatment Satisfaction Questionnaire (DTSQ) 29
6.15. Subgroup Analyses 29
6.16. Interim Analyses 30

7. Unblinding Plan 31
8. References 32
9. Appendices 33

Table of Contents

Table **Page**

Table GPGO.6.1. Analysis Populations/Data sets 10

Table GPGO.6.2. Definitions of Hypoglycemic Event Categories 20

Table GPGO.6.3. Categorical Criteria for Abnormal Blood Pressure and Pulse Measurements 27

3. Revision History

SAP Version 1 was created and approved prior to the first permanent data transfer.

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

1. Clarified baseline definition for selected measures to minimize missing values at baseline (Section 6.2).
2. Clarified the stratification factors used in the analyses will be based on the information collected in clinical database (Section 6.3).
3. Provided detailed description for supplemental analyses in Appendix 2.
4. Changed the analysis method for HbA1c and weight loss target analyses to the logistic regression using imputation from MMRM model (Section 6.12.5).
5. Update the language to handle lack of convergence in longitudinal logistic regression analysis due to low number of subjects who achieved the target for hemoglobin A1c (HbA1c) and weight loss target analyses (Section 6.12.5).
6. Add assessment of SARS-CoV-2 (COVID-19) impact (Section 6.7 and 6.11).
7. Added constipation to gastrointestinal safety (Section 6.13.4.14).
8. Move unblinding plan to the separated document.

4. Study Objectives

4.1. Primary Objective

Primary objective of the study is to demonstrate that once-weekly tirzepatide 5 mg, and/or 10 mg, and/or 15 mg are superior to dulaglutide 0.75 mg in HbA1c change from baseline to 52 weeks. in patients with T2DM who have discontinued OAM monotherapy or are OAM-naïve.

4.2. Secondary Objectives

Secondary objectives are to compare the efficacy of once-weekly tirzepatide versus dulaglutide 0.75 mg at 52 weeks relative to the following endpoints:

- Mean change in HbA1c
- Proportion of patients who achieve HbA1c <7%, ≤6.5%, and <5.7%
- Mean change in FSG
- Mean change in daily average 7-point SMBG profiles
- Mean change in body weight
- Proportion of patients who achieve weight loss of ≥5%, ≥10%, and ≥15% from baseline
- Mean change in fasting insulin
- Mean change in fasting C-peptide
- Mean change in HOMA-2

Secondary objectives are to compare the safety of once-weekly tirzepatide versus dulaglutide 0.75 mg at 52 weeks relative to the following endpoints:

· Incidence of TEAEs

· Early discontinuations of investigational product due to AEs

· Adjudicated all deaths and nonfatal major CV events

· Adjudicated pancreatic AEs

· Serum calcitonin

· Incidence of allergic and hypersensitivity reactions

· Incidence of injection site reactions

· Incidence of treatment-emergent ADAs to tirzepatide

· The 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

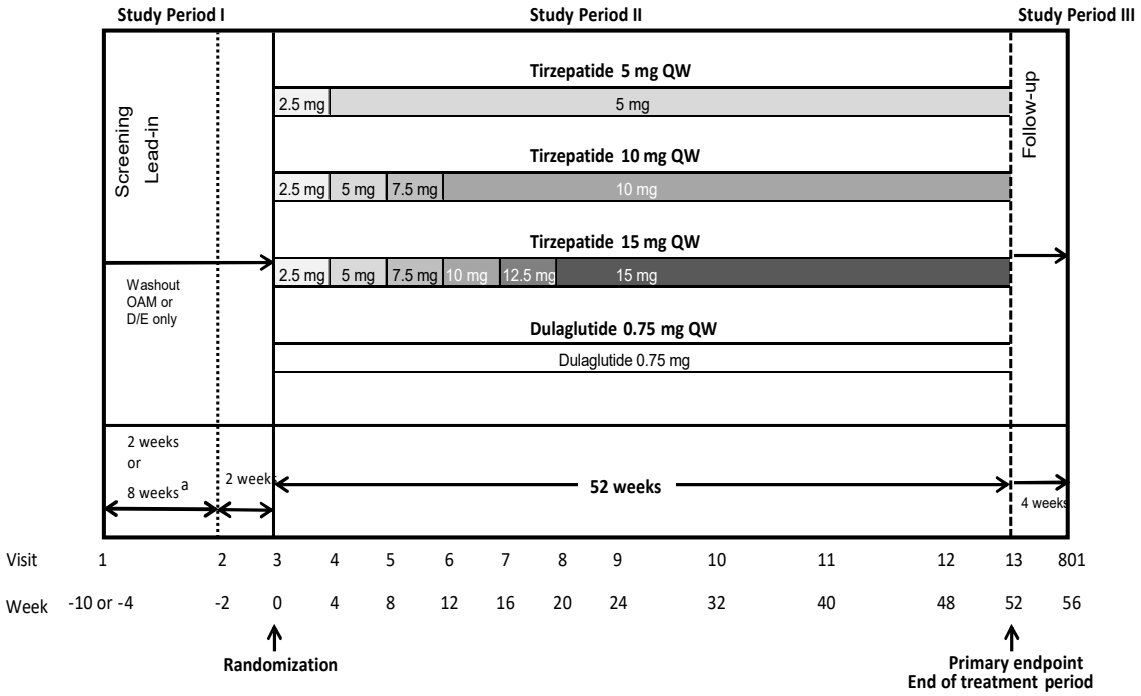
5. Study Design

5.1. Summary of Study Design

Study GPGO is a multicenter, randomized, double-blind, parallel, active-controlled, 52-week Phase 3 study which will assess the safety and efficacy of tirzepatide (5, 10, and 15 mg), compared to dulaglutide 0.75 mg in approximately 636 randomized patients with T2DM who have discontinued OAM monotherapy or are OAM-naïve.

Study GPGO will consist of 3 periods:

- a 4-week (OAM-naïve) or 10-week (at least 8-week OAM washout) screening/lead-in period
- a 52-week treatment period
- a 4-week safety follow-up period



Abbreviation: D/E = diet and exercise therapy; OAM = oral antihyperglycemic medication; QW = once -weekly.

a At least 8-weeks or more OAM monotherapy washout must be completed prior to V2.

Note: All doses will be administered QW subcutaneously using single-use pens.

Figure GPGO 5.1 Illustration of study design for Clinical Protocol I8F-JE-GPGO.

5.2. Determination of Sample Size

The trial is powered to assess superiority of tirzepatide doses (5 mg, and/or 10 mg, and/or 15 mg), versus dulaglutide 0.75 mg, relative to mean change from baseline in HbA1c at 52 weeks, under the following assumptions: use of 2 sample t-test to compare treatment means utilizing HbA1c data collected before initiation of any rescue medication and premature treatment discontinuation; up to 15% subjects initiating any rescue medication or premature treatment discontinuation; at least 0.5%, 0.5%, and 0.4% superior mean reduction in HbA1c from baseline at 52 weeks for tirzepatide 15, 10, and 5 mg, respectively, compared to dulaglutide 0.75 mg; and a common standard deviation (SD) of 1.0%. On the basis of these assumptions, randomizing 636 subjects using a 1:1:1:1 randomization ratio to tirzepatide 5 mg (159 patients), tirzepatide 10 mg (159 patients), tirzepatide 15 mg (159 patients), and dulaglutide 0.75 mg (159 patients) is required to ensure at least 90% power to establish superiority of tirzepatide 10-mg and/or 15-mg doses, compared to dulaglutide 0.75 mg at a 2-sided significance level of 0.025, followed by superiority of tirzepatide 5 mg, compared to dulaglutide 0.75 mg only if superiority of tirzepatide 10 mg or 15 mg is declared. Tirzepatide 5 mg will be tested at a 2-sided significance level of 0.05 if superiorities of both 10 mg and 15 mg are declared. Tirzepatide 5 mg will be tested at a 2-sided significance level of 0.025 if superiority of either 10 mg or 15 mg is declared. This parallel gatekeeping procedure controls family-wise Type 1 error rate at a 2-sided 0.05 level (Dmitrienko et al. 2003).

5.3. Method of Assignment to Treatment

Patients who meet all criteria for enrollment will be randomized to 1 of the double-blinded 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 0.75 mg dulaglutide. The randomization will be stratified by baseline HbA1c ($\leq$8.5% or >8.5%), baseline BMI (<25 or $\geq$25 kg/m^2), and washout of antidiabetic medication (yes or no).

6. A Priori Statistical Methods

6.1. Populations for Analyses

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

Table GPGO.6.1. Analysis Populations/Data sets

Population/Data Set	Description
Screened patients	All participants who sign informed consent
Randomized patients	All patients who are randomly assigned a treatment arm
Modified intention-to-treat (mITT) set	All randomly assigned participants who are exposed to at least 1 dose of investigational product. 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 data after initiating rescue antihyperglycemic medication or stopping investigational product (last dose date + 7 days).
Full analysis set (FAS)	Data obtained during Study Period II from mITT, regardless of adherence to investigational product or initiation of rescue antihyperglycemic medication.
Safety analysis set (SS)	Data obtained during both Study Period II and III from mITT, regardless of adherence to investigational product 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 change to the data analysis methods described in the protocol, and the justification for making the change, will be described in the statistical analysis plan (SAP) or the clinical study report (CSR). Additional exploratory data analyses may be conducted, as deemed appropriate. 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). Additional analyses of the data may be conducted as deemed appropriate without further changes made to the protocol or SAP, even after the database locks (DBL).

Unless otherwise, noted, all tests of treatment effects will be conducted at a 2-sided alpha level of 0.05, and the confidence interval (CI) will be calculated at 95%, 2-sided. In statistical summaries and analyses, patients will be analyzed as randomized.

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 prior to the first dose time will serve as baseline. For laboratoly

analyte, 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.

Efficacy and safety will be assessed using the modified intention-to-treat (mITT) population, which consists of all randomly assigned participants who are exposed to at least 1 dose of investigational product. The primary efficacy of tirzepatide versus dulaglutide 0.75 mg 52 weeks will be guided by the "efficacy" estimand using the efficacy analysis set (EAS). The "efficacy" estimand, represents efficacy prior to discontinuation of investigational product and prior to initiation of rescue therapy, that is without confounding effects of rescue therapy for persistent severe hyperglycemia.

Summary statistics for continuous measures will include sample size, mean, SD, median, minimum, and maximum.

For time to event analysis, Kaplan-Meier method will be used for estimation of cumulative event-free survival rates over time, and Cox proportional hazards regression analysis will be used to compare hazard rates among treatments.

Summary statistics for categorical measures (including categorized continuous measures) will include sample size, frequency, and percentages. Fisher's exact test will be used to examine the treatment difference in categorical outcomes. Logistic regression may be used to examine the treatment difference in binary efficacy outcomes. Summary statistics for discrete count measures will include sample size, mean, standard deviation (SD),, median, minimum, and maximum. The negative binomial regression model will be used for the treatment comparison of discrete count measures.

Statistical treatment comparisons will only be performed between tirzepatide doses and dulaglutide 0.75 mg. Since the trials are not adequately powered to detect differences among tirzepatide doses, comparisons among tirzepatide arms will not be performed unless otherwise specified.

Other statistical methods may be used, as appropriate, and details will be documented in the later sections.

6.3. Adjustments for Covariates

The randomization is stratified by baseline HbA1c group (≤8.5%, >8.5%), baseline BMI group (<25 or ≥25 kg/m^2), and washout of antidiabetic medication (yes or no). For HbA1c related analyses, baseline BMI group (<25 or ≥25 kg/m^2), and washout of antidiabetic medication (yes or no) will be used as fixed effects and baseline HbA1c as a covariate. For body weight related analyses, baseline HbA1c group (≤8.5%, >8.5%) and washout of antidiabetic medication (yes or no) will be used as fixed effects and baseline body weight as a covariate. Stratification factors used in the analyses will be based on the information collected in clinical database.

For other efficacy analyses, baseline HbA1c group (≤8.5%, >8.5%), baseline BMI group (<25 or ≥25 kg/m^2), and washout of antidiabetic medication (yes or no) will be used as fixed effects and respective baseline value as a covariate.

6.4. Handling of Dropouts or Missing Data

Missing efficacy parameter values and missing safety laboratory values will not be explicitly imputed unless otherwise stated. For health outcomes data, LOCF will be used for ANCOVA analysis and item-level missingness is dealt with as per instrument developers' instruction.

6.5. Multicenter Studies

Summary of final study disposition and study drug disposition for all randomized patients by site will be provided.

6.6. Multiple Comparisons/Multiplicity

Type 1 error rate control strategy for primary and selected efficacy objectives among tirzepatide arms compared to dulaglutide arm is described in Section 6.12.3. 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. Listing of patients who were not screen failures but who were not randomized will be provided.

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 drug disposition for all randomized patients will be provided by treatment. Patient study disposition both at the 52-week visit as well as at the end of safety follow-up visit (Visit 801) will be provided by treatment. Patient study drug disposition at the 52-week visit will be provided by treatment.

The listing for subjects who discontinued due to COVID-19 will be created if such patients exist.

6.8. Patient Characteristics

Listing of patient demographics will be provided for all randomized patients. All demographic and baseline clinical characteristics will be summarized by treatment for all randomized patients. Baseline demographic and clinical characteristics of special interest include and not limited to: age, gender, race, weight, HbA1c, fasting glucose, duration of T2DM, renal function, results of the fundal exam, and history of gallbladder disease.

6.9. Treatment Exposure and Compliance

Listing of patients randomized but not receiving study treatment will be provided, if applicable.

Summary of duration of follow-up (defined as time in days from date of randomization to date of safety follow-up) 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.

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 be provided by study treatment.

Treatment compliance, overall, is defined as taking at least 75% of the total required injections of investigational product. Compliance will be calculated by taking the number of doses administered (regardless the actual dose administered) divided by the total number of doses expected to be administered (during the actual treatment duration) ×100. Treatment compliance will be summarized by treatment groups using the mITT set.

Proportion of patients with missing dosing information, receiving no dulaglutide dose, receiving 0.75 mg dulaglutide will be presented by week from first dose for dulaglutide arm.

Proportion of patients with missing dosing information, receiving no tirzepatide dose, receiving 2.5 mg, 5 mg, 7.5 mg, 10 mg, 12.5 mg or 15 mg tirzepatide will be presented by randomized treatment and week from first dose. A listing of patients with treatment noncompliance or missed ≥3 consecutive doses will be produced.

6.10. Concomitant Therapy

The prespecified concomitant medications of interest will be summarized by treatment at randomization using the mITT set. The 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.

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

- Use of
    - Antihypertensive therapy, by type
    - Lipid lowering therapy, by type
    - Antiemetic medication
    - Antidiarrheal medication
- Rescue therapy due to severe persistent hyperglycemia

The frequencies and percentages of patients taking other concomitant medications before and after allocation (Visit 3) will be summarized. Patient listings for the use of concomitant medications may be provided.

6.11. Important Protocol Deviations

Important protocol deviations (IPD) are identified in Trial Issues Management Plan (TIMP). Trial Issue Management Plan specifies the method to capture IPD, either by programming or by non-programming. Programming method will find potential IPD. Those potential IPD will be

reviewed to identify IPD. Non-programming method is based on information from Simplicity Clinical Trial Management System (sCTMS).

A listing and a summary of IPD by treatment will be provided.

In addition, a listing and a summary of protocol deviations (PD) due to COVID-19 by treatment will be provided.

6.12. Efficacy Analysies

The efficacy assessment, guided by the "efficacy" estimand, will use Efficacy Analysis Set (EAS) which consists of data obtained before the initiation of any rescue therapy and before premature treatment discontinuation.

6.12.1. Primary Outcome and Methodology

The primary efficacy assessment, guided by the "efficacy" estimand, will use the EAS, which consists of data obtained before the initiation of any rescue therapy and before premature treatment discontinuation. The analysis model for change from baseline in HbA1c assessed over time up to the 52-week visit will be an MMRM. Restricted maximum likelihood (REML) will be used to obtain model parameter estimates and Kenward-Roger approximation 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, baseline BMI group (<25 or ≥ 25 kg/m^2) and washout of antidiabetic medication (yes or no) 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. Resulting Least Squares Mean (LSM) estimate of mean change from baseline in HbA1c will be summarized and plotted by visit and by study treatment.

With the aid of the MMRM analysis, p-values and 2-sided 95% confidence interval (CI) for mean change in HbA1c from baseline to 52-week visit for each of 5 mg, 10 mg, and 15 mg tirzepatide versus 0.75 mg dulaglutide will be derived and summarized.

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

6.12.2.1. 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. The analysis model for change from baseline in body weight assessed over time up to the 52-week visit will be an MMRM, with terms: treatment, visit, and treatment-by-visit interaction, baseline HbA1c group (≤8.5%, >8.5%) and washout of antidiabetic medication (yes or no) as fixed effects; and baseline body weight as a covariate. An unstructured covariance structure will model the relationship of within-patient errors. If this model fails to converge, the covariance structures will be tested in order described in Section 6.12.1.

6.12.3. Type 1 Error Rate Control Strategy for Primary and Secondary Objective

For analysis within each estimand, a graphical testing scheme (Bretz et al. 2009, 2011) presented in Figure GPGO.6.1 will be used to strongly control for type 1 error rate.

The primary and selected secondary objective hypotheses subject to Type 1 error rate control are as follows,

- $H_{5,1}$, $H_{10,1}$, and $H_{15,1}$: Superiority test of tirzepatide 5 mg, 10 mg, 15 mg versus 0.75 mg dulaglutide in HbA1C change from baseline at 52 weeks respectively.
- $H_{5,2}$, $H_{10,2}$, and $H_{15,2}$: Superiority test of tirzepatide 5 mg, 10 mg, 15 mg versus 0.75 mg dulaglutide in body weight change from baseline at 52 weeks respectively.

$H_{10,1}$ and $H_{15,1}$ will be initially tested each at 0.025 significance level.

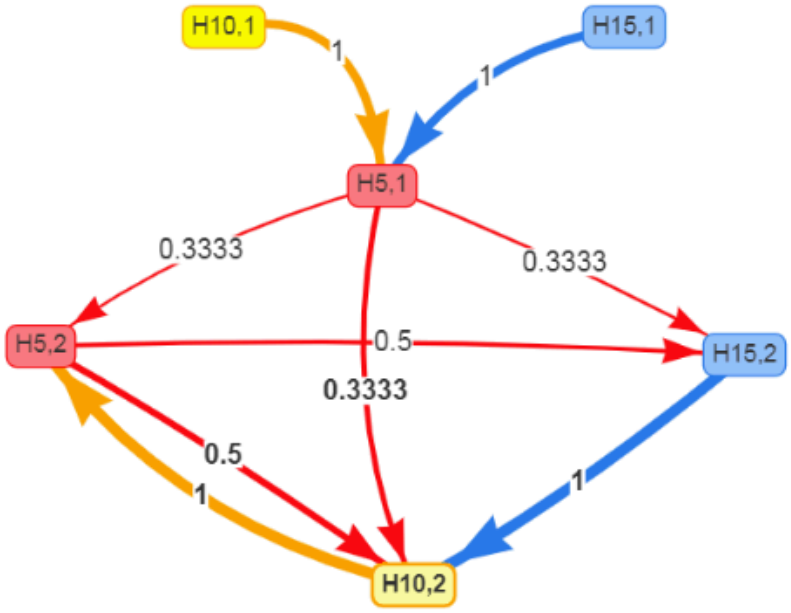


Figure GPGO.6.1. Type 1 error control strategy for primary and secondary efficacy endpoints.

6.12.4. Supplementary Analyses of the Primary and Secondary Outcomes

An efficacy assessment, guided by the "treatment-regimen" estimand, will be conducted for supplemental purpose. Details are described in Appendix 2. No multiplicity adjustments will be made for conducting 2 efficacy assessments for different estimands.

6.12.5. Secondary Efficacy Analyses

The following continuous efficacy measures will be summarized by treatment and nominal visit (week). The analyses for following efficacy measures will be conducted similar manner as HbA1c except for including baseline HbA1c group (≤8.5%, >8.5%) as fixed effect and baseline measurements as covariate in the model.

- Mean change in FSG
- Mean change in fasting insulin
- Mean change in fasting C-peptide
- Mean change in HOMA2-%B (beta-cell function) and HOMA2-%S (insulin sensitivity)
- Mean change in waist circumference

The changes from baseline to week 52 in 7-point SMBG profiles will be analyzed using an ANCOVA model with model terms of treatment, baseline HbA1c (≤8.5%, >8.5%), baseline BMI group (<25 or ≥25 kg/m^2), washout of antidiabetic medication (yes or no) as fixed effects, and

baseline measurements as a covariate. The following parameters derived from 7-point SMBG profile will be summarized and/or analyzed by treatment:

- Pre morning meal BG (mg/dL)
- 2-hour postprandial measurement for morning meal BG (mg/dL)
- Pre midday meal BG (mg/dL)
- 2-hour postprandial measurement for midday meal BG (mg/dL)
- Pre evening meal BG (mg/dL)
- 2-hour postprandial measurement for evening meals BG (mg/dL)
- Bedtime BG (mg/dL)
- Morning meal 2-hr excursion (mg/dL)
- Midday meal 2-hr excursion (mg/dL)
- Evening meal 2-hr excursion (mg/dL)
- Mean of all meals 2-hr excursion
- Mean of all 7-point BG (mg/dL)
- Mean of all pre-meals BG (mg/dL)
- Mean of all postprandial meals BG (mg/dL)
- Circadian variation in 7-point BG (mg/dL)

The following categorical efficacy measures will be summarized by treatment and nominal visit (week).

- Proportion of patients who achieve HbA1c <7%, ≤6.5%, and <5.7%
- Proportion of patients who achieve weight loss of ≥5%, ≥10%, and ≥15% from baseline

The analysis of the proportion of patients who achieve HbA1c target at the 52-week will be performed with missing values imputed from an MMRM model and then dichotomized. The MMRM model includes BMI group (<25 or ≥25 kg/m^2), washout of antidiabetic medication (yes or no), treatment, 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 BMI group (<25 or ≥25 kg/m^2), washout of antidiabetic medication (yes or no), treatment as fixed effects, and baseline HbA1c as a covariate. In addition, the analysis will be conducted utilizing data from baseline through the 52-week visit with the aid of a longitudinal logistic regression with repeated measurements with baseline BMI group (<25 or ≥25 kg/m^2), washout of antidiabetic medication (yes or no), treatment, visit, treatment-by-visit interaction as fixed effects, and baseline HbA1c as a covariate.

The analysis of the proportion of patients who achieve weight loss target at the 52-week will be performed with missing values imputed from an MMRM model and then dichotomized. The MMRM model includes baseline HbA1c group (≤8.5%, >8.5%), washout of antidiabetic medication (yes or no), treatment, visit, and treatment-by-visit interaction as fixed effects, and baseline body weight as a covariate. After dichotomizing continuous body weight change, the data is analyzed using a logistic regression model with baseline HbA1c group (≤8.5%, >8.5%), washout of antidiabetic medication (yes or no), treatment as fixed effects, and baseline body weight as a covariate. In addition, the analysis will be conducted utilizing data from baseline through the 52-week visit with the aid of a longitudinal logistic regression with repeated measurements with baseline HbA1c group (≤8.5%, >8.5%), washout of antidiabetic medication (yes or no), treatment, visit, treatment-by-visit interaction as fixed effects, and baseline body weight as a covariate.

In case longitudinal logistic model does not converge due to small number of patients who achieved the target, logistic regression may be utilized to analyze proportion of patients achieving weight loss target at nominal visits.

6.13. Safety Analyses

Unless specified otherwise, safety assessments will be guided by an estimand comparing safety of tirzepatide doses with dulaglutide 0.75 mg, irrespective of adherence to investigational product or initiation of rescue therapy. Thus, the safety analysis will be conducted using the SS (See Table GPGO.6.1).

Adverse events will be coded from the actual term using the Medical Dictionary for Regulatory Activities (MedDRA) and reported with preferred terms and system organ class. Selected notable AEs of interest may be reported using high-level terms or standardized MedDRA queries. Summary statistics will be provided for incidence of TEAEs, SAEs, study discontinuation due to AEs, investigational product discontinuation due to AEs, and deaths. Counts and proportions of subjects experiencing AEs will be reported for each treatment group. Comparisons of each tirzepatide doses to dulaglutide 0.75 mg will be performed using a Fisher's Exact test.

A selected safety analysis (e.g. hypoglycemia) will be conducted after excluding data while on rescue therapy or data after starting another antihyperglycemic medication.

6.13.1. Adverse Events

A treatment-emergent adverse event (TEAE) is defined as an event that first occurred or worsened in severity after the first dose. 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.

Overview of the number and percentage of patients who experienced at least one TEAE, serious adverse event (SAE), death , discontinued from study treatment due to an AE, discontinued from study treatment, or study due to an AE, and TEAE related to study drug, will be summarized by treatment.

The percentages of patients with all TEAEs and common TEAE (occurred in ≥5% in any treatment group), 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
- Permanent discontinuations of study or study treatment due to AEs
- Confirmed pancreatitis

"Notable" events are not limited to this list. Complete list of "notable" events will be described in Patient Narrative Planning Tool.

6.13.1.1. Deaths

A listing of all deaths and other SAEs will be provided. Listing will include patient identification including the treatment, site number, date of death, age at the time of enrollment, sex, associated AE group ID, 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, caused 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 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.

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 onanalysis of time to study drug discontinuation. by treatment as well as time to study drug discontinuation due to AE will be conducted.

6.13.1.5. Treatment of Overdose

A listing of patients reporting over-dosing of tirzepatide will be provided as a protocol deviation.

6.13.2. Special Safety Topics

6.13.2.1. Hypoglycemic Events

Definitions of different categories of hypoglycemic events are included in .

Table GPGO.6.2. Definitions of Hypoglycemic Event Categories

	Symptoms and/or Signs of hypoglycemia	Blood Glucose Level
Glucose Alert Value (Level 1):		
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):		
Documented symptomatic hypoglycemia	Yes	<54 mg/dL (3.0 mmol/L)
Documented asymptomatic hypoglycemia	No	<54 mg/dL (3.0 mmol/L)
Documented unspecified hypoglycemia	Unknown	<54 mg/dL (3.0 mmol/L)
Severe Hypoglycemia (Level 3)		

Severe hypoglycemia (Level 3): 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 as AESI.

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

To avoid duplicate reporting, all consecutive blood glucose (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 be provided for (a) hypoglycemia excluding events occurring after initiation of a new antihyperglycemic therapy and (b) hypoglycemia including events occurring after initiation of a new antihyperglycemic therapy. For each of the aforementioned hypoglycemia category, incidence as well as rate per patient year of exposure will be provided by treatment. Listing of hypoglycemic events, including events occurring after initiation of a new antihyperglycemic therapy, will be provided.

For each of the aforementioned hypoglycemia category, the incidence of hypoglycemic event will be analyzed using logistic regression with treatment, baseline BMI group (<25 or ≥25 kg/m^2), and washout of antidiabetic medication (yes or no) as a 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 mean modeled using treatment, baseline BMI group (<25 or ≥25 kg/m^2), and washout of antidiabetic medication (yes or no) as a fixed effect and baseline HbA1c as a covariate. The logarithm of days during 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. Rescue Therapy for Severe Persistent Hyperglycemia

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 time (in days) from first dose to end of study participation (study discontinuation or safety follow-up). A listing of patients initiating 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, number and proportion of patients with maximum postbaseline pancreatic enzyme value 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 with log transformed (postbaseline measure/baseline measure) as response variable and stratification factors, treatment, nominal visit, and treatment-by-nominal visit interaction as fixed effects, and baseline value as a covariate will be used to analyze each pancreatic enzyme.

6.13.2.4. Thyroid Disease, Malignancies and C-Cell Hyperplasia

Treatment-emergent Thyroid disease, C-cell hyperplasia, and neoplasms will be identified using a 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/PT within SMQ and HLT will be provided. Thyroid malignancies and C-cell hyperplasia will be considered as AESI.

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. Calcitonin

Observed calcitonin data will be summarized by treatment and by nominal visit. An MMRM analysis (log transformed) will be used similar to Section 6.13.2.3.1. Plot of time profile will be created.

Additionally, 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, >100ng/L

6.13.2.7. 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 MACE events of special interest are: 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).

A listing of patients reporting MACE events, either reported by investigator or identified by the CEC, will be provided. 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 positively adjudicated MACE will be considered as AESI.

Summaries of CEC adjudicated, and investigator-reported MACE events will be provided by treatment and event type.

Time-to- first event analysis of the composite end point consisting of CEC adjudicated: death due to cardiovascular cause, myocardial infarction, stroke, or hospitalization due to unstable angina (MACE-4) will be conducted.

6.13.2.8. Supraventricular Arrhythmias and Cardiac Conduction Disorders

Adverse event database will be searched using a pre-defined SMQ or MedDRA 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.

6.13.2.9. Hypersensitivity Events

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

a. **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.

b. **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 pre-defined 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 pre-defined SMQ search will be considered as AESIs.

6.13.2.10. 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, pruritus and edema.

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

6.13.2.12. Hepatobiliary Safety

6.13.2.13. 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.

6.13.2.13.1. 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.

6.13.2.13.2. Liver Enzymes

Analyses for laboratory analyte measurements are described in Section 6.13.4. This section describes additional analyses of liver enzymes. Summary and analysis of hepatic enzymes using CN units and SI units will be provided. In addition, the following will be provided by treatment groups:

- Shift table of maximum to maximum alanine aminotransferase (ALT) measurement from baseline (≤1 × ULN, > 1 × ULN) to postbaseline with the following categories: ≤1 × ULN, >1 to <3 × ULN, ≥3 to <5 × ULN, ≥5 to <10 × ULN, ≥10 × ULN.

- Shift table of maximum to maximum aspartate transaminase (AST) measurement from baseline (≤1 × ULN, > 1 × ULN) to postbaseline with the following categories: ≤1 × ULN, >1 to <3 × ULN, ≥3 to <5 × ULN, ≥5 to <10 × ULN, ≥10 × ULN..

- Shift tables of maximum to maximum total bilirubin and direct bilirubin from baseline to postbaseline with the following categories: ≤1 × ULN, > 1 to <2 × ULN, ≥2 × ULN.

- Shift tables of serum alkaline phosphatase (ALP) from baseline to postbaseline with the following categories: ≤1 × ULN, > 1 to <2 × ULN, ≥2 × 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.

6.13.2.14. Gastrointestinal Safety

Summaries of all events of treatment-emergent gastrointestinal (GI) AEs, including nausea, vomiting, diarrhea and constipation, will be provided by treatment and PT with decreasing frequency.

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

Onset (first occurrence) of GI AEs will be plotted by treatment.

The maximum severity and duration of treatment-emergent nausea, vomiting, diarrhoea, constipation 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. PTs with severe/serious cases in the gastrointestinal SOC will be considered as AESIs.

6.13.2.15. Acute Renal Events

Laboratory measures related to renal safety will be analyzed as specified for laboratory measurements in Section 6.13.4.

Additionally, 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 (<30, ≥30 to <45, ≥45 to <60, ≥60 to <90, and ≥90 mL/min/1.73m2). 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 as AESI.

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. Incidence of the resulting TEAEs will be summarized by treatment and PT. Severe/serious metabolic acidosis, including diabetic ketoacidosis will be considered as AESIs.

6.13.2.17. 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 AESIs.

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 as AESIs.

6.13.3. Vital Signs

Descriptive summaries by treatment and by nominal visits will be provided for the baseline and postbaseline values and 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 baseline HbA1c group (≤8.5%, >8.5%), baseline BMI group (<25 or ≥25 kg/m^2), washout of antidiabetic medication (yes or no), 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. If this model fails to converge, the covariance structures will be tested in order described in Section 6.12.1.

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 signs abnormalities are stated in Table GPGO.6.3.

Table GPGO.6.3. Categorical Criteria for Abnormal Blood Pressure and Pulse Measurements

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

Abbreviation: BP = blood pressure.

6.13.4. Clinical Laboratory Evaluation

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

A shift table with unplanned measurements. included will be provided. 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. Proportion of patients shifted will be compared between treatments. using Fisher's exact test.

For qualitative laboratory analytes, the numbers and percentages 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.

6.13.5. 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 neutralizing TE ADA, and with cross-reactive TE ADA to tirzepatide 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 will be tabulated if data warrant.

A listing will 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 PK sample, and the clinical interpretation result.

A listing of patients who are not TE ADA evaluable will be created

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

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

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

6.14. Health Outcomes

The patient-reported outcome questionnaires will be completed by the patients at baseline, at 40 weeks and at 52 weeks (or early termination visit prior to 52 weeks). These include use of the mITT population (all randomized patients who have taken at least 1 dose of study medication) on the EAS, and use of a 2-sided alpha level of 0.05 and 95% CI for comparisons of tirzepatide doses to dulaglutide 0.75 mg. No multiplicity adjustment will be made in the evaluation of health outcome measures. Item-level missingness is dealt with as per instrument developers' instruction.

6.14.1. EQ-5D-5L

Each item will be summarized descriptively by treatment at each scheduled visit at EQ-5D-5L is administered. The changes from baseline to week 40 and week 52 (LOCF) in the index and visual analog scale (VAS) scores will be analyzed using an MMRM model. The MMRM using REML will be used to fit change from baseline in each item at week 40 and week 52. The model will include baseline HbA1c group (≤8.5%, >8.5%), baseline BMI group (<25 or ≥25 kg/m^2), washout of antidiabetic medication (yes or no), 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. If this model fails to converge, the covariance structures will be tested in order described in Section 6.12.1. ANCOVA analyses will be performed as a sensitivity analysis. The ANOCVA model will include terms of treatment, baseline HbA1c group (≤8.5%, >8.5%), baseline BMI group (<25 or ≥25 kg/m^2), and washout of antidiabetic medication (yes or no) as fixed effects, and baseline EQ-5D-5L score as a covariate.

A Japanese specific EQ-5D-5L scoring algorithm will be used (Ikeda et al., 2015).

6.14.2. Diabetes Treatment Satisfaction Questionnaire (DTSQ)

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 two 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) at 40 and 52 weeks (DTSQc only) for the perceived hyperglycemia item, perceived hypoglycemia item and the six-item overall satisfaction score.

Treatment comparison in the DTSQc at week 40 and at week 52 will be analyzed using the MMRM model with terms of baseline HbA1c group (≤8.5%, >8.5%), baseline BMI group (<25 or ≥25 kg/m^2), washout of antidiabetic medication (yes or no), treatment group, visit, and treatment-by-visit interaction as fixed effects, and baseline DTSQs score 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 will be tested in order described in Section 6.12.1.

Treatment comparison in the DTSQc at week 40 and at week 52 (LOCF) will be analyzed using an ANCOVA model with model terms of treatment, baseline HbA1c group (≤8.5%, >8.5%), baseline BMI group (<25 or ≥25 kg/m^2), and washout of antidiabetic medication (yes or no) as fixed effects, and baseline DTSQs score as a covariate. ANCOVA analyses will be performed as a sensitivity analysis.

The analyses will be conducted for the perceived hyperglycemia item, perceived hypoglycemia item and the 6-item overall satisfaction score.

6.15. Subgroup Analyses

Subgroup analyses of mean change in HbA1c and mean change in body weight at week 52, AE and hypoglycemic events through safety follow-up will be provided.

The analyses model for change from baseline in HbA1c and change from baseline in body weight will be the MMRM model including terms: subgroup, subgroup-by-treatment, subgroup-by-visit and subgroup-by-treatment-by-visit interactions in addition to the models described in the Section 6.12.1 (HbA1c) and Section 6.12.2.1 (body weight). For subgroup analyses by baseline HbA1c group, baseline HbA1c as covariate will be removed from the analysis model. The subgroup-by-treatment interaction at week 52 will be evaluated using a significance level of 0.10.

Subgroup analyses will be conducted by the following baseline characteristics:

- Age group (<65 years, ≥65) and (<75 years, ≥75)
- Gender
- BMI group (<25 or ≥25 kg/m^2)

- BMI group (<27 or ≥27 kg/m^2)
- Washout of antidiabetic medication (yes or no)
- Duration of diabetes (<median, ≥median)
- HbA1c group (≤8.5%, >8.5%)
- Renal impairment (eGFR <60, ≥60 mL/min/1.73 m^2)

Other exploratory subgroup analyses may be performed as deemed appropriate.

6.16. Interim Analyses

No interim analyses are planned for this study.

7. Unblinding Plan

Unblinding plan is created 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.

Bretz F, Maurer W, Hommel G. Test and power considerations for multiple endpoint analyses using sequentially rejective graphical procedures. *Stat Med*. 2011;30(13):1489–1501.

Dmitrienko A, Offen WW, Westfall PH. Gatekeeping strategies for clinical trials that do not require all primary effects to be significant. *Stat Med*. 2003;22(15):2387-2400.

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

Ikeda S, Shiroiwa T, Igarashi I, Noto S, Fukuda T, Saito S, et al. Developing a Japanese version of the EQ-5D-5L value set (in Japanese; abstract in English). *J Natl Inst Public Health.* 2015;64:47–55.

9. Appendices

Appendix 1. Searching Criteria for Adverse Events of Special Interest (AESI)

The search criteria for each AESI are defined in AdaM ADAE specification.

Appendix 2. Supplementary Analyses of the Primary and Secondary Outcomes

An efficacy assessment, guided by the "treatment-regimen" estimand, will be conducted for supplemental purpose. This assessment will analyze change from baseline in HbA1c and body weight to the 52-week visit using an analysis of covariance (ANCOVA) model described in the Table GPGO.A.1. The ANCOVA analysis will be conducted using the Full Analysis Set (FAS) at the 52-week visit, which consists of all available data of changes from baseline in HbA1c or body weight data at the 52-week visit, irrespective of whether they were obtained while the participants had discontinued the investigational product or whether the participant had been given rescue medication.

Missing data at the 52-week for patients who prematually discontinued the IP are imputed based on "retrieved dropouts," defined as patients who discontinued IP (regardless rescue use) but had HbA1c value at 52-week visit in the same treatment arm, and for patients who were on-treatment up to 52-week visit (regardless rescue use) but did not have HbA1c value at 52-week visit (maybe rare) are imputed based on patients who were on-treatment (regardless rescue use) and had HbA1c at 52-week visit. 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.

The analysis for change in body weight will be conducted as similar manner as HbA1c. Details are described in the Table GPGO.A.1.

Analysis will be conducted with multiple imputations and statistical inference over multiple imputations will be guided by the method proposed by (Rubin 1987).

With the aid of the ANCOVA analysis, p-values and 2-sided 95% CI for mean change in HbA1c from baseline to 52-week visit for 5 mg, 10 mg, and 15 mg tirzepatide compared to dulaglutie 0.75 mg will be derived and summarized.

No multiplicity adjustments will be made for conducting 2 efficacy assessments for different estimands.

Table GPGO.A.1. Supplemental Analyses

Objective	Relative to the efficacy measure:	Analysis Model:	Additional Information for imputation
To compare tirzepatide 5 mg, 10 mg, and 15 mg to dulaglutide 0.75 mg	Change from baseline in HbA1c at 52-week visit	Fixed effect: treatment, baseline BMI group (<25 or ≥25 kg/m^2), washout of antidiabetic medication (yes or	If value of the imputed HbA1c change from baseline is <-6.0% or >6.0%, that value will be set to -6.0% or

		no), and baseline HbA1c as a covariate	6.0%, respectively, to avoid unrealistic imputed values.
To compare tirzepatide 5 mg, 10 mg, and 15 mg to dulaglutide 0.75 mg	Change from baseline in body weight at 52-week visit	Fixed effect: treatment, baseline HbA1c group (≤8.5%, >8.5%), washout of antidiabetic medication (yes or no), and baseline body weight as a covariate	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.

Leo Document ID = 1ec38521-64fd-4ec4-ab74-12ab3c551df0

Approver: PPD

Approval Date & Time: 07-Apr-2021 00:30:48 GMT

Signature meaning: Approved