Statistical Analysis Plan Version 1 I8F-MC-GPIA

Efficacy and Safety of Tirzepatide Once Weekly in Chinese Participants without Type 2 Diabetes Who Have Obesity or are Overweight with Weight-Related Comorbidities: A Randomized, Double-Blind, Placebo-Controlled Trial (SURMOUNT-CN)

NCT05024032

Approval Date: 04-Jan-2023

1. Statistical Analysis Plan I8F-MC-GPIA: Efficacy and Safety of Tirzepatide Once Weekly in Chinese Participants without Type 2 Diabetes Who Have Obesity or are Overweight with Weight-Related Comorbidities: A Randomized, Double-Blind, Placebo-Controlled Trial (SURMOUNT-CN)

LY3298176 for Chronic Weight Management

Phase-3 randomized, double-blind, placebo-controlled trial comparing 3-doses of LY3298176 to Placebo in participants without Type 2 Diabetes Who have obesity or are overweight with weight-related comorbidities.

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Protocol I8F-MC-GPIA
Phase 3

Approved on 04 Jan 2023 GMT

2. Table of Contents

Section **Page**

Table of Contents

Table **Page**

Table of Contents

Figure **Page**

Table of Contents

Appendix **Page**

3. Revision History

This is the first version of the statistical analysis plan (SAP) for Study I8F-MC-GPIA.

4. Study Objectives

4.1. Primary Objective

The primary objective of the study is to demonstrate that tirzepatide 10 mg and/or 15 mg once-weekly (QW) are superior to placebo for percent change in body weight from randomization **and** percentage of participants who achieve ≥5% body weight reduction at 52 weeks.

4.2. Key Secondary Objectives

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

- To demonstrate that tirzepatide10 mg, and/or 15 mg QW are superior to placebo for change in body weight (kg) from randomization at 20 weeks.
- To demonstrate that tirzepatide 10 mg and/or 15 mg QW are superior to placebo for percentage of participants who achieve ≥10% body weight reduction from randomization at 52 weeks.
- To demonstrate that tirzepatide 10 mg and/or 15 mg QW are superior to placebo for percentage of participants who achieve ≥15% body weight reduction from randomization at 52 weeks.
- To demonstrate that tirzepatide 10 mg and/or 15 mg QW are superior to placebo for change from randomization in waist circumference (cm) at 52 weeks.

4.3. Other Secondary Objectives

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

- to demonstrate that tirzepatide10 mg, and/or 15 mg QW are superior to placebo at 52 weeks for
  - Mean change in absolute body weight (kg) from randomization
  - Mean change in body mass index (BMI) (kg/m^2) from randomization
  - Mean change in hemoglobin A1c (HbA1c) (%, mmol/mol) from randomization
  - Mean change in fasting glucose (mg/dL) from randomization
  - Mean change in SF-36v2 acute form physical functioning domain score from randomization
  - Mean change in Impact of Weight on Quality of Life-Lite-Clinical Trials Version (IWQOL-Lite-CT) physical function composite score from randomization
- to demonstrate that tirzepatide (10 mg and 15 mg combined) QW doses are superior to placebo at 52 weeks for
  - Mean change in diastolic blood pressure (DBP) (mmHg) from randomization

- Mean change in systolic blood pressure (SBP) (mmHg) from randomization
- Mean change in fasting lipids (mmol/L) from randomization (total cholesterol, LDL cholesterol, HDL cholesterol, non-HDL cholesterol, VLDL cholesterol, triglycerides, free fatty acids)
- Mean change in fasting insulin from randomization

4.4. Tertiary/Exploratory

To assess mean change from randomization to week 52 regarding the following parameters:

- SF-36v2 acute form Mental Component Score (MCS)
- SF-36v2 acute form Physical Component Score (PCS)
- SF-36v2 acute form Role-Physical domain (RP)
- SF-36v2 acute form Bodily Pain domain (BP)
- SF-36v2 acute form General Health domain (GH)
- SF-36v2 acute form Vitality domain (VT)
- SF-36v2 acute form Social Functioning domain (SF)
- SF-36v2 acute form Role-Emotional domain (RE)
- SF-36v2 acute form Mental Health domain (MH)
- IWQOL-Lite-CT total score
- IWQOL-Lite-CT Physical composite score
- IWQOL-Lite-CT Psychosocial composite score
- EQ-5D-5L utility score
- EQ-5D-5L VAS score

4.5. Pharmacokinetics

To characterize the PK of tirzepatide 10 mg and 15 mg QW and evaluate the relationship between tirzepatide exposure and safety, tolerability, and efficacy measures.

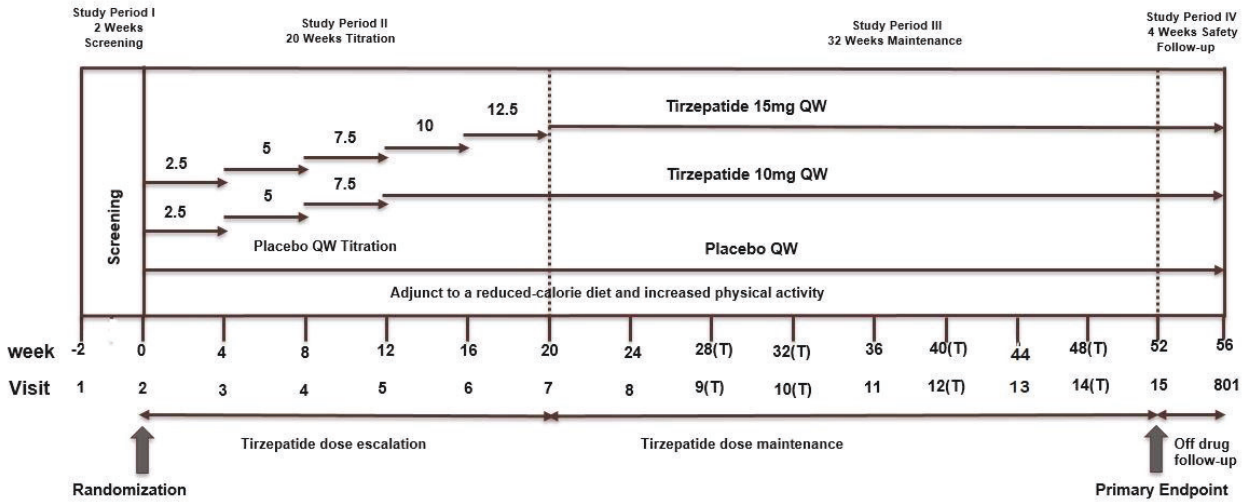
- Population PK and PD parameters

5. Study Design

5.1. Summary of Study Design

Study I8F-MC-GPIA (GPIA; SURMOUNT-CN) is a Phase 3, multicenter, randomized, placebo-controlled, double-blinded study of the safety and efficacy of 10 mg, and 15 mg tirzepatide QW, compared with placebo in Chinese participants without T2DM who either have obesity (BMI ≥28 kg/m^2) or overweight (BMI ≥24 kg/m^2) with the presence of at least one weight-related comorbidity (for example, hypertension, dyslipidemia, Obstructive sleep apnea or cardiovascular disease).

All participants will undergo a 2-week screening period and a 52-week dose escalation and treatment period, followed by a 4-week safety follow-up period.



Abbreviations: QW = once weekly; T = telephone visit.

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

The details about the overview of study periods and study visits can be found in Study GPIA protocol Section 5. The detail of the unique visits not displayed in Figure GPIA.5.1 is provided below.

Visit 99

Visit 99 is only applicable to participants who discontinue the study treatment prematurely prior to Visit 15 (Week 52). Participants will be asked to return for Visit 99 at 52 weeks ± 7 days after randomization. This visit is critical to ensure complete data collection for the primary weight-loss endpoint.

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

- measurement of weight and waist circumference
- concomitant medications

- assessment of adverse events (AEs), and
- completion of the mental health questionnaires (after the AE assessment).

Early Discontinuation of Treatment Visit

Participants unable or unwilling to continue the study treatment for any reason will perform an early discontinuation of treatment (ED) visit at the visit when the participant informs the site about the study treatment discontinuation.

Sample size determination

Approximately 210 participants will be randomly assigned to study intervention (70 participants per intervention group). The sample size determination assumes that evaluation of superiority of tirzepatide 10 mg and tirzepatide 15 mg to placebo will be conducted in parallel, each at a 2-sided significance level of 0.025 using a 2-sample t-test. Additionally, a difference of at least 11% mean body weight percentage reduction from randomization at 52 weeks for tirzepatide 10 mg and/or tirzepatide 15 mg compared with placebo, a common standard deviation (SD) of 11%, and a drop-out rate of 25% are assumed for statistical power calculations. Under the assumptions above, randomizing 210 participants in a 1:1:1 ratio to tirzepatide 10 mg (70), tirzepatide 15 mg (70), and placebo (70) provides more than 90% power to demonstrate superiority of each tirzepatide dose to placebo.

The chosen sample size and randomization ratio also provides >90% power to establish superiority of tirzepatide 10 mg and tirzepatide 15 mg doses to placebo in term of proportion of participants achieving at least 5% body weight reduction at 52 weeks, conducted in parallel using a Chi-square test, each at a 2-sided significance level of 0.025, assuming 25% placebo-treated participants and 90% tirzepatide-treated participants achieving the goal and a dropout rate of 25%.

6. A Priori Statistical Methods

6.1. Populations for Analyses

For purposes of analyses, Table GPIA.6.1 defines the analysis sets.

Table GPIA.6.1. Description of Analysis Datasets

Analysis Set	Description
Entered Participants	All participants who sign informed consent
Randomized Participants	All participants who are randomly assigned a study treatment
Modified Intent-to-Treat (mITT) Set	All randomly assigned participants who are exposed to at least 1 dose of study drug. Participants will be included in the treatment group they were randomized to.
Efficacy Analysis Set (EAS)	Data obtained during treatment period from mITT, excluding data after discontinuation of study drug (last dose date + 7 days).
Full Analysis Set (FAS)	Data obtained during treatment period from mITT, regardless of adherence to study drug.
Safety Analysis Set (SS)	Data obtained during the treatment period plus safety follow-up period from mITT, regardless of adherence to study drug.
Per Protocol Set (PPS)	Participants included in the randomized population who have completed Week 52 of study treatment without permanent discontinuation of IP and without significant protocol deviations through Week 52 that would significantly impact the primary objective. Treatment group will be defined on the basis of the treatment the participants actually receive.

6.2. General Considerations

Statistical analysis of this study will be the responsibility of Eli Lilly and Company (Lilly) or its designee. 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.

Unless otherwise noted, 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.

Unless specified otherwise, efficacy and safety will be assessed using the modified intention-to-treat (mITT) population, and data will be analyzed based on the randomized treatment (that is, not the actual treatment received by the participant). For primary efficacy analysis, the efficacy analysis will be conducted using Efficacy Analysis Set (EAS). Safety analysis will be conducted using Safety Analysis Set (SS).

Summary descriptive statistics for continuous measures will include sample size, mean, standard deviation (SD), median, minimum, and maximum. The analysis model to make comparisons among treatment groups relative to continuous measurements assessed over time will be either analysis of covariance (ANCOVA) or a mixed model for repeated measures (MMRM).

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 hazards rates among treatments.

Summary statistics for categorical measures (including categorized continuous measures) will include sample size, frequency, and percentages. Logistic regression will be used to examine the treatment difference in binary efficacy outcomes if there is a need to adjust for covariate. Otherwise, Fisher's exact test will be used to examine the treatment difference.

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

Unless specified otherwise, baseline will be defined as the last available non-missing measurement during Visit 1 to Visit 2. For the safety related parameters, the definition of baseline and postbaseline are specified in Table GPIA.6.2.

Table GPIA.6.2. Baseline and Postbaseline Definitions for Safety Groups

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).	Starts 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[a], Vital Signs, and ECGs.	Baseline will include all scheduled and unscheduled measurements during the baseline period (Visit 1 to Visit 2)	Postbaseline will be defined as measurements after Visit 2. All scheduled and unscheduled measurements will be included.
SS	1.3) Change from Baseline to Week *xx* and to Last Postbaseline for Labs[a],Vital Signs, and ECGs.	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. The early discontinuation (ED) visits are considered scheduled visits.

Abbreviations: ECGs = electrocardiogram; SS = Safety Analysis Set.

[a] Immunogenicity related analysis is specified in Section 6.13.3.6.

There will be 2 estimands of interest in evaluating primary and key secondary efficacy objectives.

- For objectives controlled for Type 1 error at 52 weeks:
  - the "efficacy" estimand, defined as the average treatment effect of tirzepatide relative to placebo at 52 weeks, in the randomized participants had they remained on their randomized treatment for the entire planned 52 weeks treatment duration.
  - the "treatment regimen" estimand, is the average treatment effect of tirzepatide relative to placebo at 52 weeks, for the randomized participants regardless of the adherence to treatment.

End of study participation for a participant 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 participants considered to be lost-to-follow-up, end of study participation will be the date of lost-to-follow-up reported by the investigator. Participant data included in the database after the last date of study participation (including safety follow-up period) will be excluded from statistical analysis.

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

Statistical summaries and results of statistical analyses will be displayed in the following order: Placebo, tirzepatide 10 mg, tirzepatide 15 mg, and pooled tirzepatide (all doses combined, if necessary).

Not all analyses described in this SAP will necessarily be included in the Clinical Study Reports (CSRs). Any analysis described in this SAP and not provided in the CSR would be available upon request. Not all displays will necessarily be created as a "static" display. Some may be incorporated into interactive display tools instead of or in addition to a static display.

6.3. Adjustments for Covariates

The study is stratified by sex (female, male) and presence of comorbidities (Y/N).

Unless otherwise specified, for efficacy related analyses at 52 weeks, sex, presence of comorbidities and corresponding baseline value will be used as a covariate.

6.4. Handling of Dropouts or Missing Data

For the primary and key secondary efficacy endpoint analyses subject to type 1 error rate control, data for participants with missing values at the 52-week visit will be imputed based on the method described in Section 6.12.1.1 and 6.12.1.3. However, for the parameters with only 1 postbaseline measure during the analysis period per schedule of activity, the last observation carried forward (LOCF) approach will be applied to impute the endpoint when ED measure is available.

6.5. Multiple Comparisons/Multiplicity

The type 1 error rate control strategy for primary and key secondary efficacy objectives is illustrated in Section 6.12.3, which is performed in "efficacy" estimand as efficacy analysis is for China registration purpose. In addition, no multiplicity adjustments will be made for evaluating other secondary and exploratory objectives, and safety assessments.

6.6. Patient Disposition

A listing of study disposition for all randomized participants will be provided at the final database lock, respectively. Summaries of study disposition and study drug disposition for all randomized participants will be provided by planned study treatment at the primary database lock (after 52-week treatment).

The study completion status is defined as follows: participants who have completed 52-week treatment period and with non-missing safety follow-up visit (Visit 801) will be considered as completers. Otherwise, they will be considered as non-completers.

6.7. Historical Illnesses and Preexisting Conditions

The count and percentages of participants with historical illnesses and preexisting conditions will be summarized by treatment group using the Medical Dictionary for Regulatory Activities (MedDRA) Preferred Term (PTs) nested within System Organ Class (SOC). The SOC will be in alphabetical order. Conditions (that is, PTs) will be ordered by decreasing frequency within SOC. This will be summarized for all randomized participants.

6.8. Patient Characteristics

A listing of participant demographics for all randomized participants will be provided. All demographic and baseline clinical characteristics will be summarized by study treatment for all randomized participants. Baseline demographic and clinical characteristics of special interest include but are not limited to: age (years), sex (female, male), height (cm), weight (kg), BMI (kg/m^2), waist circumference (cm), age group (<65 years, ≥65 years), BMI group (<28, ≥28 kg/m^2); BMI group2 (<30, ≥30 and <35, ≥35 and <40, ≥40 kg/m^2), weight-related comorbidities.

6.9. Concomitant Therapy

Concomitant medication will be summarized by PTs by treatment group by decreasing frequency for SS group.

Additionally, medications of interest (as defined below) will be summarized by treatment for SS.

Concomitant medications of interest include the following:

- baseline antihypertensive therapy
- baseline lipid lowering therapy
- changes to baseline medication in post-randomization (in term of type/class and dose):
    - antihypertensive therapy, and
    - lipid lowering therapy.
- utilization after randomization of:
    - antidiarrheal medication. and
    - antiemetic medication.

6.10. Treatment Exposure and Compliance

6.10.1. Study and Study Treatment Exposure

Summary of duration of follow-up (defined as time in days from date of randomization to the date of the last study visit) and/or 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 treatment group using data from SS in the period of the 52 weeks plus safety follow-up (Visit 801) for all the randomized participants.

For the summary of duration on study treatment, the frequency and percentage of participants falling into the following range will be summarized by planned treatment group as well:

- >0
- ≥4 weeks
- ≥8 weeks
- ≥12 weeks
- ≥16 weeks
- ≥20 weeks
- ≥24 weeks
- ≥36 weeks
- ≥44 weeks
- ≥52 weeks

In addition, the frequency and percentages of participants falling into the following study treatment exposure ranges may be summarized by planned treatment group:

- 0 weeks
- >0 to <4 weeks
- ≥4 to <8 weeks
- ≥8 to <12weeks
- ≥12 to <16 weeks
- ≥16 to <20 weeks
- ≥20 to <24 weeks
- ≥24 to <36 weeks
- ≥36 to <44 weeks
- ≥44 to <52 weeks
- ≥52 weeks

No p-values will be reported in these summaries as they are intended to describe the study populations rather than test hypotheses about them.

6.10.2. Compliance to Study Treatment

Summary of prematurely discontinuing study treatment (including discontinuation reason) will be provided by study treatment. A time-to-event analysis of premature study treatment discontinuation will also be conducted.

The analyses related to compliance will be conducted for the period during 52-week treatment period for all randomized participants.

If data warrants, the counts and percentages of participants who follow the planned escalation scheme, have dose interruption, or have dose de-escalation may be summarized for tirzepatide treatment group. In addition, the proportion of participants receiving 2.5 mg, 5 mg, 7.5 mg, 10 mg, 12.5 mg, or 15 mg may be presented by randomized tirzepatide treatment and visit during the dose escalation period.

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

6.11. Important Protocol Deviations

Important protocol deviations are identified in the Trial Issues Management Plan (TIMP). A listing and a summary of important protocol deviations by treatment will be provided at the end of 52 weeks treatment (for all randomized participants).

6.12. Efficacy Analyses

The assessment of efficacy objectives will be guided by the "efficacy" estimand using the EAS. All primary and key secondary efficacy assessments will be guided by the "treatment regimen" estimand conducted using the FAS for sensitivity analysis. Assessment of the primary and key secondary objectives will be conducted with multiple imputation of missing data (see Section 6.12.1.3).

6.12.1. Primary Efficacy Analysis

The primary efficacy measure will be percent change in body weight from randomization **and** percentage of participants who achieve ≥5% body weight reduction at 52 weeks. The percent change in body weight at each nominal visit is defined as:

*(postbaseline body weight [kg] – baseline body weight [kg]) / baseline body weight [kg] * 100%.*

Percent change in body weight will be summarized by treatment and nominal visit (week) from randomization to 52 weeks; Percentage of participants who achieve ≥5% body weight reduction will be summarized by treatment at 52-week.

6.12.1.1. Analysis Related to the Efficacy Estimand

The analysis related to efficacy estimand will be conducted utilizing data in the EAS.

For the mean percent body weight change from randomization, a MMRM will be conducted. Restricted maximum likelihood (REML) will be used to obtain model parameter estimates and the Kenward-Roger option will be used to estimate the denominator degrees of freedom. The response variable of MMRM will be the percent change in body weight from baseline values obtained at each scheduled post baseline visit.

For the percentage of participants achieving at least 5% body weight reduction from randomization over time, a logistic regression model will be used with the response variable of the percentage of participants achieving at least 5% body weight reduction from randomization at 52-week.

For MMRM the independent variables of analysis model are treatment group, visit, treatment-by-visit interaction, stratification factors (presence of comorbidities at randomization, sex) as fixed effects, and baseline body weight as a covariate. An unstructured covariance structure will model

relationship of within-patient errors. If this model fails to converge, the following variance covariance structures will be tested in order until convergence is achieved:

- heterogeneous Toeplitz
- heterogeneous first order autoregressive
- heterogeneous compound symmetry
- Toeplitz
- first order autoregressive, and
- compound symmetry.

The first covariance structure that converges will be used.

With the aid of the MMRM analysis, 2-sided 95% CIs for mean percent change in body weight from randomization to the 52-week visit for tirzepatide 10 mg and 15 mg compared to placebo will be derived and summarized. The resulting least squares mean (LSM) estimates of mean percent change in body weight from baseline will be plotted by visit and by study treatment.

A logistic regression model with terms of treatment group, sex, and presence of comorbidities at randomization as fixed effects, and baseline body weight as a covariate, will be conducted for percentage of participants achieving at least 5% body weight reduction from randomization at the 52-week visit. Missing body weight measurement at 52-week will be imputed by the predicted value from MMRM model aforementioned, then the continuous measurements will be dichotomized into status of achieving at least 5% body weight reduction (Yes or No). A logistic regression will be utilized to analyze proportion of participants with at least 5% body weight reduction by treatment at 52-week.

6.12.1.2. Analysis Related to the Treatment regimen Estimand

The analysis related to “treatment regimen” estimand will be conducted using data in the FAS.

The analysis for the mean percent change in body weight will be conducted utilizing ANCOVA. The response variable for the ANCOVA model will be percent change in body weight from randomization at 52 weeks. A logistic regression model will be used for the analysis of the percentage of participants achieving at least 5% body weight reduction obtained at the 52-week visit. Both models will include terms of treatment group, sex, and presence of comorbidities at randomization as fixed effects and baseline body weight as a covariate. The ANCOVA analysis will be conducted with multiple imputation of missing body weight at 52 weeks (see Section 6.12.1.3 for details) and statistical inference over multiple imputation of missing data guided by Rubin (1987). As for the logistic regression, missing body weight data at 52 weeks will be imputed first based on Section 6.12.1.3, then the continuous measurements will be categorized into status of achieving at least 5% body weight reduction (Yes or No).

With the aid of the ANCOVA model, 2-sided 95% CI for mean change in percent body weight from baseline to the 52-week visit between tirzepatide 10 mg and placebo, as well as tirzepatide 15 mg and placebo will be derived.

With the aid of the logistic regression model, 2-sided 95% CI and odds ratio for percentage of participants achieving at least 5% body weight reduction from baseline to the 52-week visit between tirzepatide 10 mg and placebo, as well as tirzepatide 15 mg and placebo will be derived.

6.12.1.3. Methods for Multiple Imputations

For efficacy analyses relative to "treatment regimen" estimand, the intercurrent events (ICEs) and the resulting missing values will be handled as follows:

- **Category 1:** for missing data solely due to exceptional circumstances, such as pandemic or natural disasters (after other reasons for missing data are ruled out), the analysis will consider the missing data as missing at random. The missing data will be imputed using all non-missing data of the primary outcome measurement from the same treatment arm.
- **Category 2:** for missing data due to all other ICEs: missing data will be imputed based on retrieved dropouts in the same treatment arm, defined as observed primary outcome measurements, from participants in the same treatment group, who had their efficacy assessed after early discontinuation of the study drug. In case where there are not enough retrieved dropouts to provide a reliable imputation model (for example, the model implemented by the SAS program does not converge), an alternative multiple imputation method with reference to the placebo group (that is, placebo multiple imputation) will be used.

6.12.1.4. Sensitivity Analysis Related to the Treatment regimen Estimand

Sensitivity analyses of the primary efficacy outcomes will be conducted using the FAS and guided by the "treatment-regimen" estimand, which represents the efficacy irrespective of adherence to study drug. This assessment will analyze percent change in body weight obtained at the 52-week visit using an ANCOVA and the percentage of participants achieving at least 5% body weight reduction obtained at the 52-week visit using a logistic regression model. The terms for both models will be the same as specified in Section 6.12.1.2 for "treatment regimen" estimand.

Missing values of change in body weight at the 52-week visit will be imputed based on observed body weight change from baseline values at the visit from participants in the same treatment group who had their efficacy assessed after early discontinuation of study drug. In cases where there are not enough retrieved dropouts to provide a reliable imputation model (for example, the model implemented by the SAS program does not converge), an alternative multiple imputation method with reference to the placebo group (placebo multiple imputation) will be used. Analysis will be conducted with multiple imputations.

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

Table GPIA.6.3. Secondary Measures Controlled for Type 1 Error

Objectives	Relative to the efficacy measure:	Analysis conducted in a manner similar to	Additional Information

Tirzepatide10 mg, and/or 15 mg QW is superior to placebo:	Mean change in body weight (kg) from randomization at 20 weeks	MMRM model in Section 6.12.1.1 for efficacy estimand and ANCOVA model in Section 6.12.1.2 for treatment regimen estimand	
Tirzepatide 10 mg and/or 15 mg is superior to placebo:	Percentage of participants achieving body weight reduction ≥10% (and ≥15%) at 52 weeks	logistic model in Section 6.12.1.1 for efficacy estimand and logistic model in Section 6.12.1.2 for treatment regimen estimand	
	Mean change in waist circumference (cm) from randomization at 52 weeks	MMRM model in Section 6.12.1.1 for efficacy estimand and ANCOVA model in Section 6.12.1.2 for treatment regimen estimand	LSM estimates will be plotted by treatment through 52-weeks.

Abbreviations: ANCOVA = analysis of covariance; LSM = least squares mean; MMRM = mixed model for repeated measures; QW = once-weekly;

Decision will be guided by the 2-sided p-values in each objective, and details will be included in Section 6.12.3.

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

Type I error rate control strategy for evaluation of primary and key secondary objectives is performed in "efficacy" estimand as efficacy analysis is for China registration purpose, which is illustrated in Figure GPIA.6.1. The hypotheses for the primary and key secondary objectives are as follows:

- $H_{10,1}$ and $H_{15,1}$: superiority of tirzepatide 10 mg and/or 15 mg versus placebo for percent change in body weight from randomization and percentage of participants who achieve ≥5% body weight reduction at 52 weeks
- $H_{10,2}$ and $H_{15,2}$: superiority of tirzepatide 10 mg and/or 15 mg versus placebo for change in body weight (kg) from randomization at 20 weeks.
- $H_{10,3}$ and $H_{15,3}$: superiority of tirzepatide 10 mg and/or 15 mg versus placebo for percentage of participants who achieve ≥10% body weight reduction from randomization at 52 weeks
- $H_{10,4}$ and $H_{15,4}$: superiority of tirzepatide 10 mg and/or 15 mg versus placebo for percentage of participants who achieve ≥15% body weight reduction from randomization at 52 weeks
- $H_{10,5}$ and $H_{15,5}$: superiority of tirzepatide 10 mg and/or 15 mg versus placebo for change from randomization in waist circumference (cm) at 52 weeks

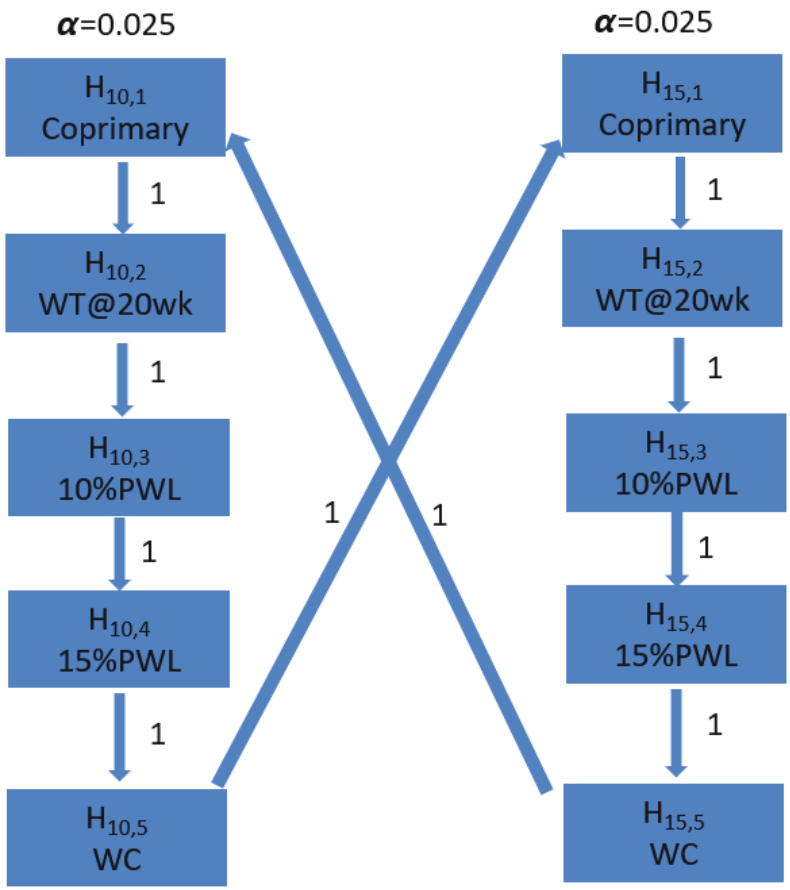


Figure GPIA.6.1 Type 1 error control strategy for primary and key secondary efficacy endpoints.

Abbreviations: PWL = percent weight loss WC = waist circumference; wk = week, WT = weight loss in kg

6.12.4. Other Secondary and Exploratory Efficacy Analyses

Other secondary and exploratory efficacy analyses will be conducted for EAS. Missing data will be imputed using LOCF or handled through MMRM without utilizing multiple imputation technique, except for mean change in fasting lipids (mmol/L) from randomization (total cholesterol, LDL cholesterol, HDL cholesterol, non-HDL cholesterol, VLDL cholesterol, triglycerides, free fatty acids), the analysis related to "efficacy estimand will be conducted in the EAS and "treatment regimen" estimand will be conducted using data in the FAS.

6.12.4.1. Other Secondary Efficacy Analyses

Table GPIA.6.4. Secondary Measures Not Controlled for Type 1 Error

Objective	Relative to the efficacy measure:	Analysis conducted in a manner similar to:	Additional Information
Tirzepatide10 mg and/or 15 mg QW is superior to placebo at 52 weeks:	Mean change in body weight (kg) from randomization	MMRM model in Section 6.12.1.1	LSM estimates through 52 weeks will be plotted by study treatment.
	Mean change in BMI (kg/m^2) from randomization	MMRM model in Section 6.12.1.1	Use baseline BMI (kg/m^2) as a covariate. LSM estimates through 52 weeks will be plotted by study treatment.
	Mean change in HbA1c (%, mmol/mol) from randomization	MMRM model in Section 6.12.1.1	Use baseline HbA1c (%, mmol/mol) as a covariate. LSM estimates through 52 weeks will be plotted by study treatment.
	Mean change in fasting glucose (mg/dL) from randomization	MMRM model in Section 6.12.1.1	Use baseline fasting glucose as a covariate. LSM estimates through 52 weeks will be plotted by study treatment.
	Mean change in SF-36v2 acute form physical functioning domain score from randomization	ANCOVA model	with terms of treatment, stratification factors, and baseline SF-36v2 PF score as a covariate. Missing data will be imputed using LOCF method.
	Mean change in IWQOL-Lite-CT Physical Function composite (PF) score from randomization	ANCOVA model	with terms of treatment, stratification factors, and baseline IWQOL-Lite CT PF score as a covariate. Missing data will be imputed using LOCF method.
Tirzepatide QW (all doses combined) is superior to placebo at 52 weeks	Mean change in DBP (mmHg), SBP (mmHg), fasting lipids (mmol/L): total-C (mg/dL), LDL-C (mg/dL), HDL-C (mg/dL), non-HDL-C, VLDL-C (mg/dL), triglycerides (mg/dL), free fatty acids (mg/dL) , fasting insulin from randomization	MMRM model in Section 6.12.1.1 ANCOVA model	All tirzepatide doses will be pooled together. LSM estimates through 52 weeks will be plotted by study treatment. with terms of treatment, stratification factors, and baseline value of lipid parameter as a covariate. Missing data will be imputed using method described in 6.12.1.3.

Abbreviations: BMI = body mass index; DBP = diastolic blood pressure; HbA1c = hemoglobin A1c; HDL-C = high-density lipoprotein cholesterol; IWQOL-Lite-CT = Impact of Weight on Quality of Life-Lite-Clinical Trials Version; LDL = low-density lipoprotein; LOCF = last observation carried forward; LSM = least squares mean; MMRM = mixed model for repeated measures; QW = once-weekly; SF-36v2 = Short Form 36 version 2 Health Survey; VLDL-C = very low-density lipoprotein cholesterol.

6.12.4.2. Exploratory Efficacy Analyses

Table GPIA.6.5. Exploratory Efficacy Analysis

Objective	Relative to the efficacy measure:	Analysis Conducted
compare tirzepatide 10 mg, and/or 15 mg with placebo at 52 weeks	Percentage of participants achieving ≥20% and 25% body weight reduction from randomization	Longitudinal logistic model in Section 6.12.1.1
compare tirzepatide 10 mg, and/or 15 mg with placebo at 52 weeks	Percentage of participants whose BMI shifts between clinically relevant categories, ie, from baseline (<28, ≥28) to postbaseline (<24, ≥24 and <28, ≥28)	Shift analysis will be conducted based on data from EAS.
visualize tirzepatide10 mg and/or 15 mg with placebo percent weight loss change from randomization up to safety follow-up after 52 weeks	Percent body weight loss measured at postbaseline from randomization to 52 weeks plus 4 weeks safety follow-up	Time-course plot will be generated. Only the participants who complete the 52 weeks treatment and have the safety follow-up (Visit 801) will be included.
compare tirzepatide 10 mg, and/or 15 mg to placebo at 52 weeks:	• SF-36v2 acute form Mental Component Score (MCS) • SF-36v2 acute form Physical Component Score (PCS) • SF-36v2 acute form Role-Physical domain (RP) • SF-36v2 acute form Bodily Pain domain (BP) • SF-36v2 acute form General Health domain (GH) • SF-36v2 acute form Vitality domain (VT) • SF-36v2 acute form Social Functioning domain (SF) • SF-36v2 acute form Role-Emotional domain (RE) • SF-36v2 acute form Mental Health domain (MH) • IWQOL-Lite-CT total score • IWQOL-Lite-CT Physical composite score • IWQOL-Lite-CT Psychosocial composite score • EQ-5D-5L utility score • EQ-5D-5L VAS score	Only data from participants from EAS will be included. ANCOVA model with terms of treatment, stratification factors, and baseline score as a covariate. Missing data will be imputed using LOCF method.

Abbreviations: BMI = body mass index; EAS = Efficacy Analysis Set; IWQOL-Lite-CT = Impact of Weight on Quality of Life-Lite-Clinical Trials Version; MMRM = mixed model for repeated measures; QW = once weekly; T2DM = type 2 diabetes mellitus; VAS = Visual Analog Scale.

In addition, the following endpoints, risk difference in proportions for an unconditional treatment effect (Ge et al.2011) among each tirzepatide dose and placebo arm may be conducted:

- percentage of participants achieving at least 5% body weight reduction at 52 weeks

- percentage of participants achieving at least 10% body weight reduction at 52 weeks
- percentage of participants achieving at least 15% body weight reduction at 52 weeks
- percentage of participants achieving at least 20% body weight reduction at 52 weeks.
- percentage of participants achieving at least 25% body weight reduction at 52 weeks.

6.13. Safety Analyses

Unless specified otherwise, safety assessments will be based on the SS (Table GPIA.6.1). All events that occur between the first dose date of study drug and the end date of study participation will be included, regardless of the adherence to study drug.

The statistical assessment of homogeneity of the distribution of categorical safety responses between tirzepatide doses and placebo will be conducted using Fisher's exact test, unless specified otherwise.

For the selected continuous safety parameters, the mean change from baseline differences vs. placebo will be assessed via an MMRM using REML. Data from scheduled visits will be utilized for this analysis unless specified otherwise, the model for the analysis during 52 weeks treatment period will include sex, presence of comorbidities at randomization, treatment group, visit and treatment-by-visit interaction as fixed effects, and baseline value of the safety parameter as a covariate. To model the covariance structure within participants, the unstructured covariance matrix will be used. If this model fails to converge, the covariance structures specified in Section 6.12.1.1 will be tested in order until met convergence. If the data does not warrant the MMRM model, then ANCOVA model will be conducted.

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

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

Unless otherwise specified, all the analyses listed in this section will be conducted using SS in the postbaseline period: up to 52 weeks treatment period plus the safety follow-up (Visit 801, if applicable) for all randomized participants.

6.13.1. Analysis of Adverse Events

6.13.1.1. Treatment Emergent Adverse Events

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

For events occurring on the day of first taking study medication, the case report form (CRF)-collected information (for example, treatment emergent flag, start time of study treatment and event) will be used to determine whether the event was pre- versus post-treatment if available. If the relevant information is not available, then the events will be counted as post-treatment.

The counts and percentages of participants with TEAEs will be summarized by treatment using MedDRA Preferred Term (PT) nested within System Organ Class (SOC). Statistical comparisons will be applied at both the SOC and PT levels. Events will be ordered by decreasing frequency within SOC. The SOC will be in alphabetical order. For events that are sex-specific, the denominator and computation of the percentage will include only participants from the given sex.

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

The counts and percentages of patients with TEAEs by maximum severity will be summarized by treatment 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 nonmissing severities. If all severities are missing for the defined postbaseline period of interest, it will show as missing in the table.

6.13.1.2. Common Adverse Events

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

6.13.1.3. Deaths

A listing of all deaths during the study will be provided. The listing will include participant identification including the treatment, site number, date of death, age at the time of enrollment, sex, cause of death as reported by investigator, cause of death as adjudicated by CEC. etc.

6.13.1.4. Other Serious Adverse Events

The counts and percentages of participants who experienced a SAE (including deaths and SAEs temporally associated or preceding deaths) during the postbaseline period will be summarized by

treatment using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency within SOC. The SOC will be in alphabetical order.

A listing of all SAEs will be provided. The listing will include treatment, participant identification including the site number, date of event, age at the time of enrollment, sex, MedDRA SOC and PT, Reported term, severity, outcome, relationship to study drug, time from first dose of study drug to the event, AE start date, AE end date, seriousness, action taken related to study treatment and event outcome.

6.13.1.5. Other Significant Adverse Events

The counts and percentages of participants who discontinued from study treatment or study due to an AE during the postbaseline period may be summarized by treatment group using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency within SOC.

6.13.2. Patient Narratives

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

- death
- serious adverse event, or
- permanent discontinuation of study treatment due to AEs.
- Pregnancy
- Adjudication Confirmed Major Adverse Cardiovascular Events
- Adjudication Confirmed Pancreatitis

Patient narratives (patient level data and summary paragraph) will be provided for participants in the randomized population with at least 1 notable event.

6.13.3. Special Safety Topics

For adverse events of special interest (AESI) or special safety topics, the counts and percentages of participants will be summarized by treatment and PT with decreasing frequency. Individual participant level data may be presented. Displays with individual participant level data may be created using various formats, such as a customized listing and/or a customized graphical participant profile. Adverse events of special interest are defined in each section of special safety topics, where applicable.

6.13.3.1. Acute Gallbladder Disease

All events of TEAE biliary colic, cholecystitis, or other suspected events related to gallbladder disease will be summarized by treatment groups by PT with decreasing frequency. Detailed searching criteria can be found in Appendix 1.

6.13.3.2. Exocrine Pancreas Safety

6.13.3.2.1. Pancreatic Enzyme

Observed pancreatic enzyme data (p-amylase and lipase) will be summarized by treatment and nominal visit.

The counts and percentages of participants with maximum postbaseline pancreatic enzyme value exceeding the following thresholds will be provided by baseline pancreatic enzyme value (≤ upper limit of normal [ULN], > ULN), and postbaseline: ≤1x ULN, (>1 to ≤3) x ULN, (>3 to ≤5) x ULN, (>5 to ≤10) x ULN, >10x ULN.

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

6.13.3.2.2. Pancreatitis Events

Summaries of adjudicated and investigator-reported pancreatic events will be provided by treatment. Detailed searching criteria can be found in Appendix 1.

Treatment emergent adjudicated pancreatitis will be considered as AESI. Listing of participants with adjudicated pancreatitis may be provided if deemed necessary.

6.13.3.3. Gastrointestinal Safety

6.13.3.3.1. Nausea, Vomiting, and Diarrhea

Summaries and analyses for incidence and severity of nausea, vomiting (including "vomiting" and "vomiting projectile"), diarrhea (including "diarrhea" and "diarrhoea"), and 3 events combined, will be provided by each treatment group.

Summary of the prevalence over time for nausea, vomiting, and diarrhea will also be presented.

Time to the onset of nausea, vomiting, and diarrhea will be plotted.

6.13.3.3.2. Severe Gastrointestinal Events

Severe gastrointestinal (GI) adverse events will be captured with AE-CRF form and serious cases will be captured with the SAE form. The PTs in the GI SOC MedDRA version at the time of database locks will be used to identify GI AEs, and only the PTs with serious/severe treatment-emergent cases will be considered as AESIs.

The counts and percentages of participants with severe/serious treatment-emergent GI events will be summarized by treatment.

6.13.3.4. Hepatic Safety

6.13.3.4.1. Hepatobiliary Disorders

Hepatobiliary disorders will be considered as AESI. The counts and percentages of participants with treatment-emergent potentially drug-related hepatic disorders will be summarized by treatment using the MedDRA PTs. Detailed searching criteria can be found in Appendix 1.

A listing of participants with treatment emergent hepatobiliary disorders may be provided if deemed necessary.

6.13.3.4.2. Liver Enzymes

Analyses for laboratory analyte measurements are described in Section 6.13.6. This section describes additional analyses of liver enzymes.

The counts and percentages of participants with the following elevations in hepatic laboratory tests at any time during the treatment period and during the entire study including follow up period will be summarized between treatment groups:

- The counts and percentages of participants with an alanine aminotransferase (ALT) measurement ≥3 times (3x), 5 times (5x), and 10 times (10x) the Covance ULN will be summarized for all participants with a postbaseline value and for subsets based on various levels of baseline value.
    - participants whose nonmissing maximum baseline value is ≤1x ULN
    - participants whose maximum baseline is >1x ULN, and
    - participants whose baseline values are missing.
- The counts and percentages of participants with an aspartate aminotransferase (AST) measurement ≥3x, 5x, and 10x the Covance ULN during the treatment period will be summarized for all patients with a postbaseline value and for subsets based on various levels of baseline, as described above for ALT.
- The counts and percentages of participants with a total bilirubin (TBL) measurement ≥2x the Covance ULN during the treatment period will be summarized for all participants with a postbaseline value, and subset into 4 subsets:
    - participants whose nonmissing maximum baseline value is ≤1 x ULN
    - participants whose maximum baseline is >1 x ULN, but <2 x ULN
    - participants whose maximum baseline value is ≥2 x ULN, and
    - participants whose baseline values are missing.
- The counts and percentages of participants with a serum alkaline phosphatase (ALP) measurement ≥2x the Covance ULN during the treatment period will be summarized for all participants with a postbaseline value and for the subsets described above to TBL.

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

6.13.3.5. Hypoglycemia

The following categories in accordance with the 2019 American Diabetes Association position statement on glycemic targets (ADA 2019b) will be defined in the database.

Glucose alert value (Level 1):

- **Documented symptomatic hypoglycemia** is defined as any time a patient feels that he/she is experiencing symptoms and/or signs associated with hypoglycemia, and has a blood glucose (BG) level of <70 mg/dL (<3.9 mmol/L).
- **Documented asymptomatic hypoglycemia** is defined as any event not accompanied by typical symptoms of hypoglycemia, but with a measured BG <70 mg/dL (<3.9 mmol/L).
- **Documented unspecified hypoglycemia** is defined as any event with no information about symptoms of hypoglycemia available, but with a measured BG <70 mg/dL (<3.9 mmol/L).

Documented Clinically Significant Hypoglycemia (Level 2):

- **Documented symptomatic hypoglycemia** is defined as any time a participant feels that he/she is experiencing symptoms and/or signs associated with hypoglycemia and has a BG level of <54 mg/dL (<3.0 mmol/L).
- **Documented asymptomatic hypoglycemia** is defined as any event not accompanied by typical symptoms of hypoglycemia but with a measured BG <54 mg/dL (<3.0 mmol/L).
- **Documented unspecified hypoglycemia** is defined as any event with no information about symptoms of hypoglycemia available but with a measured BG <54 mg/dL (<3.0 mmol/L).

Severe Hypoglycemia (Level 3):

- **Severe hypoglycemia** is defined as an episode with severe cognitive impairment requiring the assistance of another person to actively administer carbohydrate, glucagon, or other resuscitative actions. These episodes may be associated with sufficient neuroglycopenia to induce seizure or coma. Plasma glucose measurements may not be available during such an event, but neurological recovery attributable to the restoration of PG to normal is considered sufficient evidence that the event was induced by a low PG concentration.

Other hypoglycemia categories:

- Nocturnal hypoglycemia is defined as any hypoglycemic event that occurs between bedtime and waking.

If a hypoglycemic event meets the criteria of severe, the event would specifically be collected as an SAE. Serious hypoglycemia are defined by pharmacovigilance criteria and will also be captured with an SAE form.

To avoid duplicate reporting, all consecutive hypoglycemic events occurring within a 1-hour period will be considered to be a single hypoglycemic event.

Both the incidence (percent of patients experiencing ≥1 episode) and the rate (episodes/patient/year) of level 2 or level 3 hypoglycemia, and level 3 hypoglycemia will be reported by treatment group.

Severe/serious hypoglycemia will be considered as AESIs. The summaries of severe/serious hypoglycemia will be provided by treatment group. A listing of all events of severe hypoglycemia may be provided, if deemed necessary. This listing will provide treatment allocation, clinical characteristics of the hypoglycemic event, and concomitant antihyperglycemic medications.

6.13.3.6. Immunogenicity

6.13.3.6.1. Definitions of Sample ADA Status

At a high level, an individual sample is potentially examined multiple times, in a hierarchical procedure, to produce a sample anti-drug antibodies (ADA) assay result and potentially multiple cross-reactive antibodies assay results. The cut points used, the drug tolerance of each assay, and the possible values of titers are operating characteristics of the assay.

Figure GPIA. 6.2 details a flow chart that reflects the multitiered testing approach.

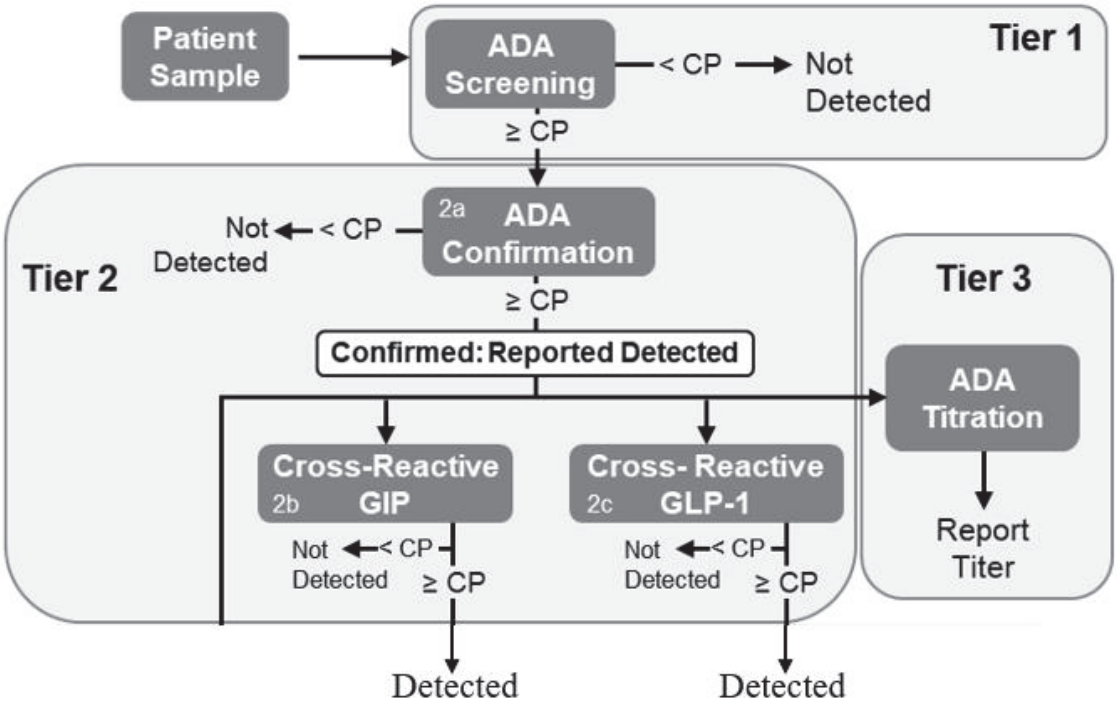


Figure GPIA. 6.2 Flowchart of immunogenicity multitiered testing approach.

Abbreviations: ADA = anti-drug antibody; CP = cut point; GIP = glucose-dependent insulinotropic polypeptide; GIPR = glucose-dependent insulinotropic polypeptide receptor; GLP-1 = glucagon-like peptide-1; GLP-1R = glucagon-like peptide-1 receptor; LY = LY3298176; NAb =neutralizing antibodies.

Table GPIA.6.6 outlines the results as reported from Tier 2a of the multitiered testing approach.

Table GPIA.6.6. Sample Anti-Drug Antibodies (ADA) Assay Results

Sample Laboratory Result	Explanation
Detected	ADA are detected and confirmed.
Not Detected	The raw result as reported from the laboratory indicates not detected. The clinical interpretation of such results depends other factors (see Table GPIA.6.7).
NO TEST, QNS, etc	Sample exists but was unevaluable by the assay.

Abbreviations: ADA = anti-drug antibodies; QNS = quantity not sufficient.

It can be the case that the presence of high concentrations of tirzepatide will affect immunoassays, and conversely high levels of antibodies may affect the measurement of tirzepatide concentration. Thus, a tirzepatide drug concentration, assessed from a sample drawn at the same time as the ADA sample, plays a key role in clinical interpretation of a sample when the laboratory results are Not Detected (see Table GPIA.6.7).

Table GPIA.6.7. Sample Clinical Anti-Drug Antibodies (ADA) Interpretation Results

Sample Clinical Interpretation	Explanation
ADA Present	ADA assay result is Detected
ADA Not Present	ADA assay result is Not Detected and simultaneous drug concentration is at a level that has been demonstrated to not interfere in the ADA detection method (ie, drug concentration is below the assay's drug tolerance level). For participants receiving placebo, drug concentration is not assessed and is assumed to be below the assay's drug tolerance level. If drug concentration was planned but is not available for a treatment-period sample, a Not Detected sample will be declared ADA Not present.
ADA Inconclusive	ADA assay result is Not Detected but drug concentration in the sample is at a level that can cause interference in the ADA detection method.
ADA Missing	ADA sample not drawn, QNS, not tested, etc., causing there to be no laboratory result reported or the result is reported as "no test."

Abbreviations: ADA = anti-drug antibodies; QNS = quantity not sufficient.

Similar terminology to Table GPIA.6.7 applies for each type of cross-reactive assay. Importantly, each of these are distinct assays and, in general, have different assay operating characteristics.

6.13.3.6.2. Definitions of Immunogenicity Assessment Periods

Immunogenicity Baseline Observations: Baseline period for immunogenicity assessment for each participant includes all observations on or prior to baseline visit. In instances where multiple baseline observations are collected, to determine participant ADA status the last nonmissing

immunogenicity assessment prior to first administration of study is used to determine treatment-emergent status (see below).

Immunogenicity Postbaseline Period Observations: Postbaseline period observations for each participant includes all observations after the first administration of study drug. There are two different periods listed below:

- The planned treatment period is defined as from the first dose of treatment to end of the treatment period
- The entire postbaseline period is defined as from the first dose of treatment to the end of safety follow-up visit or date of study withdrawal

6.13.3.6.3. Definitions of Participant ADA Status

Treatment-emergent (TE) ADA-evaluable participants: a participant with a non-missing baseline ADA result and at least 1 non-missing postbaseline ADA result.

TE ADA-unevaluable participant: any participant who does not meet the evaluable criteria.

Baseline ADA Present (preexisting antibody): ADA detected in a sample collected up to the first dose date and time.

Baseline ADA Not Present: ADA is not detected, and the corresponding PK concentration is missing or below the drug tolerance limit in a sample collected up to the first dose date and time.

Treatment-emergent ADA positive (TE ADA+) participant: an evaluable participant who had a:

- baseline status of ADA Not Present and at least 1 postbaseline status of ADA Present with titer ≥2*MRD, where the MRD is the minimum required dilution of the ADA assay.
- baseline and postbaseline status of ADA Present, with the postbaseline titer being 2 dilutions (4-fold) greater than the baseline titer. That is, the participant has baseline status of ADA Present, with titer 1:B, and at least 1 postbaseline status of ADA Present, with titer 1:P, with P/B ≥ 4.

As shown in Figure GPIA. 6.2 , a titer is expected when ADA assay result is Detected. On occasion, the corresponding assay cannot be performed, in which case a titer value will be imputed for the purpose of TE ADA determination. A baseline sample with detected ADA and no titer is imputed to be the minimum required dilution (MRD) (1:10) and a post-baseline sample with ADA detected and no titer is imputed to be one dilution above the MRD (1:20).

TE ADA Inconclusive participant: A TE ADA-evaluable participants TE ADA Inconclusive if ≥ 20% of the participant's postbaseline samples are ADA Inconclusive and all remaining postbaseline samples are ADA Not Present.

TE ADA negative (TE ADA-) participant: A TE ADA-evaluable participant is TE ADA- when the participant is not TE ADA+ and not TE ADA Inconclusive.

6.13.3.6.4. Analyses to be Performed

The count and proportion of participants who are TE ADA+ will be tabulated by treatment group, where proportion are relative to the number of participants who are TE ADA evaluable, as defined above. The tabulation will include the number and proportion of participants with ADA Present at baseline and the number and proportion of TE ADA+ participant exhibiting each type of cross-reactive antibodies.

This analysis will be performed for

- the planned treatment period
- the entire postbaseline period including follow-up.

A summary will be provided of the count and percentage of tirzepatide-treated participants experiencing specific TEAE (see Table GPIA.6.8) by participant TE ADA status (TE ADA+, TE ADA-, TE ADA Inconclusive) during the planned treatment period. The PT will be ordered by decreasing incidence in TE ADA+ status group.

Table GPIA.6.8. Adverse Events for Analysis with Immunogenicity Results

TE AE category	Criteria
Hypersensitivity reactions	Anaphylaxis SMQ (narrow or algorithm)
	Hypersensitivity SMQ (narrow)
	Angioedema SMQ (narrow)
	Severe Cutaneous Adverse Reaction SMQ (narrow)
Injection site reactions	Injection site reaction HLT
	Infusion site reaction HLT
	Administration site reaction HLT

Abbreviations: HLT= high level term; MedDRA = Medical Dictionary for Regulatory Activity; SMQ = Standardised MedDRA Query; TEAE = treatment-emergent adverse event

Additional immunogenicity analyses as determined later may be presented. The relationship between antibody titers, the PK parameters, and pharmacodynamics (PD) response to tirzepatide may also be assessed.

Cases of TE ADA that are associated with TEAEs of either hypersensitivity or injection site reaction will be classified as AESIs.

6.13.3.7. Hypersensitivity Reactions

Two main analyses are performed in support of assessment of potential immediate hypersensitivity, including anaphylaxis as well as potential nonimmediate hypersensitivity.

Time Period A, of potential immediate hypersensitivity includes all TEAEs occurring from start of study drug administration up to 24 hours after end of study drug administration. For events without the hypersensitivity electronic case report form (eCRF), only date (no time) information is collected, the event occurred on the same date as the study drug injection date will be included in Time Period A.

Time Period B, of potential non-immediate hypersensitivity, includes all TEAEs occurring more than 24 hours after the end of study drug administration, but prior to subsequent drug administration.

The counts and percentages of participants who experienced a TEAE will be summarized by PT with decreasing frequency by treatment.

Detailed searching criteria can be found in Appendix 1. Within query, individual PTs that satisfied the queries will be summarized. For Time Period A analysis, the Anaphylactic reaction SMQ algorithm will be summarized. Also, a single event may satisfy multiple SMQs, in which case the event contributes to every applicable SMQ.

6.13.3.7.1. Hypersensitivity Reactions

The cases of treatment-emergent hypersensitivity will be considered as AESIs. All serious hypersensitivity events will be collected with the SAE form. A listing of participants with serious hypersensitivity reactions may be provided if deemed necessary.

6.13.3.8. Injection Site Reaction

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

Patient-based analysis and event-based analysis may be provided if necessary. The patient-based analysis summarizes all injection-site reaction (ISR) questionnaire forms for an individual patient with a single statistic, typically an extreme value. This analysis allows each patient to contribute only once for each parameter, at the expense of a focus on the most extreme events. By contrast, the event-based analysis summarizes all ISR questionnaire forms received, without regard to individual patients. This provides characteristics of ISR events as a proportion of all events for which questionnaire responses were provided, at the expense of some potential bias due to differential contribution of individual patients to the analysis.

The counts and percentages of participants with treatment emergent injection site reaction will be summarized by treatment using the MedDRA PTs. Detailed searching criteria can be found in Appendix 1.

The PT will be listed for summary in decreasing order of incidence for tirzepatide-treated participants.

6.13.3.8.1. Injection Site Reactions

The treatment-emergent injection site reactions (for example, abscess, cellulitis, erythema, hematomas/hemorrhage, exfoliation/necrosis, pain, subcutaneous nodules, swelling, induration, inflammation) will be considered as AESIs.

The counts and percentage of participants with treatment-emergent ISRs will be summarized by treatment. A listing of participants with treatment-emergent ISRs may be provided, if deemed necessary.

6.13.3.9. Major Adverse Cardiovascular Events

The following positively adjudicated major adverse cardiovascular events (MACE) will be considered as AESIs:

- death due to cardiovascular AEs
- myocardial infarction (MI),
- hospitalization for unstable angina,
- hospitalization for heart failure,
- coronary interventions (such as coronary artery bypass graft or percutaneous coronary intervention),
- cerebrovascular events, including cerebrovascular accident (stroke) and transient ischemic attack.

The counts and percentages of participants with positively adjudicated MACE may be summarized by treatment.

In addition, MACE reported by investigator may also be summarized although a major adverse cardiovascular event reported by investigator is not considered as AESI.

A listing of participants reporting MACE events, either reported by investigator or identified by the clinical endpoint committee (CEC), may be provided.

6.13.3.10. Major Depressive Disorder/Suicidal Ideation

The treatment-emergent major depressive disorder/suicidal ideation or behavior will be captured as AESI. Adverse events will be searched using MedDRA PT terms. Detailed searching criteria can be found in Appendix 1.

The counts and percentages of participants with TEAEs will be summarized by treatment group using MedDRA PT nested within SMQ. Events will be ordered by decreasing frequency in the total tirzepatide group nested within SMQ. A listing of participants with major depressive disorder/suicidal ideation or behavior may be provided if deemed necessary.

Additionally, suicidal ideation and behavior, and depression will be assessed by the investigator via spontaneously reported AEs and through the use of the Columbia-Suicide Severity Rating Scale (C-SSRS) and the Patient Health Questionnaire (PHQ-9).

6.13.3.10.1. Patient Health Questionnaire

Total scores for the PHQ-9 range from 0 to 27 with total scores categorized as

- none (not depressed): 0 through 4
- mild: 5 through 9
- moderate: 10 through 14
- moderately severe: 15 through 19, and
- severe: 20 through 27.

Shift tables will be provided showing the counts and percentages of participants within each baseline category (maximum value) versus each postbaseline category (maximum value) by treatment.

Additionally, the following 3 outcomes of interest will be compared between treatments (based on the maximum value during baseline and postbaseline):

- any increase in depression category (ie, worsening of depression): includes participants in the none, mild, moderate, or moderately severe category during baseline and with at least 1 postbaseline measurement
- increase from No or Mild Depression to Moderate, Moderately Severe or Severe Depression: includes participants in the none or mild depression category during baseline and with at least 1 postbaseline measurement; and
- increase from Mild or Moderate Depression to Moderately Severe or Severe Depression: includes participants in the mild or moderate depression category during baseline and with at least 1 postbaseline measurement

Sensitivity analysis will be conducted using the Modified ITT - Safety Analysis Set excluded data completed by investigator.

6.13.3.10.2. Suicidal Ideation and Behavior Solicited Through C-SSRS

Suicide-related thoughts and behaviors occurring during the entire study period, based on the C-SSRS, will be summarized by treatment group. In particular, for each of the following suicide-related events, the counts and percentages of participants with the event will be summarized by treatment group:

- died by suicide
- nonfatal suicide attempt
- interrupted attempt
- aborted attempt
- preparatory acts or behavior
- active suicidal ideation with specific plan and intent
- active suicidal ideation with some intent to act without specific plan
- active suicidal ideation with any methods (no plan) without intent to act

- nonspecific active suicidal thoughts
- wish to be dead, and
- non-suicidal, self-injurious behavior.

In addition, the counts and percentages of participants who experienced at least 1 of the composite measures will be presented. The participants with at least 1 post-baseline C-SSRS assessment are included. The composite measure is determined at each assessment by the "yes" or "no" responses in C-SSRS categories by the study participant:

- Category 1 – Wish to be Dead
- Category 2 – Non-specific Active Suicidal Thoughts
- Category 3 – Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act
- Category 4 – Active Suicidal Ideation with Some Intent to Act, without Specific Plan
- Category 5 – Active Suicidal Ideation with Specific Plan and Intent
- Category 6 – Preparatory Acts or Behavior
- Category 7 – Aborted Attempt
- Category 8 – Interrupted Attempt
- Category 9 – Actual Attempt (non-fatal), and
- Category 10 – Completed Suicide.

Composite endpoints of suicidal ideation and suicidal behavior based on the above categories are defined below:

- **Suicidal ideation:** A "yes" answer at any time during study to any 1 of the 5 suicidal ideation questions (Categories 1 through 5) on the C-SSRS.
- **Suicidal behavior:** A "yes" answer at any time during study to any 1 of the 5 suicidal behavior questions (Categories 6 through 10) on the C-SSRS.
- **Suicidal ideation or behavior:** A "yes" answer at any time during study to any 1 of the 10 suicidal ideation and behavior questions (Categories 1 through 10) on the C-SSRS.

A listing contains data for each participant with suicidal ideation, suicidal behavior, or non-suicidal self-injurious behavior during the study by treatment and visit. Data from all visits are displayed, regardless of a "yes" or "no" answer, for patients with any "yes" answer for suicidal ideation, suicidal behavior, or self-injurious behavior without suicidal intent.

6.13.3.11. Malignancy

Treatment-emergent malignancy will be considered an AESI. The counts and percentages of participants with treatment emergent malignancy will be summarized by treatment and PT with decreasing order. Detailed searching criteria can be found in Appendix 1.

6.13.3.12. Renal Safety

Laboratory measures related to renal safety will be analyzed as specified for laboratory measurements in Section 6.13.6.

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).

Mixed model repeated measure analyses for eGFR and UACR will be provided. Log transformation will be performed for UACR.

6.13.3.12.1. Acute Renal Events

Because severe GI events may lead to dehydration, which could cause a deterioration in renal function including acute renal failure, dehydration events will be analyzed. Acute renal events associated with chronic renal failure exacerbation will also be captured.

Acute renal events from the following SMQ search will be considered as AESIs.

The counts and percentages of participants with acute renal events will be summarized by treatment by using the MedDRA PTs contained in any of the following SMQs:

- Acute renal failure: Narrow terms in Acute renal failure SMQ (20000003) and
- Chronic kidney disease: Narrow terms in Chronic kidney disease SMQ (20000213).

In addition, a listing of participants with treatment-emergent acute renal events may be provided, if deemed necessary.

6.13.3.12.2. Dehydration

Dehydration events will be captured in the Narrow terms in Dehydration SMQ (20000232). Severe/serious dehydration events will be considered AESIs.

A listing of participants with treatment-emergent dehydration events will be provided.

6.13.3.13. Thyroid Safety Monitoring

6.13.3.13.1. Calcitonin

Observed calcitonin data (a thyroid-specific laboratory assessment) will be summarized by treatment and nominal visit.

The counts and percentages of participants with a maximum postbaseline calcitonin value in the following thresholds will be provided by treatment and baseline calcitonin value (≤20 ng/L, >20 ng/L to ≤35 ng/L, >35 ng/L). Postbaseline: ≤20 ng/L, >20 ng/L to ≤35 ng/L, >35 ng/L to ≤50 ng/L, >50 ng/L to ≤100 ng/L, and >100 ng/L.

6.13.3.13.2. C-Cell Hyperplasia and Thyroid Malignancies

Treatment-emergent thyroid malignancies and C-Cell hyperplasia will be considered AESIs. Detailed searching criteria can be found in Appendix 1.

The counts and percentages of participants with treatment-emergent thyroid C-cell hyperplasia and malignancies will be summarized by treatment and PT ordered with decreasing frequency. In addition, a listing of participants with treatment-emergent thyroid C-cell hyperplasia and neoplasms may be provided if deemed necessary.

6.13.3.14. Treatment-Emergent Arrhythmias and Cardiac Conduction Disorders

Treatment-emergent arrhythmias and cardiac conduction disorders will be considered as AESIs.

The treatment-emergent arrhythmias and cardiac conduction disorders events will be included using the MedDRA PTs. Detailed searching criteria can be found in Appendix 1.

The counts and percentages of participants with treatment emergent arrhythmias and cardiac conduction disorders will be summarized by treatment and PT nested within SMQ. The PT will be ordered with decreasing frequency within SMQ. A listing of participants with treatment-emergent arrhythmias and cardiac conduction disorders may be provided if deemed necessary.

6.13.3.15. Overdose

A listing of patients reporting AEs related to overdosing of tirzepatide will be provided.

6.13.3.16. Abuse Liability

The counts and percentages of participants with treatment emergent potential abuse liability events will be summarized by treatment group with decreasing frequency. Detailed searching criteria can be found in Appendix 1.

6.13.4. Vital Signs

In the case that multiple records of an individual vital sign are collected at the same visit, they will be averaged prior to being used for data summaries and analyses.

Descriptive summaries by treatment and by nominal visit will be provided for baseline and postbaseline values as well as change from baseline values.

An MMRM and/or an ANCOVA model as in Section 6.13 might be conducted if necessary.

Counts and percentages of participants with treatment-emergent abnormal sitting SBP, sitting DBP, and pulse will be presented by treatment for participants who have both baseline and at least 1 postbaseline result. A treatment-emergent high result is defined as a change from a value less than or equal to the high limit at baseline to a value greater than the high limit at any time that meets the specified change criteria during the postbaseline period. A treatment-emergent low result is defined as a change from a value greater than or equal to the low limit at baseline to a value less than the low limit at any time that meets the specified change criteria during the postbaseline period. To assess decreases, change from the minimum value during the baseline period to the minimum value during the postbaseline period will be used. To assess increases, changes from the

maximum value during the baseline period to the maximum value during the postbaseline period will be used. Both planned and unplanned measurements will be included in the analysis. The criteria for identifying patients with treatment-emergent vital sign abnormalities are stated in Table GPIA.6.9.

Table GPIA.6.9. 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

In addition, following analyses will be conducted by treatment:

- counts and percentages of participants who had resting heart rate changes from baseline at 2 consecutive visits of more than 10 bpm and/or 20 bpm
- counts and percentages of participants who had resting heart rate maximum changes from baseline at visit <20 bpm or ≥20 bpm
- counts and percentages of participants who had at least 1 resting heart rate exceeding 100 bpm, and
- counts and percentages of participants who had at least 1 resting heart rate exceeding 100 bpm occurring at 2 consecutive study visits.

6.13.5. Electrocardiograms

For electrocardiogram (ECG) parameters that collect in triplicates (ie, Visit 3 and Visit 12), the average from the 3 measurements for the same parameter at the same visit will be calculated and used for all the subsequent summaries and analyses.

Summary statistics by treatment and by nominal visit will be provided for ECG parameters (heart rate, PR, QRS, QT, and QT corrected using Fredericia's correction factor [$QTcF = QT / RR^{0.333}$]). When the QRS is prolonged (eg, a complete bundle branch block), QT and QTc should be used to assess ventricular repolarization. Thus, for a particular ECG, the following will be set to missing (for analysis purposes) when QRS is ≥120 msec: QT and QTcF.

Change from baseline to postbaseline values for ECG parameters will be summarized for patients who have both a baseline and at least 1 postbaseline result. Only planned measurements will be included in the mean change analyses.

The criteria for identifying participants with treatment-emergent quantitative ECG abnormalities is based on Table GPIA.6.10.

The counts and percentages of participants who meet following criteria at any time during the entire study period (including the off-drug follow up time period) will be summarized by treatment group:

- treatment-emergent ECG abnormalities as listed in Table GPIA.6.10
- QT greater than 500 msec
- QTcF greater than 500 msec, and
- treatment-emergent increase from the maximum baseline in QTcF interval of greater than 30 msec, 60 msec, or 75 msec. Maximum baseline will be the maximum nonmissing observation in the baseline period. The maximum value during the treatment period will be analyzed. Scheduled and unscheduled measurements will be included.

Table GPIA.6.10. Selected Categorical Limits for ECG Data

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

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

6.13.6. Clinical Laboratory Evaluation

Limits from the performing lab will be used to define low (L) and high (H). 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. The associated descriptive will be presented in System International (SI) units and in conventional (CN) units.

Observed and change from baseline values for each visit may be displayed in plots for participants who have both a baseline and at least 1 postbaseline planned measurement. Baseline will be the last nonmissing observation prior to taking first study drug. Unplanned measurements will be excluded from plots.

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

For qualitative laboratory analytes, the number and percentage of participants 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 participant identification, treatment group, laboratory collection date, study day, analyte name, and analyte finding.

The MMRM model or ANCOVA (if MMRM model is not applicable) will be used for the analysis during the treatment period for the continuous measurements for selected lab tests.

6.14. Health Outcomes

The patient-reported outcome questionnaires will be analyzed using the mITT population on the EAS, unless specified otherwise.

Item-level missingness is dealt with as per the instrument developers' instruction.

Additional psychometric analyses will be performed by Global Patient Outcomes Real World Evidence at Lilly and documented in a separate analysis plan.

6.14.1. Patient Global Impression of Status for Physical Activity

The counts and percents of participants for Patient Global Impression of Status for Physical Activity (PGIS) response categories at each time point will be summarized by nominal visit and by treatment. For time points >52 weeks, only the participants with prediabetes at randomization will be included so that the denominator for percents will be participants with prediabetes at randomization only. A shift table from baseline to postbaseline of 5 PGIS response categories will be created at each postbaseline visit.

6.14.2. Short-Form-36 Health Survey Version 2, Acute Form

Per copyright owner, the QualityMetric Health Outcomes™ Scoring (PRO_CoRe V2.0) Software will be used to derive the following domain and component scores:

- Mental Component Score (MCS)
- Physical Component Score (PCS)
- Physical Functioning domain (PF)
- Role-Physical domain (RP)
- Bodily Pain domain (BP)
- General Health domain (GH)
- Vitality domain (VT)
- Social Functioning domain (SF)
- Role-Emotional domain (RE), and
- Mental Health domain (MH).

The following analyses for the actual value and change from baseline value for each domain and component score will be conducted:

- descriptive summaries by treatment group, and
- analysis described in Table GPIA.6.4.

If data allowed, analysis for SF-36 physical function domain score analysis described in Table GPIA.6.4 and in Table GPIA.6.4 will be conducted to evaluate the treatment effect in participants who have impaired physical function at baseline, which is defined as PGI-S response at baseline of "moderately limited", "very much limited," or "extremely limited". The empirical cumulative distribution function (eCDF) curves of the change from baseline to Week-52 in SF-36 physical function domain score will be provided by treatment group.

6.14.3. Impact of Weight on Quality of Life-Lite Clinical Trials

The following parameters will be included from IWQOL-Lite-CT:

- IWQOL Lite CT total score (all items: items 1 through 20)
- Physical Function composite score (5 items: items 1 through 3, 16, 17)
- Physical composite score (7 items: item 1 through 5, 16, 17), and
- Psychosocial composite score (13 items: item 6 through 15, 18, 19, 20).

IWQOL-Lite-CT total and composite scores range from 0 to 100, with higher scores reflecting better levels of functioning.

IWQOL-Lite-CT scores are computed according to the IWQOL-Lite scoring rules (Kolotkin et al. 2002) as following:

- Each composite raw score will be calculated if a minimum of 50% of the items for that composite has a non-missing value; the total score will be calculated if a minimum of 75% of all 20 items has a non-missing value.
    - physical composite score: 4 of 7 items
    - physical function composite score: 3 of 5 items
    - psychosocial composite score: 7 of 13 items
    - IWQOL Lite CT total score: 15 of 20 items
- If the minimum required number of items are answered for a composite then:
    - Calculate the average of the valid non-missing responses corresponding to the items in the total or each composite (1 = "never" or "not at all true" and 5 = "always" or "completely true").
    - The composite score is then calculated by transforming the raw composite score to the 0 (worst)-to-100 (best) metric using the following formula for every participant at each time point:

$$100\ (S_{max} - C_{avg})/(S_{max} - S_{min})$$

- C_{avg} is the raw average score of all nonmissing item responses in the composite; this average must be a number between 1 and 5, inclusive
- S_{max} is the maximum possible raw score value (that is, 5)
- S_{min} is the minimum possible raw score value (that is, 1)
- Inserting the maximum and minimum possible score values, the formula is reduced to $100\ (5 - C_{avg})/4$.

For total and each composite score, the actual value and change from baseline value following analyses will be conducted:

- descriptive summaries by treatment group and
- ANCOVA analysis described in in Table GPIA.6.4

6.14.4. EQ-5D-5L

For the utility score and the Visual Analog Scale (VAS) scores, following analyses of the actual value and change from baseline value will conducted:

- descriptive summaries by treatment group and
- analysis of covariance described in Table GPIA.6.4.

6.15. Subgroup Analyses

Efficacy subgroup analyses will be guided by the efficacy estimand in EAS.

6.15.1. Subgroup Analysis of Body Weight Change

Subgroup analyses by the following baseline characteristics will be provided:

- age group (<65, ≥65 years)
- sex
- body mass index group (<28, ≥28 kg/m^2),
- body mass index group2 (<30, ≥30 and <35, ≥35 kg/m2), and
- presence of comorbidities at randomization (Y vs N).

The outcome measures for the subgroup analyses will include:

- percent change in body weight from randomization at 52 weeks and
- percentage of participants achieving at least 5% body weight reduction at 52 weeks.

For the percentage change in body weight from randomization at 52 weeks, for each subgroup analyses aforementioned, the following 2 models will be conducted:

- Conduct MMRM model on the subgroup only with terms of treatment group, visit, treatment-by-visit-interaction, sex, and presence of comorbidities at randomization as fixed effects, and baseline body weight as a covariate. Variance-covariance structure for within-patient errors will be same as Section 6.12.1.1
- Full MMRM model: treatment group, visit, subgroup, treatment-by-visit-interaction, treatment-by-subgroup-interaction, subgroup-by-visit-interaction, treatment-visit-subgroup-interaction, sex, and presence of comorbidities at randomization as fixed effects, and baseline body weight as a covariate. Variance-covariance structure for within-patient errors will be same as Section 6.12.1.1.

For the percentage of participants achieving at least 5% body weight reduction at 52 weeks, for each subgroup analyses aforementioned, the following 2 models will be conducted:

- Conduct logistic regression model on the subgroup only with terms of treatment group, sex, and presence of comorbidities at randomization as fixed effects, and baseline body weight as a covariate. Missing body weight measurement at 52 weeks will be imputed by the predicted value from MMRM model on the subgroup aforementioned, then the continuous measurements will be dichotomized into status of achieving at least 5% body weight reduction (Yes or No).
- Conduct logistic regression model with terms of treatment group, subgroup, treatment-by-subgroup-interaction, sex, and presence of comorbidities at randomization as fixed effects, and baseline body weight as a covariate. Missing body weight measurement at 52 weeks will be imputed by the predicted value from full MMRM model aforementioned, then the continuous measurements will be dichotomized into status of achieving at least 5% body weight reduction (Yes or No).

7. Unblinding Plan

Details of the blinding and unblinding will be provided in Blinding and Unblinding Plan document for Study GPIA.

8. COVID-19 Pandemic Impact

This section lists additional statistical analyses that may be performed at the primary database lock and final database lock to assess the impact of COVID-19 pandemic if the data warrants.

8.1. General Consideration

Percentage and count of randomized participants who followed the COVID-19 mitigation plan may be summarized by treatment group. This includes, but not limited to, participants rescreened, procedures conducted via remote visit or mobile home health visit, visits occurred using the extended visit windows, alternative way of investigator product shipment/dispensing, use of a local lab, etc. A listing of randomized participants who followed the COVID-19 mitigation plan may be provided. Similar analyses may be provided by country and by treatment group.

Percentage and count of randomized participants whose study visits were impacted by COVID-19 pandemic may also be summarized. A listing may be provided.

8.2. Exposure

A listing of randomized participants who had study drug temporary interruption due to COVID-19 pandemic may be provided.

8.3. Protocol Deviation

Percentage and count of randomized participants having important protocol deviation related to COVID-19 pandemic will be summarized by treatment.

Percentage and count of randomized participants with protocol deviation related to COVID-19 pandemic may also be summarized by treatment.

A listing of all randomized participants who had important protocol deviation due to COVID-19 pandemic may be provided.

8.4. Patient Disposition

A summary table for all randomized participants that discontinue study or study treatment due to COVID-19 pandemic will be provided by treatment.

A listing of randomized participants who discontinued the study or study treatment due to COVID-19 infection will be provided.

8.5. Adverse Events

A listing of randomized participants who had COVID-19 infection, including death due to COVID-19, during the post-randomization period will be provided.

8.6. Major Depressive Disorder/Suicidal Ideation

The counts and percentages of participants with TEAEs for major depression may be summarized by treatment group using MedDRA PT nested within SMQ by COVID-19 subgroup (that is, participants without impact versus with impact) for SS group.

A participant is defined as impacted by COVID-19 if either one of the following is satisfied:

- no COVID-19 illness, but impacted by quarantine and travel restrictions, clinics closing, visits being canceled, delay or nondelivery of the investigational product, virtual visits, etc.

OR

- with COVID-19 illness.

The suicidal ideation and behavior solicited through C-SSRS may be summarized by treatment group by COVID-19 subgroup (ie, participants without impact vs with impact) for SS group.

8.7. Missing Data Due to COVID-19

For the primary endpoints and key secondary endpoints, missing data due to COVID-19 will be handled as described in Section 6.12.1.3. In addition, a summary table for participants whose primary or key secondary measurements were impacted by COVID-19 (including missing, collected using alternative options) may be provided. A listing of participants whose primary or key secondary measurements were impacted by COVID-19 (including missing, collected using alternative options) may be provided if deemed necessary.

9. References

[ADA] American Diabetes Association. 2. Classification and diagnosis of diabetes: standards of medical care in diabetes-2019. *Diabetes Care*. 2019a; 42(suppl 1):S 13-S28. https://doi.org/10.2337/dc19-S002

[ADA] American Diabetes Association. 6. Glycemic targets: standards of medical care in diabetes-2019. *Diabetes Care*. 2019b;42(suppl 1):S61-S70. https://doi.org/10.2337/dc19-S006

Anwar YA, White WB. Ambulatory monitoring of the blood pressure devices, analysis, and clinical utility. In: William White, editor. *Blood Pressure Monitoring in Cardiovascular Medicine and Therapeutics*. Humana Press, Inc; 2001:p 64.

Armato JP, DeFronzo RA, Abdul-Ghani M, Ruby RJ. Successful treatment of prediabetes in clinical practice using physiological assessment (STOP DIABETES). *Lancet Diabetes Endocrinol*. 2018;6(10):781-789. https://doi.org/10.1016/S2213-8587(18)30234-1

Ernst ME, Bergus GR. Noninvasive 24-hour ambulatory blood pressure monitoring: overview of technology and clinical applications. *Pharmacotherapy*. 2002;22(5):597-612. https://doi.org/10.1592/phco.22.8.597.33212

Pickering TG, Hall JE, Appel LJ, et al. Recommendations for blood pressure measurement in humans and experimental animals. Part 1: blood pressure measurement in humans: a statement for professionals from the Subcommittee of Professional and Public Education of the American Heart Association Council on High Blood Pressure Research. *Hypertension*. 2005;45(1):142-161. https://doi.org/10.1161/01.HYP.0000150859.47929.8e

Kolotkin RL, Crosby RD, Rhys Williams G. Health-related quality of life varies among obese subgroups. *Obes Res*. 2002;10(8):748-756. https://doi.org/10.1038/oby.2002.102

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

10. Appendices

Appendix 1. Searching Criteria for Special Safety Topics

Abuse Liability

To identify AE terms suggestive of potential abuse liability, narrow terms from SMQ of Drug abuse and dependence (20000101) will be used.

Acute Gallbladder Disease

All biliary colic, cholecystitis, or other suspected events related to gallbladder disease will be identified using the MedDRA PTs in any of the following:

- Narrow PTs in Gallbladder related disorders SMQ (20000124)
- Narrow PTs in Biliary tract disorders SMQ (20000125), and
- Narrow PTs in Gallstone related disorders SMQ (20000127).

Amputation/Peripheral Revascularization

Amputations/peripheral revascularization events will be identified using the following MedDRA PTs:

- Amputation
- Peripheral revascularization.

C-cell Hyperplasia and Thyroid Malignancies

Thyroid malignancies and C-Cell hyperplasia will be identified using MedDRA HLT for Thyroid neoplasms and PT for thyroid C-cell hyperplasia.

Hepatic Treatment-Emergent Adverse Events

Treatment-emergent potentially drug-related hepatic disorders will be identified using the MedDRA PTs contained in any of the following:

- Broad and narrow terms in the Liver related investigations, signs and symptoms SMQ (20000008)
- Broad and narrow terms in the Cholestasis and jaundice of hepatic origin SMQ (20000009)
- Broad and narrow terms in the Hepatitis non-infections SMQ (20000010)
- Broad and narrow terms in the Hepatic failure, fibrosis and cirrhosis and other liver damage SMQ (20000013)
- Narrow terms in the Liver-related coagulation and bleeding disturbances SMQ (20000015)

- Narrow PTs in Gallbladder related disorders SMQ (20000124)
- Narrow PTs in Biliary tract disorders SMQ (20000125); and
- Narrow PTs in Gallstone related disorders SMQ (20000127).

Hypersensitivity Reactions

Analyses are based on the following:

- Narrow and algorithm terms in Anaphylactic reaction SMQ (20000021)
- Narrow terms in Angioedema SMQ (20000024)
- Narrow terms in Severe cutaneous adverse reactions SMQ (20000020)
- Narrow terms in Hypersensitivity SMQ (20000214).

For the Anaphylactic reaction SMQ, each term is classified by scope (Narrow, Broad) and by category (A, B, C, D). All Narrow terms are category A, and all Broad terms are category B, C, or D. In addition to the usual Narrow and Broad searches, the SMQ defines an algorithm to further refine the cases of interest. For Time Period A analysis, the Anaphylactic reaction SMQ algorithm will be included. The algorithm is based upon events that occur within Time Period A. The counts and percentages of participants who experienced a TEAE for the following will be analyzed for each of the 2 time periods:

- any narrow term from any one of the 4 SMQs indicated above (that is, combined search across narrow of all 4 SMQs)
- any narrow scope term within each SMQ, separately (that is, narrow SMQ search). For Time Period A analysis, any term from Anaphylactic reaction SMQ algorithm.

Injection Site Reactions

The ISR AE will be identified using the MedDRA PT in any of the following:

- HLT of Injection site reaction
- HLT of Administration site reaction
- HLT of Infusion Site Reactions

Pancreatitis Events

Determination of investigator-reported events will be through the “Acute pancreatitis” MedDRA SMQ (20000022, narrow scope) and a “Chronic pancreatitis” PT search of the AE database, while adjudication-confirmed pancreatitis are found from adjudication forms.

Malignancy

The malignancy events will be identified using the MedDRA PT contained in Malignant tumours SMQ (20000194) narrow scope or Tumours of unspecified malignancy SMQ (20000195) narrow scope.

Major Depressive Disorder/Suicidal Ideation

AEs will be searched using MedDRA PTs from SMQs narrow scope: 20000037 [Suicide/self-injury] and 20000167 [Depression (excl suicide and self injury)]).

All Arrhythmias and Cardiac Conduction Disorders

All the arrhythmias and cardiac conduction disorders events will be identified using the MedDRA PT contained in any of the following SMQs:

1) All Arrhythmias:
    - For symptoms: Arrhythmia related investigations, signs and symptoms SMQ (20000051), narrow and broad terms
    - For-all arrhythmias: In Cardiac arrhythmia SMQ, under tachyarrhythmia sub SMQ
        - Supraventricular tachyarrhythmia SMQ (20000057), broad and narrow terms
        - Tachyarrhythmia terms, nonspecific SMQ (20000164), narrow terms only; and
        - Ventricular tachyarrhythmia SMQ (20000058), narrow terms only.
2) Cardiac Conduction Disorders
    - Conduction defects SMQ (20000056), narrow terms only; and
    - Cardiac conduction disorders High Level Term (HLT; 10000032), all PTs.

Signature Page for VV-CLIN-076865 v1.0

Approval	PPD 04-Jan-2023 03:48:10 GMT+0000

Signature Page for VV-CLIN-076865 v1.0