

Statistical Analysis Plan Addendum 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 Addendum:
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)

**Meal Tolerance Test, Continuous Glucose Monitoring and
Body Composition**

Confidential Information

The information contained in this document is confidential and the information contained within it may not be reproduced or otherwise disseminated without the approval of Eli Lilly and Company or its subsidiaries.

Note to Regulatory Authorities: This document may contain protected personal data and/or commercially confidential information exempt from public disclosure. Eli Lilly and Company requests consultation regarding release/redaction prior to any public release. In the United States, this document is subject to Freedom of Information Act (FOIA) Exemption 4 and may not be reproduced or otherwise disseminated without the written approval of Eli Lilly and Company or its subsidiaries.

LY3298176 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 Addendum: 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) Meal Tolerance Test, Continuous Glucose Monitoring and Body Composition.....	1
2. Table of Contents.....	2
3. Revision History	3
4. Study Objectives	4
4.1. Objectives and Endpoints	4
4.2. Determination of Sample Size	5
5. A Priori Statistical Methods	6
5.1. General Considerations	6
5.2. Multiple Comparisons/Multiplicity.....	6
5.3. Patient Disposition	6
5.4. Patient Characteristics	6
5.5. Meal Tolerance Test.....	6
5.6. Continuous Glucose Monitoring.....	7
5.7. Body Composition Analyses	9
6. References	11

3. Revision History

SAP addendum Version 1 was created and approved prior to the first unblinding.

SAP addendum Version 2 was created and approved prior to the primary outcome database lock.

- Clarified CGM parameters to be analyzed (Section 5.6).
- Clarified analysis population for CGM (Section 5.6).
- Changed the threshold for the time in range and hypoglycemic for CGM data (Section 5.6) according to the guideline for glycemic target of American Diabetes Association (2021).
- Added body composition parameters and changed the analysis model (Section 5.7).

4. Study Objectives

4.1. Objectives and Endpoints

Table GPGO.1 shows the objectives and endpoints of Study GPGO protocol addendum, which are in addition to the Study GPGO protocol body.

Table GPGO.1. Objectives and Endpoints

Objectives	Endpoints
<p>Pharmacokinetics</p> <ul style="list-style-type: none"> To evaluate the pharmacokinetic profile of once-weekly tirzepatide in Japanese patients with T2DM using the commercial formulation 	<ul style="list-style-type: none"> AUC and C_{\max} of tirzepatide
<p>Exploratory</p> <ul style="list-style-type: none"> To compare the effect of once-weekly tirzepatide versus dulaglutide 0.75 mg on the pharmacodynamic profile and satiety after a standardized test meal at Week 32 To compare the effect of once-weekly tirzepatide versus dulaglutide 0.75 mg on daily glucose variability measured by CGM at Week 32 To compare the effect of once-weekly tirzepatide versus dulaglutide 0.75 mg on body composition at Week 12, 32, and 52 	<ul style="list-style-type: none"> The change from baseline in serum glucose $AUC_{(0-6h)}$ The change from baseline in insulin $AUC_{(0-6h)}$ The change from baseline in C-peptide $AUC_{(0-6h)}$ The change from baseline in glucagon $AUC_{(0-6h)}$ The change from baseline in triglyceride $AUC_{(0-6h)}$ The change from baseline in satiety (full and hungry) in VAS Daily average glucose (mg/dL) Time in range ≤ 70 mg/dL (minutes) Time in range > 180 mg/dL (minutes) Daily within-day SD (mg/dL) Daily MAGE (mg/dL) The change from baseline in total body water (L) The change from baseline in protein (kg) The change from baseline in minerals (kg) The change from baseline in body fat mass (kg) The change from baseline in lean body mass (kg)

Abbreviations: AUC = area under the concentration versus time curve; $AUC_{(0-6h)}$ = area under the concentration versus time curve from time zero to 6 hours after dose; C_{\max} = maximum concentration; CGM = continuous glucose monitoring; MAGE = mean amplitude of glycemic excursion; T2DM = type 2 diabetes mellitus; VAS = visual analog scale.

4.2. Determination of Sample Size

A total of 48 patients will be randomized in a 1:1:1:1 ratio to 1 of 4 treatments. Assuming 1 dropout, the sample size of 11 patients for each treatment group would provide 79% power to demonstrate a statistically significant difference between the tirzepatide dose levels and dulaglutide 0.75 mg. This computation assumes the true treatment difference in glucose AUC(0-6h) is 330 mg*hr/dL, a common standard deviation of 250 mg*hr/dL.

5. A Priori Statistical Methods

5.1. General Considerations

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.

Exploratory efficacy parameters 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 at 32 weeks will be guided by the “efficacy” estimand using the efficacy analysis set (EAS).

5.2. Multiple Comparisons/Multiplicity

No multiplicity adjustments will be made for evaluating objectives in this analysis addendum.

5.3. Patient Disposition

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

5.4. Patient Characteristics

All demographic and baseline clinical characteristics will be summarized by study treatment for the patients in the mITT set, for those patients who participated in the protocol addendum.

5.5. Meal Tolerance Test

Exploratory objective relative to meal tolerance test is to compare the effect of once-weekly 5 mg, 10 mg and 15 mg tirzepatide versus dulaglutide 0.75 mg on the pharmacodynamic (PD) profile and satiety after a standardized test meal at Week 32 for the following endpoints:

- The change from baseline in serum glucose $AUC_{(0-6h)}$
- The change from baseline in insulin $AUC_{(0-6h)}$
- The change from baseline in C-peptide $AUC_{(0-6h)}$
- The change from baseline in glucagon $AUC_{(0-6h)}$
- The change from baseline in triglyceride $AUC_{(0-6h)}$
- The change from baseline in satiety (full and hungry) in VAS

PD analyses will be conducted using data from all patients who receive at least 1 dose of the investigational product and have evaluable PD parameter. The AUCs will be calculated using the linear-trapezoidal method and actual times.

The change in PD parameters from baseline to Week 32 will be analyzed by an analysis of covariance (ANCOVA) model with treatment, baseline HbA1c group ($\leq 8.5\%$, $> 8.5\%$), baseline BMI group (< 25 or ≥ 25 kg/m²) and washout of antidiabetic medication (yes or no) and baseline value (Week -1 [Day-7 of Week 0]) of the dependent variable as a covariate. The least squares

mean (LSM), standard error (SE) and the 95% confidence interval (CI) derived from the model will be displayed for change from baseline by treatment group. Treatment comparisons will be displayed showing the treatment difference LSM and the 95% CI of treatment differences along with the p-values for the treatment comparison.

Satiety after a standardized test meal (Satiety, Hunger and Satiety/Hunger) in the VAS will be summarized by treatment and time after a standardized test meal (0 [pre-meal], 60, 120, 180, 240, 300, and 360 minutes).

A comparison of tirzepatide doses and dulaglutide 0.75 mg will be conducted by ANCOVA model with treatment, baseline HbA1c group ($\leq 8.5\%$, $> 8.5\%$), baseline BMI group (< 25 or ≥ 25 kg/m²) and washout of antidiabetic medication (yes or no) and baseline value of the dependent variable as a covariate. The LSM, SE and the 95% CI derived from the model will be displayed for change from baseline by treatment group. Treatment comparisons will be displayed showing the treatment difference LSM and the 95% CI of treatment differences along with the p-values for the treatment comparison.

5.6. Continuous Glucose Monitoring

Exploratory objective relative to CGM is to compare the effect of each once-weekly 5 mg, 10 mg and 15 mg tirzepatide versus dulaglutide 0.75 mg on daily glucose control and variability measured by CGM at Week 32 for the following endpoints:

- Daily average glucose (mg/dL)
- Duration of time in range < 70 mg/dL (minutes)
- Percentage of time per day in range < 70 mg/dL
- Duration of time in range > 180 mg/dL (minutes)
- Percentage of time per day in range > 180 mg/dL
- Duration of time in range ≥ 70 mg/dL and ≤ 180 mg/dL (minutes)
- Percentage of time per day in range ≥ 70 mg/dL and ≤ 180 mg/dL
- Daily within-day SD (mg/dL)
- Daily mean amplitude of glycemic excursions (MAGE) (mg/dL)

The CGM system will be offered to a subgroup of patients at selected sites in the study [Protocol Addenda I8F-JE-GPGO (1)]. Interstitial glucose (ISIG) values will be collected by the CGM at 15-minute intervals, for a maximum of 96 measurements each day the sensor is worn.

All CGM variables will be derived for each patient and valid CGM day (unless otherwise specified). CGM parameters will then be calculated per visit as the aggregate over CGM days. Continuous glucose monitoring visits will be analyzed under the visit identified in the raw CGM data. All CGM glucose derivations will be calculated in units of mg/dL and then converted to mmol/L, by dividing by 18.

To ensure that CGM outcome variables are only calculated from CGM session days with sufficient data within the period, the following criteria will be used to determine a valid CGM day to be counted in the calculation for a visit (baseline or endpoint):

- Minimum number of measures per day – at least 70% of the total measures that are supposed to be obtained. For example, 70% of 96 measures for the 24-hour period.
- Maximum allowable missing interval – no interval greater than 3 hours between non-missing measures.
- At least 3 days (midnight to next midnight each day) of data during the CGM period, which may require the patient to wear the device for at least 5 days

For each occasion, based on the above criteria, the CGM system data collected on the first day (the date of first sensor insertion) and on the last day (between midnight and when patients remove the CGM) will not be used for the derivation and analysis of any CGM outcome variables. Only readings collected from valid CGM days before randomization will be included in the derivations for baseline. Only readings collected from valid CGM days while patients are on investigational product will be included in the derivations for post-baseline occasion at Week 32.

The percentage of time within a glucose range (target, hypo- or hyperglycemia ranges) will be calculated as the number of observations within the specified range divided by the total number of observations in the time interval (midnight to midnight 24-hour period). The duration (in minutes) within the glucose range will then be calculated as the percentage of time within the glucose range times the length of the period (eg, 1440 minutes in 24 hours).

The patient baseline characteristics and demographic variables (including but not limited to age, age group, gender, race, body mass index (BMI), duration of diabetes, baseline hemoglobin A1c (HbA1c), all stratification factors, etc.) will be obtained at entry, for those patients who participated in the CGM analyses will be summarized by treatment and overall.

The CGM data at 2 occasions (Week 0 and Week 32) will be summarized and analyzed. The CGM parameters listed above will be derived from glucose data collected by a CGM device and will be summarized by treatment and day for each occasion. The baseline daily CGM parameters will be derived as the average from several days of the data measured prior to administration of investigational product. The post baseline daily CGM parameters at Week 32 will be derived as the average of valid CGM days.

A comparison of tirzepatide doses and dulaglutide 0.75 mg relative to the daily CGM parameters at Week 32 will be conducted by a constrained longitudinal data analysis model (cLDA). Both baseline and post baseline CGM measures are considered as dependent variables, in conjunction with the constraint of a common baseline mean across the treatment groups (Liu et al. 2009). All patients who are enrolled in this addendum, randomized to 1 of the study treatments, received at least 1 dose of study drug, and have either baseline or post-baseline (Week 0 or Week 32) CGM data will be included in the analysis. The data collected after discontinuing study drug will be censored.

The following variables will be included as fixed effects: indicator variables of each treatment group at each post baseline time point, baseline HbA1c ($\leq 8.5\%$, $> 8.5\%$), baseline BMI group (< 25 or ≥ 25 kg/m²) and washout of antidiabetic medication (yes or no). More specifically, the indicator variables are defined as $\{I(\text{treatment}=i \text{ and time}=j) : i = 1,2,3,4 \text{ for treatment groups; } j = 0 \text{ (for baseline), } 1 \text{ (for post baseline)}\}$, which give a total of 8 indicator variables. An unstructured covariance matrix is used in this model to account for within subject correlation at all time points and can be different between treatment groups. If the analysis fails to converge, the following variance-covariance matrices will be used (in order) until it converges:

- Heterogeneous compound symmetry
- Compound symmetry

The Kenward-Roger approximation will be used to estimate denominator degrees of freedom for the mixed-effect model repeated measure (MMRM) models (constrained longitudinal data analysis [cLDA]).

With the aid of the cLDA analysis, 2-sided 95% confidence intervals (CI) for CGM parameter at the 32-week visit for the tirzepatide doses and the dulaglutide 0.75 mg group will be derived and the difference between each of the tirzepatide doses and the dulaglutide 0.75 mg group will also be shown.

If the cLDA model fails to converge, more simple cLDA model or ANCOVA model similar to Section 5.5 will be used.

Individual (by patient) or mean by treatment, glucose data will be plotted, if deemed appropriate.

5.7. Body Composition Analyses

Exploratory objective relative to body composition, measured by InBody 770 (Inbody Japan), is to compare the effect of once-weekly 5 mg, 10 mg and 15 mg tirzepatide versus dulaglutide 0.75 mg on body composition at Week 12, 32, and 52 for the following parameters:

- The change from baseline in total body water (L)
- The change from baseline in protein (kg)
- The change from baseline in minerals (kg)
- The change from baseline in body fat mass (kg)
- The change from baseline in lean body mass (kg)
- The change from baseline in muscle mass (kg)
- The change from baseline in body weight (kg)

All patients who are enrolled in this addendum, randomized to 1 of the study treatments, received at least 1 dose of study drug, and have baseline and at least 1 post-baseline data will be included in the analysis. The data collected after discontinuing study drug will be censored.

The data of total body water, protein, minerals, and body fat mass will be collected. Following 3 parameters are derived as below:

- Lean body mass (kg): the sum of total body water, protein, and minerals
- Muscle mass (kg): the sum of total body water and protein
- Body weight (kg): the sum of total body water, protein minerals and body fat mass

Baseline and change in body composition parameters from baseline will be summarized. A comparison of tirzepatide doses and dulaglutide 0.75 mg will be conducted by the MMRM model with terms: treatment, visit, and treatment-by-visit interaction, baseline HbA1c group ($\leq 8.5\%$, $> 8.5\%$) and washout of antidiabetic medication (yes or no) as fixed effects; and baseline values as a covariate. An unstructured covariance structure will model the relationship of 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 LSM of mean change from baseline will be plotted by visit and by study treatment. With the aid of the MMRM analysis, 2-sided 95% CI for mean change from baseline to 52-week visit for each of 5 mg, 10 mg and 15 mg tirzepatide versus dulaglutide 0.75 mg will be derived.

6. References

Liu GF, Lu K, Mogg R, Mallick M, Mehrotra DV. Should baseline be a covariate or dependent variable in analyses of change from baseline in clinical trials? *Stat Med*. 2009;28(20):2509–2530.

American Diabetes Association. 6. Glycemic Targets: Standards of Medical Care in Diabetes—2021. *Diabetes Care* 2021 Jan; 44(Supplement 1): S73-S84.

Leo Document ID = 7a0d600b-3c41-4a64-a797-e655834a66ef

Approver: PPD

Approval Date & Time: 07-Apr-2021 00:29:33 GMT

Signature meaning: Approved