Statistical Analysis Plan: I8F-MC-GPIU

A Randomized, Double-blind, Placebo-controlled Study to Investigate the Efficacy and Safety of Tirzepatide Monotherapy Compared with Placebo in Chinese Participants with Type 2 Diabetes (SURPASS-CN-MONO)

NCT05963022

Approval Date: 22-Oct-2024

Title Page

Protocol Title: A Randomized, Double-blind, Placebo-controlled Study to Investigate the Efficacy and Safety of Tirzepatide Monotherapy Compared with Placebo in Chinese Participants with Type 2 Diabetes (SURPASS-CN-MONO).

Protocol Number: I8F-MC-GPIU

Compound Number: LY3298176

Short Title: A Study to Investigate the Efficacy and Safety of Tirzepatide Monotherapy in Chinese Participants with Type 2 Diabetes.

Acronym: SURPASS-CN-MONO

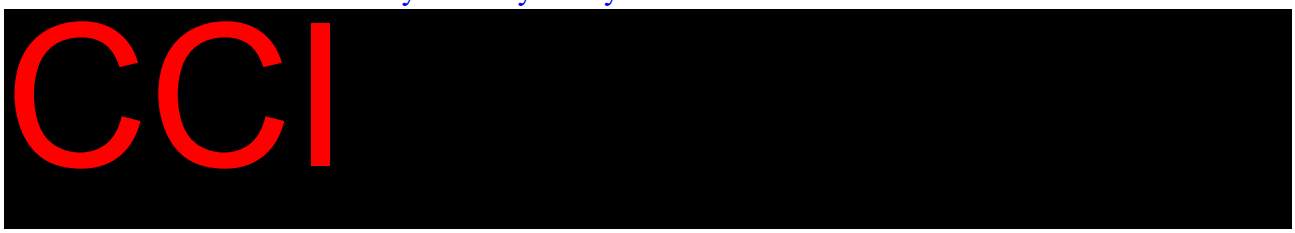
Sponsor Name: Eli Lilly and Company

Legal Registered Address: Lilly Corporate Center, Indianapolis, Indiana 46285, USA

Regulatory Agency Identifier Number(s): Not Applicable

Document ID: VV-CLIN-142838

Table of Contents



Version history

This Statistical Analysis Plan (SAP) for study I8F-MC-GPIU (GPIU) is based on the protocol amendment (b) dated 05 December 2023.

Table 1.1. SAP Version History Summary

SAP Version	Approval Date	Change	Rationale
1	31 January 2024	Not Applicable	Original version
2		**Section 4.1:** • Revise the analysis software as: "All statistical analyses will be conducted with SAS Version 9.4 or higher **and R Version 4.1.2 or higher** unless otherwise stated." • Add baseline definition for lab and patient reported outcome: "Baseline is defined as the last non-missing measurement recorded on or before the randomization visit, prior to first dose of treatment, unless otherwise specified. **For lab, 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.**"	• Added for clarification • Changed to minimize missing baseline data relevant to dosing.
		Section 4.8.3: • Removed baseline*visit interaction in MMRM	• The covariates were sufficient for covariate adjustment. Removing the baseline* visit

SAP Version	Approval Date	Change	Rationale
			interaction to avoid possible model convergence issue and keep consistent with all the SURPASS studies.
		Section 4.11.3: • Pregnancy and AESI added as notable events	• Add additional criteria for patient narrative
		Section 4.11.4.13: ● Onset plots added: "The time courses of prevalence, ~~and~~ incidence **and onset** (newly-occurring episodes) of nausea, vomiting, diarrhea, and combined will be plotted by treatment and maximum severity." • The following description added: "The gastrointestinal TEAE will also be summarized by participants with de-escalation or not. "	• Added onset plots and de-escalation related analysis for GIAE
		Section 4.11.7 added for product complaint	• Add description for product complaint analysis
		Section 4.12.1: • Removed treatment regimen estimand from subgroup analysis:	• Efficacy estimand is sufficient for subgroup

SAP Version	Approval Date	Change	Rationale
		"Efficacy subgroup analyses will be guided by the ~~treatment regimen estimand and~~ efficacy estimand. "	analysis to explore subgroup efficacy results.
		• Revised the wording: "Subgroup analyses of the change in HbA1c and the change in weight from baseline to Week 40 ~~will~~**may** be made to assess consistency of the intervention effect across the following subgroups"	• Flexible wording to reflect planned subgroup analysis
		• Duration of diabetes subgroup revised from "≤5 vs >5 to ≤10 years vs >10 years" to "≤5 vs >5 years"	• Due to the limited number of participants in ">5 to ≤10 years" and ">10 years", these two groups were combined
		• Subgroup analysis of treatment-emergent adverse event added	• Added subgroup analysis for TEAE

1. Introduction

Study GPIU (monotherapy) is a Phase 3, multicenter, randomized, double-blind, parallel, placebo-controlled, 40-week study that investigates the efficacy and safety of 3 doses of tirzepatide compared with placebo in Chinese participants with T2DM, inadequately controlled with diet and exercise alone, and have not been treated with any antihyperglycemic medication during the 90 days preceding the start of the study. This study aims to provide evidence of the tirzepatide as monotherapy for the treatment of Chinese patients with T2DM in early phase.

Approved on 22 Oct 2024 GMT

1.1. Objectives, Endpoints, and Estimands

Objectives	Endpoints
Primary	
• To demonstrate that tirzepatide QW 5 mg, and/or 10 mg, and/or 15 mg are superior to placebo QW in HbA1c change from baseline to Week 40	• Change in HbA1c from baseline
Key Secondary (controlled for type 1 error)	
• To demonstrate superiority of tirzepatide QW 5 mg, and/or 10 mg, and/or 15 mg to placebo QW at Week 40 for:	• Proportion of participants with HbA1c target values of <7.0% (<53 mmol/mol) • Change in fasting serum glucose (central laboratory) from baseline • Proportion of participants with HbA1c target values of ≤6.5% (≤48 mmol/mol) • Change in body weight from baseline • Proportion of participants with HbA1c target values of <5.7% (<39 mmol/mol)
Additional Secondary (not controlled for type 1 error)	
• To compare tirzepatide QW 5 mg, 10 mg, and 15 mg to placebo QW at Week 40 for:	• Change in daily average 7-point self-monitored blood glucose profiles from baseline • Proportion of participants who achieved weight loss of ≥5%, ≥10%, and ≥15%
Safety	
• To compare the safety of tirzepatide QW 5 mg, 10 mg, and 15 mg to placebo QW for:	• Incidence of treatment-emergent adverse events • Incidence of early discontinuation of study intervention due to AEs • Incidence of adjudicated pancreatic AEs • Change in serum calcitonin from baseline • Incidence of allergic and hypersensitivity reactions • Change in systolic blood pressure, diastolic blood pressure and pulse rate from baseline • Occurrence of hypoglycemic episodes

Objectives	Endpoints
	• Time to initiation of rescue therapy for severe persistent hyperglycemia

Primary estimand

The primary efficacy assessment is guided by the "treatment-regimen" estimand. This estimand aims at reflecting how participants with T2DM are treated in clinical practice and takes into account both tolerability and efficacy.

The treatment-regimen estimand answers the following question of interest for the primary objective: What is the treatment difference in change from baseline in HbA1c after 40 weeks of treatment in participants with T2DM regardless of treatment discontinuation for any reason and regardless of initiation of antihyperglycemic rescue?

The treatment-regimen estimand is described by the following attributes.

- *Population:* Participants with T2DM who are inadequately controlled on diet and exercise alone, and have not been treated with any antihyperglycemic medication during the 90 days preceding the start of the study.
- *Endpoint:* Change from baseline to week 40 in HbA1c.
- *Treatment condition*: The randomized treatment regardless of adherence to treatment with or without antihyperglycemic rescue medication. Further details on study treatment and rescue therapy can be found in Study Protocol Section 6.

- *Intercurrent events:* no ICEs are defined since treatment adherence and the initiation of additional antihyperglycemic medications are a part of the treatment condition.
- *Population-level summary*: Difference in mean changes between treatment conditions.

Secondary estimand

A robustness of the primary efficacy assessment, guided by the "efficacy" estimand, will be conducted. This estimand focuses on the treatment effect if participants who underwent randomization continued to receive the study treatment without treatment discontinuation or rescue antihyperglycemic medication for severe, persistent hyperglycemia.

The efficacy estimand answers the following question of interest for the primary objective: What is the treatment difference in HbA1c change from baseline to week 40 after randomization in participants with T2DM assuming that participants had stayed on treatment and not taken antihyperglycemic rescue medication?

The efficacy estimand is described by the following attributes.

- *Population*: Participants with T2DM who are inadequately controlled on diet and exercise alone, and have not been treated with any antihyperglycemic medication during the 90 days preceding the start of the study.
- *Endpoint*: Change from baseline to Week 40 in HbA1c.
- Treatment condition: The randomized treatment. Further details on study treatment and rescue therapy can be found in Study Protocol Section 6.
- *Intercurrent event*: Intercurrent events of interest: "Treatment discontinuation for any reason" and "Initiation of antihyperglycemic rescue treatment" will be addressed using the following strategies:
    - Had participants stayed on treatment (hypothetical strategy).
    - Had participants not taken antihyperglycemic rescue medication (hypothetical strategy).
- *Population-level summary*: Difference in mean changes between treatment conditions.

The efficacy estimand will be evaluated for all primary, secondary, and tertiary efficacy objectives. The treatment-regimen estimand will be evaluated for the primary and all key secondary objectives. The population, treatment condition, and intercurrent events specified above for each estimand for the primary objective will also apply to the other efficacy objectives, and the population-level summary might be adjusted according to the endpoints of the objectives.

1.2. Study Design

Study GPIU is a multicenter, randomized, double-blind, parallel, placebo-controlled, Phase 3 study which will investigate the efficacy and safety of tirzepatide monotherapy compared with placebo in Chinese participants with T2DM, inadequately controlled with diet and exercise alone, and have not been treated with any antihyperglycemic medication during the 90 days preceding the start of the study.

Approximately, 200 participants will be randomized in a CCI ratio (tirzepatide 5 mg QW, 10 mg QW, 15 mg QW, and placebo QW). Participants will be stratified based on HbA1c (≤8.5% [≤69 mmol/mol] or >8.5% [>69 mmol/mol]), and any past prior use (Yes or No) of any antihyperglycemic medication.

Study GPIU will consist of 3 periods:

- Period I: approximately CCI screening/lead-in period
- Period II: CCI treatment period, and
- Period III: CCI safety follow-up period.

Study Period I (Screening and Lead-in)

Screening (Visit 1)

The purpose of screening procedures at Visit 1 is to establish initial eligibility and to obtain blood samples for laboratory assessments needed to confirm eligibility at Visit 2. The participant will sign the informed consent form (ICF) before any study procedures are performed. Procedures at this visit will be performed as shown in the Study Protocol Schedule of Assessment (Section 1.3).

Lead-in (Visit 2 to Visit 3)

At Visit 2, the screening laboratory results will be reviewed. For those participants meeting all other eligibility requirements, a dilated fundoscopic examination, performed by an ophthalmologist or optometrist, will be completed between Visit 2 and Visit 3 to confirm eligibility unless a dilated fundoscopic examination was performed by an ophthalmologist or optometrist within 90 days of study entry (Visit 1).

Additionally, at Visit 2, participants and their caregiver(s), if applicable, will receive a glucometer and training on how to perform SMBG. Participants will be asked to monitor fasting blood glucose (FBG) 4 times a week and two 7-point SMBG profiles done on 2 nonconsecutive days in the 2-week period prior to Visit 3 (randomization), Visit 9 (Week 24), Visit 11 (Week 40).

Participants will be provided diaries and will be trained as appropriate to record BG values and hypoglycemic events. During this period, participants will also be trained on disease management (symptoms of hypoglycemia and hyperglycemia) and study procedures; this training can be repeated at subsequent visits as deemed appropriate. Participants should monitor BG any time a hypoglycemic event is suspected.

Study Period II (CCI treatment period)

Randomization (Visit 3)

At Visit 3, eligible participants will perform all required baseline study procedures (including the collection of all baseline laboratory measures and electrocardiogram [ECG]) prior to randomization and prior to administering the first dose of study intervention. Participant should arrive to the clinic in the fasting state; the fasting state should have lasted at least 8 hours (water is allowed). The questionnaires (European Quality of Life-Dimensions [EQ-5D-5L] and Impact of Weight on Quality of Life-Lite Clinical Trials Version [IWQOL-Lite-CT]) should be completed before any other study procedures if the participant is not adversely affected by the fasting condition or completed after the participant has sufficiently recovered from the preceding visit procedures.

Following randomization, participants will participate in a CCI treatment period.

Postrandomization period (end of Visit 3 to Visit 11):

The starting dose of study intervention will be 2.5 mg QW for 4 weeks, followed by an increase to 5 mg QW, for the duration of the study in the 5-mg group.

For the 10-mg group, the starting dose of study intervention will be 2.5 mg QW for 4 weeks, then the dose will be increased by 2.5 mg every 4 weeks (2.5 mg to 5 mg to 7.5 mg to 10 mg) until the 10-mg dose is reached and maintained for the duration of the study.

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

For the placebo group, participants will inject matched QW placebo for the duration of the study.

Visit 99

Visit 99 is only applicable to participants who discontinue the study treatment prematurely before Week 40 and decline to complete the remaining study visits but are willing to return for Visit 99 at Week 40 after randomization. This visit is critical to ensure complete data collection for the primary and key secondary endpoints.

Participants should attend this visit in the fasting state. Procedures to be completed are:

- measurement of HbA1c, FSG and weight
- listing of concomitant medications, and
- assessment of AEs.

For participants unwilling to attend this visit, their refusal to attend should be documented in the participant medical record.

Study Period III (safety follow-up period)

Safety follow-up (Visit 801) visits:

All participants who complete the treatment period (complete Visit 11 or Visit 99) are required to complete Visit 801, a safety follow-up visit approximately 4 weeks after their last treatment visit. Participants 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, participants will not receive study intervention. Participants may be treated with another glucose-lowering intervention decided upon by the investigator (except for prohibited antihyperglycemic medications, see Study Protocol section 6.9.1). If any new antihyperglycemic medication is initiated during the safety follow-up period, it will not be classified as rescue therapy.

Participants are also required to return any remaining study diaries to the study site at the end of this period.

The participants who complete the treatment period (complete Visit 11 or Visit 99) and safety follow-up (Visit 801) are considered as complete the study.

2. Statistical Hypotheses

The alternative hypotheses for the primary and key secondary objective are as follows,

- $H_{5,1}$, $H_{10,1}$, and $H_{15,1}$: Superiority test of tirzepatide 5 mg, 10 mg, 15 mg versus placebo in HbA1c change from baseline at Week 40 respectively.
- $H_{5,2}$, $H_{10,2}$, and $H_{15,2}$: Superiority test of tirzepatide 5 mg, 10 mg, 15 mg versus placebo in proportion of participants achieve HbA1c <7% at Week 40 respectively.
- $H_{5,3}$, $H_{10,3}$, and $H_{15,3}$: Superiority test of tirzepatide 5 mg, 10 mg, 15 mg versus placebo in FSG change from baseline at Week 40 respectively.
- $H_{5,4}$, $H_{10,4}$, and $H_{15,4}$: Superiority test of tirzepatide 5 mg, 10 mg, 15 mg versus placebo in proportion of participants achieve HbA1c ≤6.5% at Week 40 respectively.
- $H_{5,5}$, $H_{10,5}$, and $H_{15,5}$: Superiority test of tirzepatide 5 mg, 10 mg, 15 mg versus placebo in body weight change from baseline at Week 40 respectively.
- $H_{5,6}$, $H_{10,6}$, and $H_{15,6}$: Superiority test of tirzepatide 5 mg, 10 mg, 15 mg versus placebo in proportion of participants achieve HbA1c <5.7% at Week 40 respectively.

2.1. Multiplicity Adjustment

The primary and the key secondary objectives will be evaluated align to both the treatment-regimen estimand and the efficacy estimand. No multiplicity adjustment will be made for conducting analysis aligned to the two estimands.

For analysis within each estimand, $H_{5,1}$, $H_{10,1}$, and $H_{15,1}$ will be initially tested each at 0.01667 significance level. A graphical testing scheme (Bretz et al. 2009, 2011) presented in Figure GPIU.2.1 will be used to strongly control for type 1 error.

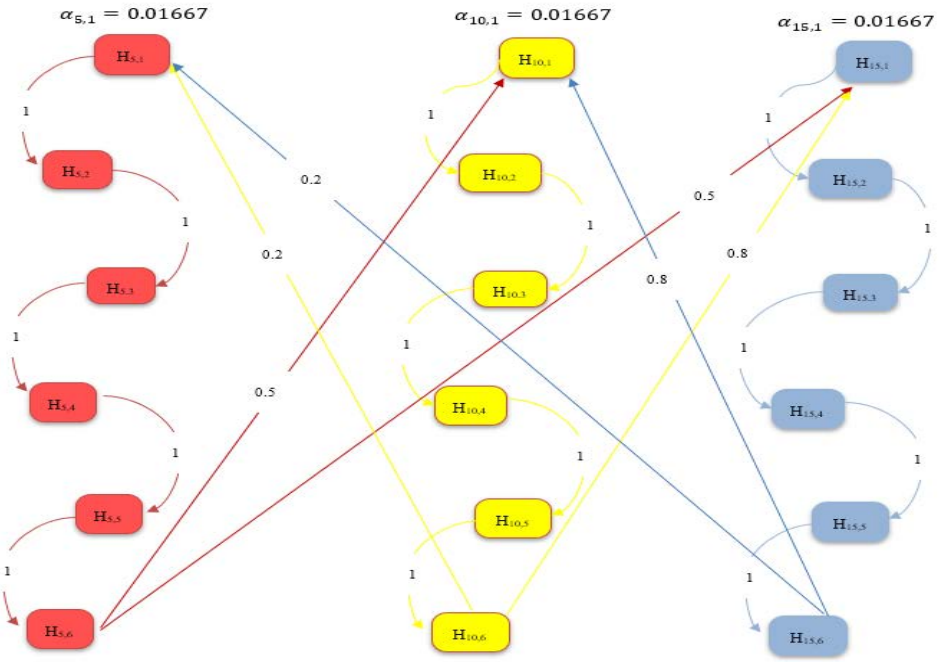


Figure GPIU.2.1. Type 1 error control strategy for primary and key secondary.

3. Analysis Sets

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

Population/Analysis Set	**Population/Analysis Set**
Screened population	All participants who signed informed consent
Randomized population	All participants who are randomly assigned to a treatment group.
Modified intention-to-treat (mITT) population	All randomly assigned participants who are exposed to at least 1 dose of study intervention.
Efficacy analysis set (EAS): This analysis set will be used to estimate the efficacy estimand for all efficacy objectives	Data obtained during Study Period II from the mITT population, excluding data after initiating rescue antihyperglycemic medication or stopping study intervention.
Full analysis set (FAS): This analysis set will be used to estimate the treatment-regimen estimand for the primary and key secondary objectives	Data obtained during Study Period II from the mITT population, regardless of adherence to study intervention or initiation of rescue antihyperglycemic medication.
Safety analysis set (SS): This analysis set will be used to assess the safety of study treatment	Data obtained during Study Periods II or III from the mITT population, regardless of adherence to study intervention or initiation of rescue antihyperglycemic medication.

4. Statistical Analyses

4.1. General Considerations

This information will be copied directly from the protocol. Additional general considerations may be added. 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 and R Version 4.1.2 or higher unless otherwise stated. Any change to the data analysis methods described in the protocol will require an amendment ONLY if it changes a principal feature of the protocol. Any other changes to the data analysis methods described in the protocol, and the justification for making the change, will be described in the SAP or the clinical study report (CSR). Some analyses and summaries described in this analysis plan may not be conducted if not warranted by data (eg, few events to justify conducting an analysis). Listing of events will be provided in such situations. Additional analyses of the data may be conducted as deemed appropriate without further changes made to the protocol or SAP, even after the primary or final database locks (DBL).

Unless otherwise noted, all tests of treatment effects will be conducted at a 2-sided alpha level of 0.05, and the confidence interval will be calculated at 95%, 2-sided. In statistical summaries and analyses, all data will be analyzed by randomized treatment assignment. Participants will be analyzed according to the treatment they were randomly assigned to, regardless of the treatment actually received.

Baseline is defined as the last non-missing measurement recorded on or before the randomization visit, prior to first dose of treatment, unless otherwise specified. For lab, 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 “treatment regimen” estimand, represents the efficacy irrespective of adherence to investigational product or introduction of rescue therapy for persistent severe hyperglycemia. Analysis relative to “treatment-regimen” estimand will be conducted using full analysis set. Second estimand, the “efficacy” estimand, represents efficacy prior to discontinuation of study drug or initiating rescue therapy for persistent severe hyperglycemia. Analysis relative to “efficacy” estimand will be conducted using efficacy analysis set. Safety will be assessed using safety analysis set. Selected safety analyses may be conducted after excluding data on rescue therapy or data after starting another antihyperglycemic medication.

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 database after the last date of study participation (date of death, date of early termination, or date of safety follow-up) will be excluded from statistical analysis. Listing of such data may be provided.

Summary statistics for continuous measures will include sample size, mean, standard deviation (SD), median, minimum, and maximum. 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. The negative binomial regression model will be used for the treatment comparison of discrete count measures if deemed appropriate.

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

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

4.2. 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 describe in Section 4.8.2. 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 unless otherwise specified.

4.3. Participant Dispositions

Frequency, counts and percentages of all participants screened, randomized, and receiving at least 1 dose of study intervention will be presented by treatment groups. Reasons for screen failure as reported by investigators will be summarized. A listing of randomized participants not receiving study intervention will be provided. Of the participants in the randomized population, frequency, counts and percentages of participants completing the study, prematurely discontinuing study, including the reason for premature discontinuation; and prematurely discontinuing study intervention, including the reason for premature discontinuation of study intervention will be presented by treatment group. A Kaplan-Meier analysis of time from randomization to premature discontinuation from study and premature discontinuation from study intervention by treatment group will be provided.

4.4. Patient Characteristics

All demographic and baseline clinical characteristics will be summarized by study treatment for all randomized patients. Baseline demographic and clinical characteristics of special interest include and not limit to: age, gender, weight, HbA1c, fasting serum glucose, duration of T2DM, eGFR.

Listing of patient demographics will be provided for all randomized patients.

4.5. Medical History and Preexisting Conditions

Medical history and pre-existing conditions will be summarized using the mITT population. Specific medical history regarding gallbladder diseases, pancreas, ketosis and diabetes complications may be summarized.

Listing of medical history and pre-existing conditions will be provided for mITT population.

4.6. Concomitant Therapy

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. Concomitant medications will be summarized using the safety analysis set in mITT population. The prespecified concomitant medications of interest may be summarized by the planned treatment.

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

- baseline antihypertensive therapy, by type
- postbaseline antihypertensive therapy in Study Period II, by type
- baseline lipid lowering therapy, by type
- postbaseline lipid lowering therapy in Study Period II, by type
- prior use of antihyperglycemic therapy
- utilization of other antihyperglycemic therapy in Study Period II, and also in Study Period II and III
- rescue therapy due to severe persistent hyperglycemia
- initiation of following medications in Study Period II:
    - antidiarrheal medication
    - antiemetic medication

4.7. Important Protocol Deviation

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

4.8. Primary Endpoint(s) Analysis

4.8.1. Definition of Endpoint(s)

The primary endpoint for this study is change in HbA1c (% and mmol/mol) from baseline at Week 40. This endpoint will be used to evaluate the primary objective of the study for both the treatment-regimen and efficacy estimands.

Both HbA1c values as well as change from baseline in HbA1c will be summarized by treatment and nominal visit (week). 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.

4.8.2. Main Analytical Approach

The main analysis of primary endpoint will be based on treatment-regimen estimand. The analysis will be conducted utilizing HbA1c data in FAS at baseline and at the 40-week visit with the aid of an analysis of covariance (ANCOVA). The response variable will be the change in HbA1c from baseline at Week 40 and model terms will include treatment, any past prior use (yes or no) of any antihyperglycemic medication, and baseline HbA1c as covariates. Missing HbA1c values at Week 40 will be imputed using multiple imputation. Statistical inference over multiple imputations will be guided by the method proposed by (Rubin 1987).

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 drug, or used rescue therapy for persistent severe hyperglycemia. In cases where there are not enough retrieved dropouts to provide a reliable imputation model, an alternative multiple imputation method with reference to the placebo group (placebo 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.

With the aid of the ANCOVA analysis, p-values and 2-sided 95% confidence interval (CI) for mean change in HbA1c from baseline through 40-week visit for 5 mg, 10 mg, and 15 mg tirzepatide compared to placebo will be derived and summarized.

4.8.3. Supplementary Analyses

The analysis will be conducted utilizing HbA1c data in 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 Kenward-Roger option will be used to estimate denominator degrees of freedom. The response variable of the MMRM model will be the change in HbA1c from baseline at Week 40 and model terms will include treatment, visit, treatment by visit interaction, any past prior use (yes or no) of any antihyperglycemic medication, baseline HbA1c. Missing data will be addressed by the MMRM model under missing at random assumption. No explicit imputation methods for missing data will be employed.

An unstructured covariance structure will be used to model the within-patient errors. If this model fails to converge, the following covariance structures will be tested in order:

- Heterogeneous Toeplitz
- Heterogeneous First Order Autoregressive
- Heterogeneous Compound Symmetry
- Toeplitz
- First Order Autoregressive
- Compound Symmetry.

The first covariance structure that converges will be used. The resulting least squares mean (LSM) estimate of mean change from baseline in HbA1c will be summarized 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 through 40-week visit for 5 mg, 10 mg, and 15 mg tirzepatide compared to placebo will be derived and summarized.

4.9. Secondary Endpoints Analysis

4.9.1. Key Secondary Endpoints (Controlled for Type 1 Error)

The secondary study objectives subject to type 1 error rate control are as follows:

- Proportion of participants with HbA1c target values of <7.0% (<53 mmol/mol)

- Change in fasting serum glucose from baseline
- Proportion of participants with HbA1c target values of ≤6.5% (≤48 mmol/mol)
- Change in body weight from baseline
- Proportion of participants with HbA1c target values of <5.7% (<39 mmol/mol)

Key secondary endpoints will be evaluated based on the efficacy and treatment-regiment estimands (Section 3), similar to the primary endpoint.

4.9.1.1. Proportion of Participants with HbA1c Target Values of <7.0%, ≤6.5% and <5.7% at Week 40

Analysis relative to treatment-regimen estimand will be conducted by utilizing HbA1c data in 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. The imputation method will be the same as the analysis of primary endpoint (see Section 4.8.2 for details). Model terms will include treatment, any past prior use (yes or no) of any antihyperglycemic medication, and baseline HbA1c as covariates. Statistical inference over multiple imputations will be guided by (Rubin 1987).

The analysis relative to the efficacy estimand for the endpoints at 40 weeks will be performed using EAS with missing values imputed from an MMRM model and then dichotomized. The MMRM model includes treatment, visit, treatment-by-visit interaction, and any past prior use (yes or no) of any antihyperglycemic medication, baseline HbA1c and baseline HbA1c-by-visit interaction. After dichotomizing continuous HbA1c, the data is analyzed using a logistic regression model with treatment, any past prior use (yes or no) of any antihyperglycemic medication, and baseline HbA1c as covariates.

4.9.1.2. Change in Fasting Serum Glucose and Body Weight From Baseline at the 40-Week Visit

The analysis for change in body weight and fasting serum glucose from baseline will be conducted in a manner similar to the primary analysis in Section 4.6. 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 and fasting serum glucose from baseline will be summarized by nominal visit and by study treatment. 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; if value of the imputed fasting serum glucose change from baseline is <-20 mmol/L or > 20 mmol/L, that value will be set to -20 mmol/L or 20 mmol/L, respectively, to avoid unrealistic imputed values.

4.9.2. Additional Secondary Efficacy Analysis

Other secondary efficacy measures will be summarized by treatment and nominal visit. Unless otherwise specified, missing data will not be imputed and assessments are not subject to type 1 error rate control. The analyses will be conducted based on the efficacy estimand.

4.9.2.1. Analysis of 7-point SMBG Profile

The 7-point SMBG profiles consist of pre-meal and 2-hour postprandial SMBG measurements for the morning, midday, and evening meals, and one measurement at bedtime. Patients will be asked to perform two 7-point SMBG profiles over a 24-hour period on 2 nonconsecutive days, in the 2-week period prior to prespecified office visits (Visits 3 [Week 0], 9 [Week 24], and 11 [Week 40]). In general, the mean of the two daily values will be used for reporting purpose. If more than 2 SMBG profiles are available, the 2 most recent profiles on nonconsecutive days should be used.

The following variables for 7-point SMBG profile will be analyzed using MMRM model:

1. Pre morning meal BG
2. 2-hour postprandial measurement for morning meal BG
3. Pre midday meal BG
4. 2-hour postprandial measurement for midday meal BG
5. Pre evening meal BG
6. 2-hour postprandial measurement for evening meals BG
7. Bedtime BG
8. Morning pre-meal to 2-hour postprandial BG excursion
9. Midday pre-meal to 2-hour postprandial BG excursion
10. Evening pre-meal to 2-hour postprandial BG excursion
11. Mean of all pre-meal BG
12. Mean of all 2-hour postprandial BG
13. Mean of all meals 2-hour excursion
14. Mean of all 7-point BG

For before and 2 hours after meal BG and bedtime BG (Reports 1-7), the mean BG value at each visit is calculated as the average of 2 daily BG values. The glucose excursion for morning, midday and evening meals is calculated as the difference between 2-hour postprandial measurements and pre-meal measurements.

For mean of all pre-meal BG (Report 11), the pre-meal daily mean is calculated as the average BG values collected for before morning, midday and evening meals on a particular day. The mean of all pre-meal BG at each visit is calculated as the average of 2 pre-meal daily means.

For mean of all postprandial BG (Report 12), the post-meal daily mean is calculated as the average of 2-hour postprandial BG values of morning, midday and evening meals on a particular day. The mean of all postprandial meals BG at each visit is calculated as the average of 2 post-meal daily means.

For mean of all meals 2-hour excursion at each visit (Report 13), the daily mean for all meals is calculated as the average of glucose excursion for morning, midday and evening meals on a

particular day. The mean of all meals 2-hour excursion at each visit is calculated as the average of 2 daily means.

For mean of all 7-point BG (Report 14), the daily mean is calculated as the average of 7 BG values collected on a particular day. The mean of all 7-point BG at each visit is calculated as the average of 2 daily means.

Actual and change from baseline of 7-point SMBG will be summarized using mITT population, and analyzed using MMRM with treatment, any past prior use (yes or no) of any antihyperglycemic medication, visit, baseline HbA1c strata (<8.5%, and ≥8.5%) and treatment-by-visit interaction, baseline 7-point SMBG, baseline 7-point SMBG-by-visit interaction as covariates. An unstructured covariance structure will be used to model the within-patient variability and implicitly adjusts for missing data. If this analysis fails to converge, then other covariance structures will be tested in the order mentioned in Section 4.8.3.

4.9.2.2. Proportion of Participants with Weight Loss of ≥5.0%, ≥10% and ≥15% at Week 40

The analysis for the proportion of participants with target weight loss at Week 40 will be conducted in a manner similar to Section 4.9.1.1 based on the efficacy estimand. 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.

4.11. Safety Analyses

Unless specified otherwise, safety assessments will be based on the SS (see Section 3). All events that occur between the date of first dose of study drug to the date of patient's safety follow-up visit or patient's end of study participation will be included. Selected safety analyses may be conducted after excluding data after the introduction of another antihyperglycemic therapy. For rare events (<10 patients have the events), summary tables may not be generated, and individual patient level data will be listed.

Unless specified otherwise, comparisons of tirzepatide doses to placebo will be performed. For selected continuous safety parameters, difference among treatment mean change from baseline at scheduled visits will be assessed via a MMRM using REML. The model will include baseline HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]), any past prior use (yes or no) of any antihyperglycemic medication, treatment group, visit, treatment-by-visit interaction, baseline

value of the safety parameter and baseline of the safety parameter-by-visit interaction. 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 4.8.3 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 an event, "time-to-first-event" will be the time to the first occurrence of the event. For patients without the event, "time-to-event" will be censored at end of study participation (study discontinuation, safety follow-up visit, or date of death).

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 specified time interval between patients.

For safety interpretation, p-values will not be used for hypothesis testing. P-values will be considered as a number between 0 and 1 that gives an idea of how strong the evidence is for an imbalance between study intervention arms. The evidence for an imbalance is stronger toward 0 and weaker toward 1. Similarly, confidence intervals will not be used for hypothesis testing. They reflect the uncertainty of an estimate.

The definition of baseline and postbaseline are provided in Table GPIU.4.1.

Table GPIU.4.1. Baseline and Postbaseline Definition for Safety Analysis

Analysis Set	Analysis Type	Baseline	Postbaseline
SS	1.1) Treatment-Emergent Adverse Events	The baseline period is defined as the start of screening and ends prior to the first dose of study treatment (typically at Week 0). If the first dose date is missing then the randomization date will be used instead of first dose date.	Starts at or after the first dose of study treatment and ends at the end of the study period (including off-drug follow up visit).
SS	1.2) Treatment-Emergent Abnormal Labs and Vital Signs.	For labs, baseline period is defined as prior to the first dose time and will include all scheduled and unscheduled measurements. If the first dose time is missing then any data collected on the date of the first dose will be treated as baseline. For vital signs, baseline period is defined as measurements collected prior to the first dose. If the first dose date is missing then the randomization date will be used instead of first dose date.	Postbaseline will be defined as after the baseline period through the end of the study participation. All scheduled and unscheduled measurements will be included.
SS	1.3) Change from Baseline for Labs, and Vital Signs.	The last scheduled and unscheduled non-missing assessment recorded during the baseline period defined above (1.2).	Postbaseline will be defined as above (1.2). Only scheduled visits will be included.

Abbreviations: ED = early discontinuation; SS = Safety Analysis Set.

4.11.1. Extent of Exposure and Treatment Compliance

For the summary of duration on study treatment and study, the frequency and percentage of participants falling into the following categorical ranges will also be summarized by planned treatment group as well: >0 week, ≥4 weeks, ≥8 weeks, ≥12 weeks, ≥16 weeks, ≥20 weeks, ≥24 weeks, ≥38 weeks.

Treatment compliance is defined as taking at least 75% of required injections of study intervention. 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.Frequency, counts and percentages of participants compliant to study intervention will be summarized by treatment arm using the mITT population. Overall study drug compliance will be evaluated for participants administered >4 planned doses.

A listing of patients who re-initiate due to missing ≥3 consecutive doses may be produced.

4.11.2. Adverse Events

For adverse events (AEs), the most recent Medical Dictionary for Regulatory Activities (MedDRA) version that accommodates the data lock timeline is to be applied. Adverse events will be coded from the actual term using the 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.

Treatment-emergent 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. Statistical comparisons will be applied at both the SOC and PT levels. Events will be ordered by decreasing frequency within SOC, with the SOC in alphabetical order. For events that are sex-specific, the denominator and computation of the percentage will include only patients from the given sex. The planned summaries for AEs are provided in Table GPIU.4.2.

Table GPIU.4.2. Tables and Figures Related to Adverse Events

Analysis	Details	Sample
Overview of AEs	Number and percentage of participants who reported • death • SAE • permanent discontinuation from study intervention due to an AE • permanent discontinuation from study due to an AE, and • TEAE.	Controlled

Analysis	Details	Sample
TEAEs by PT within SOC	Number and percentage of participants with TEAEs using MedDRA PT nested within SOC. For controlled analysis sets, events will be ordered by decreasing frequency (total column) within SOC. SOCs will be ordered by alphabetical order.	Controlled
Maximum Severity TEAEs by PT	The number and percentage of participants with TEAEs by maximum severity using MedDRA PT. For each participant 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. Events will be ordered by decreasing frequency (total column) within SOC. SOCs will be ordered by alphabetical order.	Controlled
TEAEs Occurring at ≥5% by PT	Presented as a figure using study size adjusted percentages and risk differences (95% confidence intervals). The number and percentage of participants with TEAEs using MedDRA PT for the TEAEs occurring in ≥5% before rounding in any column in the table. The cut off is determined based on the percentage not adjusted for study size. Events will be ordered by decreasing frequency (total column) within SOC. SOCs will be ordered by alphabetical order.	Controlled
SAEs by PT within SOC	The number and percentage of participants who reported a SAE (including deaths and SAEs temporally associated or preceding deaths) during the treatment period using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency (total column) within SOC. SOCs will be ordered by alphabetical order.	Controlled

Approved on 22 Oct 2024 GMT

Analysis	Details	Sample
Primary AE leading to permanent discontinuation of study intervention by PT within SOC	The number and percentage of participants who permanently discontinued from study intervention due to an AE (including AEs that led to death) during the treatment period using MedDRA PT nested within SOC. Time-to-event analyses will be conducted by treatment on time to study drug discontinuation. Primary AE listed on the disposition form will be used. Events will be ordered by decreasing frequency (total column) within SOC. SOCs will be ordered by alphabetical order.	Controlled
Primary AE leading to permanent discontinuation of study by PT within SOC	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 (total column) within SOC. SOCs will be ordered by alphabetical order.	
Listing of SAEs	A listing of all SAEs will be provided. Listing will include and not limit to treatment, patient identification including the site number, treatment group, date of event, age at the time of enrollment, gender, MedDRA SOC and PT, whether it is TEAE, severity, action taken, outcome, relationship to study drug, time from first dose of study drug to the event, and event duration. A listing of all deaths will be provided. 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, caused of death as adjudicated by Clinical Endpoint Committee (CEC).	mITT Population

Abbreviations: AE = adverse event; HLT = High Level Term; LLT = Lowest Level Term; LY = Lilly product number; MedDRA = Medical Dictionary for Drug Regulatory Activities; PT = Preferred Term; SAE = serious adverse event; SOC = System Organ Class; TEAE = treatment-emergent adverse event.

4.11.3. Narratives

The following are “notable” events, from start of study intervention through end of study participation (or data cutoff for the submission if earlier):

- deaths
- SAEs
- Severe AESI (referred to 6.1 [Appendix 1])

- pregnancy, and
- permanent discontinuations of study intervention due to AEs.

Narratives (patient-level data and summary paragraph) will be provided for participants in the mITT population with at least 1 notable event.

Safety topics of interest are not considered notable events, unless 1 of the above criteria is met. Displays with individual participant-level data will be created for safety topics of interest using various formats such as a customized listing and/or a customized graphical patient profile as specified in the section associated with the safety topic of interest. Medical case summaries or vignettes will be provided if deemed relevant for the discussion of the safety topic of interest.

4.11.4. Additional Safety Assessments

4.11.4.1. Hypoglycemic Events

Definitions of different categories of hypoglycemic events are included in Table GPIU.4.3.

Table GPIU.4.3. Definitions of Hypoglycemic Event Categories

Glycemic Criteria/Description	
Level 1 (Glucose Alert Value)	Blood Glucose <70 mg/dL (3.9 mmol/L) and >= 54 mg/d (3.0 mmol/L)
Level 2 (Clinically Significant Hypoglycemia)	Glucose <54 mg/dL (3.0 mmol/L)
Level 3 (Severe Hypoglycemia)	A severe event characterized by altered mental and/or physical status requiring assistance for recovery (no specific glucose threshold).

Nocturnal hypoglycemia: a hypoglycemia event (including severe hypoglycemia) that occurs between bedtime and waking.

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

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.

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 hypoglycemic events will also be provided. Safety analyses might be conducted after including data after the introduction of another antihyperglycemic therapy.

The incidence of hypoglycemic event will be analyzed using logistic regression with treatment and stratification factors 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 stratification factors and treatment as fixed effects. When the number of hypoglycemic events is less 10, a listing of hypoglycemic events will be provided instead.

4.11.4.2. 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", event time will be censored at end of treatment period. A listing of patients who initiated rescue therapy will be provided.

4.11.4.3. 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 predefined SMQ search for acute pancreatitis and MedDRA PT of pancreatitis chronic. Detailed searching criteria can be found in Section 6.1, Appendix 1. Treatment-emergent adjudication-confirmed pancreatitis will be considered as AESI.

4.11.4.3.1. Pancreatic Enzyme Assessment

Observed pancreatic enzyme data (p-amylase and lipase) will be summarized by treatment and nominal visit. The number and proportion of patients with 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 (postbaseline measure/baseline measure) response variable and stratification factors, treatment, nominal visit, and treatment-by-nominal visit interaction as fixed effects, and baseline value as a covariate.

4.11.4.4. Thyroid Malignancies and C-Cell Hyperplasia

Treatment-emergent thyroid malignancies and C-cell hyperplasia, will be identified using predefined MedDRA High Level Terms (HLTs) of thyroid neoplasms malignant, and PT of thyroid C-cell hyperplasia. Detailed searching criteria can be found in Section 6.1, 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.

4.11.4.5. Malignancy

The AE database will be searched using pre-defined SMQs to identify events consistent with malignancy. Detailed searching criteria can be found in Section 6.1, Appendix 1. A summary by treatment and PT within SMQ and a listing of TEAEs will be provided. Malignancy will be considered as AESI.

4.11.4.6. 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.

4.11.4.7. 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).

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

4.11.4.8. Arrhythmias and Cardiac Conduction Disorders

The AE database will be searched using pre-defined SMQ or MedDRA HLT to identify events consistent with supraventricular arrhythmias and cardiac conduction disorders. Detailed searching criteria can be found in Section 6.1, 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.

4.11.4.9. Hypersensitivity Events

Hypersensitivity reactions and related information reported via the “Vital Signs: Hypersensitivity Event” electronic case report form (eCRF) will be summarized by treatment. Two main analyses are performed:

- Potential Immediate Hypersensitivity: Analysis of TEAEs occurring from the start of study drug administration up to 24 hours after the end of study drug administration. For events withonly date (no time) information 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. 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 Section 6.1, Appendix 1. Severe/serious hypersensitivity events identified by predefined SMQ search will be considered as AESIs.

4.11.4.10. Injection Site Reaction

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 Section 6.1, 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.

4.11.4.11. 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 Section 6.1, Appendix 1 will be considered as AESI and summarized.

4.11.4.12. Hepatobiliary Safety

4.11.4.12.1. Hepatobiliary Disorders

The AE database will be searched using SMQs to identify events consistent with hepatobiliary disorders. Detailed searching criteria can be found in Section 6.1, Appendix 1. A summary by treatment and PT within SMQ will be provided. Severe/serious hepatobiliary disorders will be considered as AESI.

4.11.4.12.2. 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 Section 6.1, Appendix 1. A summary by treatment and PT within SMQ will be provided. Severe/serious acute gallbladder diseases will be considered as AESI.

4.11.4.12.3. Liver Enzymes

Analyses for laboratory analyte measurements are described in Section 4.11.4.5. 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 (≤1 × ULN, > 1 × ULN) to postbaseline with the following categories: ≤1 × ULN, >1 to <3 × ULN, ≥3 to <5 × ULN, ≥5 to <10 × ULN, ≥10 × ULN.
- A shift table of maximum to maximum aspartate transaminase (AST) measurement from baseline (≤1 × ULN, > 1 × ULN) to postbaseline with the following categories: ≤1 × ULN, >1 to <3 × ULN, ≥3 to <5 × ULN, ≥5 to <10 × ULN, ≥10 × ULN.
- Shift tables of maximum to maximum total bilirubin and direct bilirubin from baseline to postbaseline with the following categories: ≤1 × ULN, > 1 to <2 × ULN, ≥2 × ULN.
- Shift tables of serum alkaline phosphatase (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.

4.11.4.13. Gastrointestinal Safety

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

The maximum severity and duration of treatment-emergent nausea, vomiting, diarrhea, and combined through the end of the study will be summarized by treatment. The PTs in the gastrointestinal SOC will be used to identify gastrointestinal AEs. The incidence of the resulting TEAEs will be summarized by treatment and PT. PTs with severe/serious cases in the gastrointestinal SOC will be considered as AESIs. The gastrointestinal TEAE will also be summarized by participants with de-escalation or not.

4.11.4.14. Acute Renal Events

Laboratory measures related to renal safety will be analyzed as specified for laboratory measurements in Section 4.11.5.

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/1.73m^2). A max-to-max shift table of urine albumin-to-creatinine ratio (UACR), using the categories UACR <30 mg/g, 30 mg/g ≤ UACR ≤300 mg/g, UACR >300 mg/g (respectively, these represent normal, microalbuminuria, and macroalbuminuria).

The AE database will be searched using SMQs of acute renal failure and chronic kidney disease to identify events consistent with acute renal events. The incidence of the resulting TEAEs will be summarized by treatment and PT. Detailed searching criteria can be found in Section 6.1, Appendix 1. Severe/serious acute renal events will be considered as AESI.

4.11.4.15. Dehydration

The AE database will be searched using SMQ of dehydration to identify events consistent with dehydration. Detailed searching criteria can be found in Section 6.1, Appendix 1. Severe/serious dehydration events will be considered as AESIs.

4.11.4.16. 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 Section 6.1, 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 AESIs.

4.11.4.17. 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.

4.11.4.18. 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 Section 6.1, 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.

4.11.5. Clinical Laboratory Evaluations

Summaries will be created for the laboratory analyte measurements as noted in Appendix 2 of Study Protocol.

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.

4.11.6. Vital Signs, Physical Findings, and Other Observations Related to Safety

4.11.6.1. Vital Signs and Physical Characteristics

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 baseline HbA1c (≤8.5%, >8.5% [≤69, >69 mmol/mol]), any past prior use (yes or no) of any antihyperglycemic medication, treatment group, visit, and treatment-by-visit interaction as fixed effects, and baseline value of the dependent variable as a covariate. An unstructured covariance structure will model the relationship of within-patient errors. If this model fails to converge, additional covariance structures will be fitted under the same order as specified in Section 4.8.3.

Counts and percentages of patients with abnormal sitting systolic blood pressure (BP), sitting diastolic BP, and pulse will be presented by treatment. The criteria for identifying patients with treatment-emergent vital sign abnormalities are stated in Table GPIU.4.4.

Table GPIU.4.4. 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.

4.11.7. Device Product Complaints

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

4.12. Other Analyses

4.12.1. Subgroup Analyses

Efficacy subgroup analyses will be guided by the efficacy estimand.

Subgroup analyses of the change in HbA1c and the change in weight from baseline to Week 40 may be made to assess consistency of the intervention effect across the following subgroups:

- Age group: < 65 vs ≥ 65 years
- Age group: < 75 vs ≥ 75 years
- Sex: female vs male
- Duration of diabetes: <median vs ≥median

- Duration of diabetes: ≤5 vs >5 years
- Baseline HbA1c: ≤8.5% [≤69 mmol/mol] vs >8.5% [>69 mmol/mol]
- BMI: <28 vs ≥28 kg/m^2
- BMI: <25 vs ≥25 to <30 vs ≥30 kg/m^2
- Prior use of antihyperglycemic medication: yes vs no

Subgroups with few subjects may be excluded from subgroup analyses when appropriate.

Subgroup analysis regarding TEAE may be made to assess TEAE across different subgroups if deemed appropriate.

4.13. Interim Analyses

No interim analysis is planned for this study.

Approved on 22 Oct 2024 GMT

5. Sample Size Determination

Participants will be randomized in a CCI ratio to tirzepatide 5 mg, 10 mg, 15 mg, or placebo.

The trial is powered to assess superiority of tirzepatide 5 mg, 10 mg, or 15 mg versus placebo in parallel relative to change from baseline in HbA1c at Week 40 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 CCI participants in tirzepatide arms and up to CCI participants in placebo arm initiating any rescue medication or premature treatment discontinuation;
- at least CCI (placebo adjusted) mean reduction in HbA1c for the tirzepatide doses; and a common SD of CCI

On the basis of these assumptions, randomizing 200 participants using a CCI randomization ratio to tirzepatide 5 mg (50 participants), tirzepatide 10 mg (50 participants), tirzepatide 15 mg (50 participants), and placebo (50 participants) will provide at least CCI power to establish superiority for a tirzepatide dose compared to placebo at a 2-sided significance level of CCI. Furthermore, this sample size will ensure at least CCI power for the HbA1c superiority evaluation utilizing all available HbA1c data at Week 40 with missing data imputed with a conservative multiple imputation method, provided that the mean reduction in HbA1c for the tirzepatide doses is at least CCI (placebo adjusted) and SD increases to no more than CCI due to the inclusion of data on rescue medications, inclusion of data after premature treatment discontinuation, and missing data imputation.

6. Supporting Documentation

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

The AESI analyses are detailed in Section 4.11.4. The search criteria for each AESI are stored in CLUWE: T:\prd\ly3298176\common\AESI_Lab\Search criteria AESIs_TZP.xlsx.

7. References

Bretz F, Maurer W, Brannath W, Posch M. A graphical approach to sequentially rejective multiple test procedures. *Stat Med*. 2009;28(4):586-604. https://doi.org/10.1002/sim.3495

Bretz F, Posch M, Glimm E, et al. Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes, or parametric tests. *Biom J*. 2011;53(6):894-913. https://doi.org/10.1002/bimj.201000239

Carpenter JR, Roger JH, Kenward MG. Analysis of longitudinal trials with protocol deviation: a framework for relevant, accessible assumptions, and inference via multiple imputation. *J Biopharm Stat*. 2013;23(6):1352-1371. https://doi.org/10.1080/10543406.2013.834911

Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons, Inc.; 1987.

Signature Page for VV-CLIN-142838 v2.0

Approval	PPD Statistician 22-Oct-2024 11:00:21 GMT+0000

Signature Page for VV-CLIN-142838 v2.0