Statistical Analysis Plan 3: I8F-MC-GPHJ

A Phase 3b, Randomized Controlled Study to Evaluate the Efficacy and Safety of Tirzepatide Compared to Semaglutide in Adults Who Have Obesity or Overweight with Weight-Related Comorbidities (SURMOUNT-5)

NCT05822830

Approval Date: 23-Oct-2024

Title Page

Protocol Title: A Phase 3b, Randomized Controlled Study to Evaluate the Efficacy and Safety of Tirzepatide Compared to Semaglutide in Adults Who Have Obesity or Overweight with Weight-Related Comorbidities (SURMOUNT-5)

Protocol Number: I8F-MC-GPHJ

Compound Number: LY3298176

Short Title: Efficacy and Safety of Tirzepatide Compared to Semaglutide in Adults Who Have Obesity or Overweight

Acronym: SURMOUNT-5

Sponsor Name: Eli Lilly and Company

Legal Registered Address: Indianapolis, Indiana, USA 46285

Regulatory Agency Identifier Number(s)

Registry	ID
IND:	139721
EU trial number:	2022-501106-35-00

Document ID: VV-CLIN-131848

Table of Contents

Version history

The Statistical Analysis Plan (SAP) Version 1 for study I8F-MC-GPHJ was based on the protocol amendment (a) dated 29Jul2022. The SAP Version 2 and Version 3 were based on the protocol amendment (c) dated 24 Aug 2023.

Table 1 SAP Version History Summary

SAP Version	Approval Date	Change	Rationale
1.0	17 Apr 2023	Not Applicable	Original version
2.0	2 May 2024	Page 1 Protocol Title changed	Protocol Amendment (c)
		Section- 1.1 Added text "or MTD (1.7 mg or 2.4 mg)" with 'semaglutide 2.4 mg' in Primary objective.	Protocol Amendment (c)
		Section- 1.1 Added text "or MTD (1.7 mg or 2.4 mg)" with 'semaglutide 2.4 mg' in Key Secondary Objective.	Protocol Amendment (c)
		Section- 1.1 Added text "or MTD (1.7 mg or 2.4 mg)" with 'semaglutide 2.4 mg' in Additional Secondary Objective..	Protocol Amendment (c)
		Section- 1.1 New Additional Secondary objective and Endpoint added.	Protocol Amendment (c)
		Section- 1.1 Added text "or MTD (1.7 mg or 2.4 mg)" with 'semaglutide 2.4 mg' in Tertiary Objective.	Protocol Amendment (c)

SAP Version	Approval Date	Change	Rationale
		Section- 1.1 Added text "or MTD (1.7 mg or 2.4 mg)" with 'semaglutide 2.4 mg' in Tertiary (Not specified in Protocol) Objective.	Protocol Amendment (c)
		Section- 1.2 Added text "or MTD (1.7 mg or 2.4 mg)" with 'semaglutide 2.4 mg' in Study Design	Protocol Amendment (c)
		Section- 1.2 Removed text "and a 4-week safety follow-up period" in Study design	Protocol Amendment (c)
		Section- 1.2 Updated study schema.	Protocol Amendment (c)
		Section- 1.2 Added text "or MTD (1.7 mg or 2.4 mg)" with 'semaglutide 2.4 mg' in description of 'End of Visit 2 to Visit 8'.	Protocol Amendment (c)
		Section- 1.2 Removed text "Safety Follow Up Period" in Study design	Protocol Amendment (c)
		Section-1.2 Added percent to the Population-level summary for Secondary Estimand	Clarification
		Section- 2 Added text "or MTD (1.7 mg or 2.4 mg)" with 'semaglutide 2.4 mg' in Statistical Hypotheses	Protocol Amendment (c)

SAP Version	Approval Date	Change	Rationale
		Section- 3 Replaced the text 'Treatment Period plus Safety Follow up' with "Treatment Period and any follow up data" in definition of Safety Analysis Set (SS) and Safety Analysis Set (SS) Excluding Other AOMs or Bariatric Surgery	Protocol Amendment (c)
		Section-3 Row added to define Efficacy Analyses Set for High Dose (EAS-HD)	Protocol Amendment (c)
		Section- 4.1 Replaced the text "Safety follow up visit (Visit 801)" with "last study visit" in General Considerations section.	Protocol Amendment (c)
		Section- 4.3.1 Added text "or MTD (1.7 mg or 2.4 mg)" with 'semaglutide 2.4 mg' in description.	Protocol Amendment (c)
		Section- 4.3.2 Added text "or MTD" with 'semaglutide 2.4 mg' in description.	Protocol Amendment (c)
		Section- 4.4.2 Added text "Change from baseline in weight (%) in participants who reached 15 mg of tirzepatide or 2.4 mg of semaglutide and did not have a dose decreased starting at Week 48 until the end of the study." In Additional Secondary Endpoint.	Protocol Amendment (c)

SAP Version	Approval Date	Change	Rationale
		Section- 4.4.2 Analysis method and new Estimand added for analysis of percent change from baseline in body weight at 72 weeks for tirzepatide at 15 mg versus semaglutide at 2.4 mg	Protocol Amendment (c)
		Section- 4.4.2.1 Added text "Change from baseline in weight (%) for comparing tirzepatide 15 mg versus semaglutide 2.4 mg will also be analyzed using the treatment-regimen estimand using the same statistical methods as the primary endpoint (see Section 4.3.1)." In Analysis for Continuous Outcomes.	Protocol Amendment (c)
		Section- 4.4.2.2 Section modified after removal old and addition of new text. Removed text "with missing values imputed by the predicted value from the MMRM model specified in Section 4.3.2 and dichotomized data will then be derived based on continuous imputed values. After dichotomizing, the data is analyzed using a logistic regression model same as section 4.4.1.1." Text added "following the methodology specified in section 4.4.1.1 for the efficacy estimand"	Protocol Amendment (c)

SAP Version	Approval Date	Change	Rationale
		Section- 4.4.2.2 Added text "or MTD (1.7 mg or 2.4 mg)" with 'semaglutide 2.4 mg' in first row of Additional Secondary Endpoint.	Protocol Amendment (c)
		Section- 4.4.2.2 New row added for objective "compare tirzepatide at 15 mg to semaglutide at 2.4 mg for the mean percent decrease in weight loss" in Additional Secondary Endpoint.	Protocol Amendment (c)
		Section- 4.5.1.1 Added text "or MTD (1.7 mg or 2.4 mg)" with 'semaglutide 2.4 mg' in Tertiary Objective.	Protocol Amendment (c)
		Section- 4.5.1.2 Added text "or MTD (1.7 mg or 2.4 mg)" with 'semaglutide 2.4 mg' in Tertiary (Not specified in Protocol) Objective.	Protocol Amendment (c)
		Section- 4.6 Replaced the text "Safety follow up" with "last study visit" in Safety Analyses section.	Protocol Amendment (c)
		Section- 4.6.2.10 Removed text "including follow up period" from Liver Enzymes section.	Protocol Amendment (c)
		Section- 4.6.2.18 Analysis modified in the section.	Update in the analysis.

SAP Version	Approval Date	Change	Rationale
		Section- 4.6.3 Added text “or more” after “If 2” and removed text “geographical area” from Vital Signs section.	Protocol Amendment (c)
		Section- 4.9 Removed text “Due to changes in the countries where the study will be enrolled, country was removed as a stratification factor. Therefore, country has been removed as a factor in the statistical model. Instead, BMI category was added as a stratification factor and was added to the statistical models specified in this SAP. Due to the new FDA draft guidance for conditional versus unconditional treatment effects, the methodology for the analysis of binary data in section 4.4.1.1. was updated.” From Changes to Protocol-Planned Analyses.	Protocol Amendment (c)
3.0	See Date on Page 1	Section 1.1: Removed the words “Coprimary Estimand” which is not applicable.	Consistency with protocol.
		Section 3: Define a single mITT population regardless of type of analysis.	Include participants who are inadvertently enrolled for both efficacy and safety analyses.
		Section 3: Removed Safety Analysis Set excluding other AOMs, bariatric surgery or other weight loss procedures.	This analysis set is not used for any pre-specified analysis.
		Section 4.1: Removed analysis methods not used in the study.	Consistency with the planned analyses.

SAP Version	Approval Date	Change	Rationale
		Section4.3.1: Add language to clarify that participants who switch to non-study treatment are included in the treatment discontinuation category.	Clarification
		Section 4.3.1: Removed caps for imputed data	Detail information in separate document
		Section 4.4.1.1: Updated references for FDA guidance and delta method	Use most current information.
		Section 4.4.1.2: Remove caps for imputed data.	Detail information in separate document
		Section 4.5.1.1: Added log-transformation for lipid parameters and insulin.	Consistency with recommended transformation for these endpoints.
		Section 4.5.1.2: Added log-transformation for additional lipid parameters.	Consistency with recommended transformation for these endpoints.
		Section 4.6.1: Removed option not used in CRF	Clarification
		Section 4.6.2.2: Added covariates to MMRM model	Consistency with MMRM model in efficacy section
		Section 4.6.2.4: Added ‘maximum’ to calcitonin.	Clarification
		Section 4.6.2.6.: Remove ventricular.	Consistency with reporting all arrhythmias.
		Section 4.6.2.10: Added Maximum to Maximum to alkaline phosphatase.	Clarification

SAP Version	Approval Date	Change	Rationale
		Section 4.6.3: Clarify that only pulse will be analyzed with MMRM model.	Other vital parameters are analyzed with MMRM model for efficacy analyses.
		Section 4.6.2.9: Created separated paragraph for hepatic events and gallbladder disease.	Consistency with how summaries are created.
		Section 4.6.2.12: Include planned eGFR analyses.	Consistency with planned outputs.
		Section 4.6.2.15 and Section 4.6.2.17: Removed	Consistent with updated AESIs for indication.
		Section 4.6.3: Added BMI category to MMRM model.	Clarification
		Section 4.7.2: Added MMRM model to subgroup analysis.	Clarification
		Section 6.1: Updated text to be consistent with previous updates.	Consistency
		Section 6.3: Removed age ranges not applicable in the study	Clarity
		Section 6.4: Update location of AESI specifications	Clarity
		Section 6.5: Added appendix to provide clarity for concomitant therapy and procedures of interest.	Clarity
		Section 7: Updated references	Consistency

1. Introduction

This SAP describes the pre-specified statistical analyses for study GPHJ. These analyses apply to efficacy and safety data.

Changes to the protocol-planned analyses are described in Section 4.9.

1.1. Objectives, Endpoints, and Estimands

Objectives	Endpoints
Primary[a]	
To demonstrate that tirzepatide at 15 mg or MTD (10 mg or 15 mg) is superior to semaglutide 2.4 mg or MTD (1.7 mg or 2.4 mg) for the mean percent decrease in weight loss	Percent change from baseline in body weight at 72 weeks
Key Secondary[a]	
To demonstrate that tirzepatide at 15 mg or MTD (10 mg or 15 mg) is superior to semaglutide 2.4 mg or MTD (1.7 mg or 2.4 mg) at Week 72 for	
• Body weight	Body weight reduction from baseline of • ≥10% • ≥15% • ≥20%, and • ≥25%
• Waist circumference	Change from baseline in waist circumference (cm)
Additional Secondary	
To compare tirzepatide 15 mg or MTD (10 mg or 15 mg) to semaglutide 2.4 mg or MTD (1.7 mg or 2.4 mg) at Week 72 for	
• Body weight	Change from baseline in body weight (kg) Body weight reduction from baseline of ≥30%
• BMI	Change from baseline in BMI (kg/m^2)
To compare tirzepatide at 15 mg to semaglutide at 2.4 mg for the mean percent decrease in weight loss	Percent change from baseline in body weight at 72 weeks
Tertiary	
To compare tirzepatide 15 mg or MTD (10 mg or 15 mg) to semaglutide 2.4 mg or MTD (1.7 mg or 2.4 mg) at Week 72 for	
• Lipid parameters	Change from baseline in • triglycerides (mg/dL) • VLDL (mg/dL) • non-HDL cholesterol (mg/dL)
• Insulin	Change from baseline in fasting insulin (pmol/L)
• HbA1c	Change from baseline in HbA1c (%)

Objectives	Endpoints
Tertiary (Not specified in Protocol)	
To compare tirzepatide 15 mg or MTD (10 mg or 15 mg) to semaglutide 2.4 mg or MTD (1.7 mg or 2.4 mg) at Week 72 for	
• Blood pressure parameters	Change from baseline in • Systolic blood pressure (SBP) • Diastolic blood pressure (DBP)
• Lipid parameters	Change from baseline in • LDL (mg/dL) • HDL (mg/dL)

Abbreviations: BMI = body mass index; HDL = high-density lipoprotein; LDL = low-density lipoprotein; MTD = maximum tolerated dose; SF36 physical function = Short Form 36 physical function; VLDL = very low-density lipoprotein.

[a] Primary and key secondary endpoints are controlled for multiplicity.

Primary estimand

The primary estimand evaluated in this study is a modified treatment-regimen estimand. This estimand aims at reflecting how participants with obesity or overweight with at least 1 weight-related comorbid condition are treated in clinical practice and takes into account both tolerability and efficacy.

This modified treatment-regimen estimand answers the following question of interest for the primary objective: What is the treatment difference in mean percent change from baseline in body weight after 72 weeks of treatment as an adjunct to a reduced-calorie diet and increased physical activity in participants with obesity or overweight with at least 1 weight-related comorbid condition*? Under this estimand framework, participants will be included regardless of treatment discontinuation for any reason and regardless of initiation of other anti-obesity medication (except for participants who are randomized to semaglutide and start taking non-study tirzepatide or for patients who are randomized to tirzepatide and start taking non-study semaglutide). This estimand also assumes that participants who had bariatric surgery or another weight loss procedure did not get any benefit or might have worsened from their randomized study treatment.*

The *modified treatment-regimen estimand is described by the following attributes:*

- *Population*: Participants with obesity (BMI ≥30 kg/m^2) or overweight (BMI ≥27 kg/m^2) with at least 1 weight-related comorbid condition, but without Type 2 diabetes. Further details can be found in Section 5 of Protocol.
- *Endpoint: Percent change from baseline to Week 72 in body weight.*
- *Treatment cond*ition*: The randomized treatment as an adjunct to a reduced-calorie diet and increased physical activity regardless of adherence to treatment with or without other anti-obesity medications, except for non-study tirzepatide or semaglutide treatment*

switches. Further details on study treatment and concomitant therapy can be found in Section 6 of Protocol.

- *Intercurrent events: Intercurrent events of interest: "treatment discontinuation for any reason" and "initiation of other anti-obesity medication except for non-study tirzepatide or semaglutide" are addressed by the treatment condition. Bariatric surgery or other weight loss procedures will be addressed by the hypothetical strategy. It will be assumed that patients who undergo any of these procedures did not get any benefits or might have worsened from their randomized study treatment.*
- *Population-level summary: Difference in mean percent changes between treatment conditions.*

Secondary estimand(s)

The secondary estimand evaluated in this study is the efficacy estimand. This estimand focuses on the treatment effect if participants who underwent randomization continued to receive the study treatment without taking other anti-obesity therapies *such as* weight management drugs, bariatric surgery, or weight loss procedure. This estimand will be used in publications to inform prescribers or physicians and may be submitted to other regulatory agencies outside the United States.

The efficacy estimand answers the following question of interest for the primary objective: What is the treatment difference in percent change in body weight from baseline after Week 72 of treatment as an adjunct to a reduced-calorie diet and increased physical activity in participants with obesity or overweight with at least 1 weight-related comorbid condition assuming that participants had stayed on treatment and not taken other anti-obesity therapies (that is, other weight management drugs, bariatric surgery, or other weight management procedures)?

The efficacy estimand is described by the following attributes.

- *Population: Participants who have obesity (BMI ≥30 kg/m^2) or overweight (BMI ≥27 kg/m^2) with at least 1 weight-related comorbid condition, but without Type 2 diabetes. Further details can be found in Section 5 of Protocol.*
- *Endpoint: Percent change from baseline to Week 72 in body weight.*
- *Treatment condition: The randomized treatment as an adjunct to a reduced-calorie diet and increased physical activity. Further details on study treatment can be found in Section 6 of Protocol.*
- *Intercurrent event*s*: "Treatment discontinuation for any reason" and "Initiation of other anti-obesity therapies such as weight management drugs, bariatric surgery or other weight management procedures" will be addressed using the hypothetical strategy:*
    - Had *participants stayed on treatment.*
    - Had participants not taken other weight management drugs or bariatric surgery or weight loss procedure.
- *Population-level summary*: *Difference in mean percent changes between treatment conditions.*

Approved on 23 Oct 2024 GMT

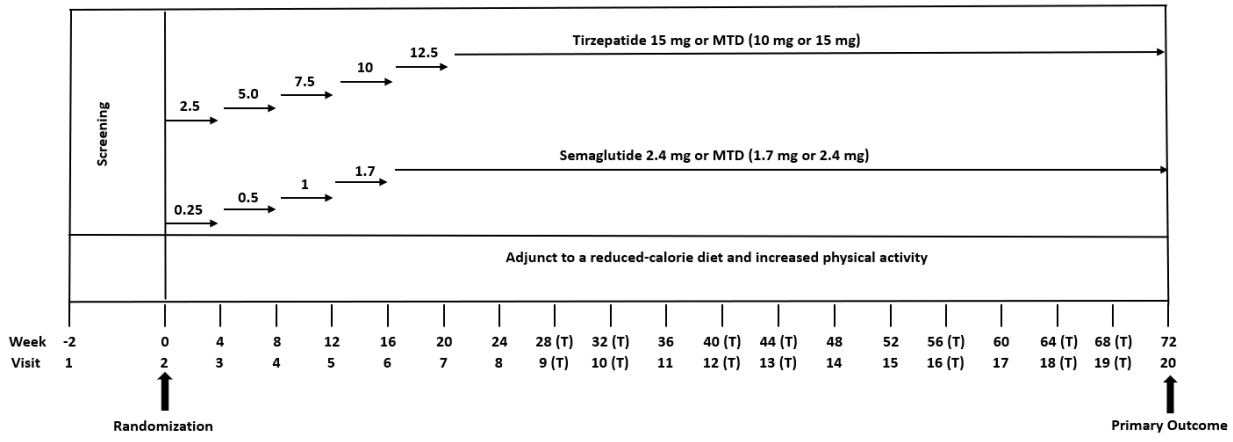
Both the efficacy and treatment-regimen estimands will be evaluated for the primary and all key secondary objectives. The population, treatment condition, intercurrent events, and population-level summary specified above for each estimand for the primary objective will also apply to the key secondary objectives. The endpoint for each key secondary objective is defined in the table above.

1.2. Study Design

Study GPHJ is a Phase 3b, multicenter, randomized, parallel-arm, open-label, comparator-controlled, 72-week trial that will investigate the efficacy and safety of tirzepatide 15 mg or MTD (10 mg or 15 mg) compared with semaglutide (2.4 mg) or MTD (1.7 mg or 2.4 mg) in study participants who have obesity (BMI ≥30 kg/m^2) or overweight (BMI ≥27 kg/m^2) with weight-related comorbidities.

Study GPHJ will consist of 2 periods: a 2-week screening period; and a 72-week, open-label tirzepatide and semaglutide treatment period (including a 20-week dose escalation period for tirzepatide and a 16-week dose escalation period for semaglutide). The study participants will be randomly assigned in a 1:1 ratio to tirzepatide 15 mg or MTD (10 mg or 15 mg) or semaglutide 2.4 mg or MTD (1.7 mg or 2.4 mg) at the end of the screening period. An upper limit of 70% enrollment of women will be used to ensure a sufficient enrollment of men.

Schema



Abbreviations: MTD = maximum-tolerated dose; (T) = Telephone Visit.

Figure 1.1 Illustration of study design for clinical protocol study I8F-MC-GPHJ.

Screening Period

Visit 1

The purpose of screening procedures at Visit 1 is to establish initial eligibility. The participant must sign the ICF before the study procedures are performed. Since some screening procedures need to be completed in the fasting state (at least 8 hours without eating, drinking [except water], or any significant physical activity), Visit 1 may be conducted over more than 1 day to ensure necessary conditions are met**.**

Treatment Period

Visit 2

At Visit 2, eligible participants will perform all required baseline study procedures, including the collection of all baseline laboratory measures and questionnaires, prior to enrollment and prior to **taking the first dose of study drug. Participants will be provided diaries and be trained to record key study information, as appropriate.**

End of Visit 2 to Visit 8

At the end of Visit 2, participants that meet all the entry criteria are enrolled in this study and randomized to either tirzepatide 15 mg or MTD (10 mg or 15 mg) or semaglutide 2.4 mg or MTD (1.7 mg or 2.4 mg) treatment arms. The date and time of the first dose of study drug will be recorded on the e-diary.

P**articipants will undergo either a 20-week tirzepatide or a 16-week semaglutide dose escalation.**

The starting dose of tirzepatide is 2.5 mg for 4 weeks, then the dose is increased by 2.5 mg every 4 weeks (2.5 to 5 to 7.5 to 10 to 12.5 to 15 mg) up to 15 mg or MTD (10 mg or 15 mg).

The starting dose of semaglutide is 0.25 mg for 4 weeks, then the dose is increased every 4 weeks (0.25 to 0.5 to 1.0 to 1.7 to 2.4 mg) to the 2.4 mg dose **or MTD (1.7 mg or 2.4 mg).**

End of Visit 8 to Visit 20

Participants complete all visit procedures.

Early discontinuation of treatment visit

Participants unable or unwilling to continue the study treatment for any reason will perform an early discontinuation of treatment visit**.**

Randomization and Blinding

This is an open-label study.

All participants will be centrally randomized in a 1:1 ratio using an IWRS.

Potential bias will be reduced by central randomization, and the randomization will be stratified by the following factors:

- Prediabetes Status (yes/no)
- Sex (female, male)
- Baseline BMI (<35 kg/m^2, ≥ 35 kg/m^2)

2. Statistical Hypotheses

The *null hypothesis corresponding to the primary objective of this study is as follows:*

- **$H_{1,0}$**: *tirzepatide at 15 mg or MTD (10 mg or 15 mg) is not superior to semaglutide 2.4 mg or MTD (1.7 mg or 2.4 mg) with respect to mean percent change in body weight from baseline at Week 72*

The *null hypotheses corresponding to the key secondary objectives is as follows:*

- **$H_{2,0}$**: *tirzepatide at 15 mg or MTD (10 mg or 15 mg) is not superior to semaglutide 2.4 mg or MTD (1.7 mg or 2.4 mg) for the percentage of participants who achieve ≥10% body weight reduction from baseline at Week 72*
- **$H_{3,0}$**: *tirzepatide at 15 mg or MTD (10 mg or 15 mg) is not superior to semaglutide 2.4 mg or MTD (1.7 mg or 2.4 mg) for the percentage of participants who achieve ≥15% body weight reduction from baseline at Week 72*
- **$H_{4,0}$**: *tirzepatide at 15 mg or MTD (10 mg or 15 mg) is not superior to semaglutide 2.4 mg or MTD (1.7 mg or 2.4 mg) for the percentage of participants who achieve ≥20% body weight reduction from baseline at Week 72*
- **$H_{5,0}$**: *tirzepatide at 15 mg or MTD (10 mg or 15 mg) is not superior to semaglutide 2.4 mg or MTD (1.7 mg or 2.4 mg) for the percentage of participants who achieve ≥25% body weight reduction from baseline at Week 72*
- **$H_{6,0}$**: *tirzepatide at 15 mg or MTD (10 mg or 15 mg) is not superior to semaglutide 2.4 mg or MTD (1.7 mg or 2.4 mg) with respect to change in waist circumference (cm) from baseline at Week 72*

Operationally*, the hypotheses will be evaluated by 2-sided tests.*

2.1. Multiplicity Adjustment

The primary and the key secondary objectives will be evaluated using both the treatment regimen and the efficacy estimands. Since the purpose for each estimand is different, multiplicity adjustment will be conducted for each estimand separately.

A prespecified graphical scheme (Bretz et al. 2009, 2011) will be implemented to control the family-wise error rate at a 1-sided alpha level of 0.025 for testing the null hypotheses stated in Section 2. More specifically, multiple testing adjusted p-values described by Bretz et al. (2009) will be calculated, and any hypothesis tests with a multiple testing adjusted 1-sided p-value of less than 0.025 will be considered statistically significant. This graphical approach is a closed testing procedure; hence, it strongly controls the family-wise error rate across all endpoints (Bretz et al. 2009, 2011; Alosh et al. 2014).

Unless otherwise specified, there will be no adjustment for multiple comparisons for any other analyses outside the primary and key secondary endpoints.

For analysis within each estimand, type 1 error rate control strategy for evaluation of primary and key secondary objectives is illustrated below:

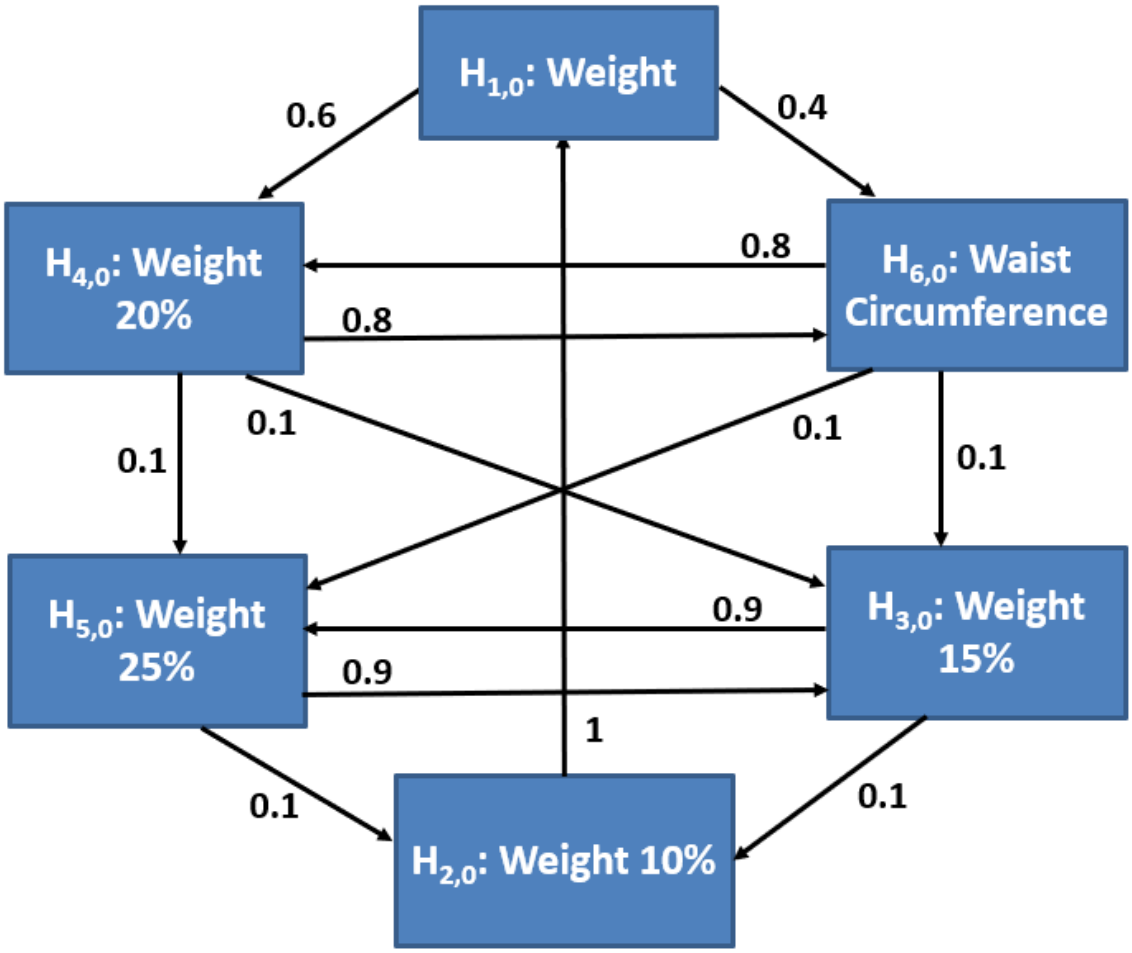


Figure 2.1. **Type 1 error control strategy for primary and key secondary efficacy endpoints.**

3. Analysis Sets

The tables define the analysis populations and datasets for the purposes of analysis based on the estimands defined in Section 1.1.

Population	Description
Screened population	All participants who signed informed consent.
Randomized population	All participants who are randomly assigned to a treatment arm.
Modified intent-to-treat population (mITT)	All randomly assigned participants who are exposed to at least 1 dose of study intervention. Participants will be analyzed according to the treatment they were randomly assigned to regardless of the treatment actually received.

Analysis Sets	Description
Efficacy Analysis Set (EAS): This analysis set will be used to estimate the efficacy estimand	Data obtained during the Treatment Period from the mITT, excluding data after permanent discontinuation of treatment or initiation of other anti-obesity medication, bariatric surgery, or other weight management procedures.
Full Analysis Set (FAS): This analysis set will be used to estimate the modified treatment-regimen estimand	Data obtained during the Treatment Period from the mITT, regardless of adherence to treatment and regardless of initiation of other anti-obesity medication except for non-study tirzepatide/semaglutide. Data after taking non-study tirzpepatide for participants randomized to semaglutide or after taking semaglutide for participants randomized to tirzepatide, having bariatric surgery, or other weight management procedures will be excluded.
Sensitivity Analysis Set: This analysis will be used to conduct sensitivity analyses for the modified treatment-regimen estimand	Data obtained during the Treatment Period from the mITT , regardless of adherence to treatment and regardless of initiation of other anti-obesity medication. Data after having bariatric surgery, or other weight management procedures will be excluded.
Efficacy Analyses Set for High Dose (EAS-HD)	Data obtained during the Treatment Period from the mITT, excluding data after permanent discontinuation of treatment, initiation of other anti-obesity medication, bariatric surgery, or other weight management procedures, and after not achieving high dose of study treatment by Week 40 at the latest.
Safety Analysis Set (SS): This analysis will be used to assess the safety of study treatment	Data obtained during the Treatment Period and any follow up data from the mITT, regardless of adherence to treatment and regardless of initiation of other anti-obesity medication.

4. Statistical Analyses

4.1. General Considerations

Statistical analysis of this study will be the responsibility of Lilly or its designee. The statistical analyses will be performed using SAS® Version 9.4 or higher.

Any change to the data analysis methods described in the protocol will require an amendment ONLY if it changes a principal feature of the protocol. Any other change to the data analysis methods described in the protocol, and the justification for making the change, will be described in the SAP or Clinical Study Report (CSR). Some analyses and summaries described in this analysis plan may not be conducted if not warranted by data (for example, few events to justify conducting an analysis). Listings of events will be provided in such situations. Additional exploratory analyses of the data will be conducted as deemed appropriate without further changes made to the protocol or SAP, even after final database lock.

Baseline is defined as the last non-missing measurement recorded on or before the randomization visit, prior to first dose of treatment, unless otherwise specified. For laboratory values for which time of collection is available, baseline needs to be prior the first dose time.

All tests of treatment effects will be conducted at a 2-sided alpha level of 0.05/1-sided alpha level of .025, unless otherwise stated, and all confidence intervals will be given at a 2-sided 95% level. In statistical summaries and analyses, all data will be analyzed by randomized treatment assignment. Participants will be analyzed according to the treatment they were randomly assigned to, regardless of the treatment actually received.

Efficacy analyses will use the efficacy analysis set (EAS) to evaluate the efficacy estimand and the full analysis set (FAS) to evaluate the modified treatment-regimen estimand. Safety will be assessed using Safety Analysis Set (SS).

End of study participation for a patient will be the earliest of date of death, date of withdrawal from further participation in the study, or date of last study visit. For patients considered to be lost to follow-up, end of study participation will be the date of lost to follow-up reported by the investigator. Patient data included in the database after the last date of study participation (date of death, date of ED or date of last study visit) will be excluded from statistical analysis. Listings of such data may be provided.

Summary statistics for continuous measures may include sample size, mean, standard deviation, median, minimum, and maximum. The analysis model to make comparisons between treatment groups relative to continuous measurements assessed at a specific time point will be an Analysis of Covariance (ANCOVA) and for over time will be Mixed Model Repeated Measures (MMRM).

The Kaplan-Meier method will be used for estimation of cumulative event-free survival rates over time.

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 when one or more covariates are included in the model. Otherwise, Fisher’s exact text will be used.

4.2. Participant Dispositions

Reasons for screen failures as reported by investigators will be summarized.

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

The number and percentage of participants prematurely discontinuing study treatment and study prior to the 72-week visit will be provided by study treatment. Reasons for prematurely discontinuing study treatment and study prior to the 72-week visit will be provided by study treatment.

4.3. Primary Endpoint Analysis

The primary endpoint for this study is percentage change from baseline in body weight to the 72-week visit (Visit 20). This endpoint will be used to evaluate the primary objective of the study for both the modified treatment-regimen and the efficacy estimands (Section 1.1).

The primary efficacy measure is mean percent body weight change from baseline. Body weight measurements will be collected at specific clinic visits as summarized in the Schedule of Activities. The null hypothesis corresponding to the primary objective is specified in Section 2.

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%*

4.3.1. Primary Analysis Relative to the Modified Treatment-Regimen Estimand

The primary efficacy analysis relative to the modified treatment-regimen estimand will be conducted using weight data in the Full Analysis Set (FAS) at baseline and at Week 72 with the aid of an ANCOVA model. The response variable will be the primary measure and model terms will include treatment, prediabetes status at randomization, sex, and BMI category, and baseline body weight as a covariate.

Missing body weight for participants who had bariatric surgery or another weight loss procedure during the study will be imputed using the worst value (for example, highest body weight value) from baseline until the last observed value before bariatric surgery or another weight loss procedure. Missing body weight values at the 72-week visit in patients who discontinued study drug early (including those participants who took non-study tirzepatide or semaglutide treatment switch) and did not have bariatric surgery or another weight loss procedure will be imputed based on observed data in the same treatment group from participants who had their efficacy assessed after early discontinuation of study drug. This analysis will be conducted with multiple imputations, and statistical inference over multiple imputations will be guided by the method proposed by Rubin (1987). Spurious missing data at Week 72 due to other reasons not related to early study treatment discontinuation or bariatric surgery or other surgical procedures will not be imputed.

With the aid of the ANCOVA analysis, p-values, and 2-sided 95% confidence intervals (CIs) for mean percent change in body weight from baseline to the 72-week visit will be derived and

summarized for the tirzepatide 15 mg or MTD arm compared to semaglutide 2.4 mg or MTD (1.7 mg or 2.4 mg).

4.3.2. Primary Analysis Relative to the Efficacy Estimand

The primary efficacy analysis relative to the efficacy estimand will be conducted using weight data over time in the efficacy analysis set (EAS). The primary analysis model for the mean percent change in body weight over time will be an MMRM. Restricted maximum likelihood (REML) will be used to obtain model parameter estimates and the Kenward-Roger option will be used to estimate denominator degrees of freedom. The response variable of MMRM will be the mean percent change in body weight from baseline values obtained at each scheduled post baseline visit.

The independent variables of the analysis model are treatment group (tirzepatide 15 mg or MTD, and semaglutide 2.4 mg or MTD), visit, and treatment-by-visit interaction, prediabetes status at randomization, sex, and BMI category as fixed effects, and baseline body weight as a covariate. Missing data will be addressed by the MMRM model. No explicit imputation methods for missing data will be employed.

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

1. Heterogenous Toeplitz
2. Heterogenous First Order Autoregressive
3. Heterogeneous Compound Symmetry
4. Toeplitz
5. First Order Autoregressive
6. Compound Symmetry

The first covariance structure that converges will be used. The resulting least squares mean (LSM) estimate of mean percent change from baseline in body weight will be summarized by visit and by study treatment.

With the aid of the MMRM analysis, p-values, and 2-sided 95% confidence intervals (CIs) for mean percent change in weight from baseline to the 72-week visit will be derived and summarized for the tirzepatide 15 mg or MTD arm compared to the semaglutide 2.4 mg or MTD arm.

4.3.3. Sensitivity Analyses

A sensitivity analysis for the primary endpoint relative to the modified treatment-regimen estimand will be conducted using the same methodology described in Section 4.3.1 using the Sensitivity Analysis Set.

In addition, a 2-way tipping point analysis will be conducted for the primary endpoint relative to the treatment-regimen estimand. This analysis investigates deviations from missing at random across both treatment arms by adding an arm-specific shift parameter to imputations of missing data points. More specifically, shift parameters δ_0 and δ_1 are added to imputed missing values for the comparator and active arms, respectively. In this study, where decreases in percentage

change from baseline in body weight are considered favorable, shift parameters of interest in a sensitivity analysis are those for which:

- δ_0 is less than or equal to 0, and
- δ_1 is greater than or equal to 0.

Missing data for this sensitivity analysis will be imputed based on the methodology described in Section 4.3.1 and parameters δ_0 and δ_1 will be added to imputed missing data. Likely or unlikely scenarios for shift parameters will be based on clinical judgment.

4.4. Secondary Endpoint Analysis

4.4.1. Key Secondary Endpoint

The endpoints corresponding to the secondary study objective subject to Type 1 error rate control are specified in Section 1.1 under "Key Secondary" (Controlled for Type 1 error) endpoints. The null hypotheses corresponding to the key secondary objectives can be found in Section 2.

Key secondary objectives will be evaluated based on the modified treatment-regimen and the efficacy estimands, similar to the primary objective.

4.4.1.1. Body Weight reduction from Baseline at the 72-week visit

Body weight reductions from baseline of ≥10%, ≥15%, ≥20%, and ≥25% will be analyzed at Week 72.

The missing data for body weight measurement at week 72 will be imputed as follows:

- For the modified treatment-regimen estimand: the same imputation methods described in section 4.3.1 will be applied.
- For the efficacy estimand: missing body weight in participants who discontinue study drug early or initiate other AOMs will be imputed by multiple imputation based on missing at random (MAR) assumption using the EAS.

Dichotomized data (Yes or No) will then be derived based on the continuous imputed values for each weight loss category.

A logistic regression model with terms of treatment group, BMI category, sex, and prediabetes status at randomization as fixed effects, and baseline body weight as a covariate will be used. The marginal mean for each treatment group will be calculated based on G-computation (FDA final guidance, 2023) and then the benefit difference (the difference in the proportions) and relative benefit (the ratio of the proportions) will the calculated. The standard errors for these quantities will be estimated using the delta-method (Ye et al., 2023). Two-sided 95% CI and p-values will be derived.

Approved on 23 Oct 2024 GMT

4.4.1.2. Mean change in Waist Circumference from Baseline at the 72-week visit

Assessment of superiority of mean change in waist circumference of tirzepatide 15 mg or MTD compared with semaglutide 2.4 mg or MTD at Week 72 will be conducted using the same statistical techniques as used for evaluating the primary objective in Section 4.3.1 (treatment regimen estimand) and Section 4.3.2 (efficacy estimand).

- For the treatment regimen estimand, the ANCOVA model (Section 4.3.1) will be updated by adding baseline waist circumference as a covariate instead of baseline body weight in the statistical model. Imputation of missing values at either Week 72 will be done in a similar manner as described in Section 4.3.1.
- For the efficacy estimand, the MMRM model (Section 4.3.2) will be updated by adding baseline waist circumference as a covariate instead of baseline body weight. Missing data will be addressed by the MMRM model and no explicit imputation methods will be used.

4.4.2. Additional Secondary Endpoint

Additional secondary endpoints specified below will use the efficacy estimand using EAS and will be summarized by treatment and nominal visit.

- Change from baseline in body weight (kg)
- Body weight reduction from baseline of ≥30%
- Change from baseline in BMI (kg/m2)

Percent change from baseline in body weight at 72 weeks will be used to compare tirzepatide at 15 mg to semaglutide at 2.4 mg. This objective will be assessed using an additional efficacy estimand, efficacy estimand for high dose, that includes not achieving the high dose of study treatment as an intercurrent event. This estimand answers the following question: What is the treatment difference in percent change in body weight from baseline after Week 72 of treatment as an adjunct to a reduced-calorie diet and increased physical activity in participants with obesity or overweight with at least 1 weight-related comorbid condition assuming that participants had stayed on treatment, had reached the high dose of study treatment and not taken other anti-obesity therapies (that is, other weight management drugs, bariatric surgery, or other weight management procedures)?

This efficacy for high dose estimand is described by the following attributes.

- Population: Participants who have obesity (BMI ≥30 kg/m2) or overweight (BMI ≥27 kg/m2) with at least 1 weight-related comorbid condition, but without Type 2 diabetes. Further details can be found in Section 5 of Protocol.
- Endpoint: Percent change from baseline to Week 72 in body weight.
- Treatment condition: The randomized treatment as an adjunct to a reduced-calorie diet and increased physical activity. Further details on study treatment can be found in Section 6 of Protocol.
- Intercurrent events: “Treatment discontinuation for any reason”, “Initiation of other anti-obesity therapies such as weight management drugs, bariatric surgery or other

weight management procedures" and 'Not achieving high dose of study drug' will be addressed using the hypothetical strategy:
    - Had participants stayed on treatment.
    - Had participants not taken other weight management drugs or bariatric surgery or weight loss procedure.
    - Had participant achieved the high dose of study drug by Week 40 or later and had not had any dose reduction, unless the dose reduction is per protocol with a BMI $\leq$ 22 kg/m^2.
- Population-level summary: Difference in mean percent changes between treatment conditions.

Analysis of percent change from baseline in body weight at 72 weeks for tirzepatide at 15 mg versus semaglutide at 2.4 mg will be conducted using the methodology described in Section 4.3.2.

Unless otherwise specified, additional secondary endpoints are not subject to type 1 error rate control. Some parameters may be log transformed, if necessary.

4.4.2.1. Analysis for Continuous Outcomes

The analysis to make comparisons between treatment groups for additional secondary continuous outcomes relative to the "efficacy" estimand will be conducted similar to the primary efficacy analysis in Section 4.3.2. The MMRM model will be updated by corresponding baseline values as a covariate. Missing data will be addressed by the MMRM model and no explicit imputation methods will be used.

4.4.2.2. Analysis for Binary Outcomes

The analysis for binary outcomes relative to the "efficacy" estimand will be performed using EAS following the methodology specified in section 4.4.1.1 for the efficacy estimand.

Objective	**Relative to the efficacy measure:**	**Analysis conducted in a manner similar to:**	**Additional Information**
Compare tirzepatide 15 mg or MTD (10 mg or 15 mg) to semaglutide 2.4 mg or MTD (1.7 mg or 2.4 mg) at Week 72			
Body Weight	Percentage of participants achieving ≥30% body weight reduction from baseline	logistic regression model in Section 4.4.1.1	
	Mean change in body weight (kg) from baseline	MMRM model in Section 4.3.2	
BMI	Mean change in BMI (kg/m2) from baseline	MMRM model in Section 4.3.2	Use baseline BMI as a covariate

Objective	Relative to the efficacy measure:	Analysis conducted in a manner similar to:	Additional Information
Compare tirzepatide 15 mg to semaglutide 2.4 mg at Week 72 for percent decrease in weight loss	Percent change from baseline in body weight	MMRM model in Section 4.3.2	Use baseline weight as a covariate. Use the EAS-HD

4.5. Tertiary Analysis

4.5.1. Tertiary Analysis Specified in the Protocol

Objective	Relative to the efficacy measure:	Analysis Conducted	Additional Information
Compare tirzepatide 15 mg or MTD (10 mg or 15 mg) to semaglutide 2.4 mg or MTD (1.7 mg or 2.4 mg) at Week 72			
Lipid parameters	Mean change in triglycerides(mg/dL),	MMRM model in Section 4.3.2	Use corresponding baseline lipid parameters as a covariate. Data will be log-transformed prior to fitting the MMRM model.
	Mean change in VLDL(mg/dL) from baseline	MMRM model in Section 4.3.2	Use corresponding baseline lipid parameters as a covariate. Data will be log-transformed prior to fitting the MMRM model.
	Mean change in non-HDL cholesterol (mg/dL) from baseline	MMRM model in Section 4.3.2	Use corresponding baseline lipid parameters as a covariate. Data will be log-transformed prior to fitting the MMRM model.

Objective	Relative to the efficacy measure:	Analysis Conducted	Additional Information
Insulin	Change from baseline in fasting insulin (pmol/L)	MMRM model in Section 4.3.2	Use corresponding baseline insulin values as a covariate. Data will be log-transformed prior to fitting the MMRM model.
HbA1c	Change from baseline in HbA1c (%)	MMRM model in Section 4.3.2	Use baseline HbA1c as a covariate.

4.5.2. Other Tertiary Analysis Not Specified in Protocol

Objective	Relative to the efficacy measure:	Analysis conducted in a manner similar to:	Additional Information
Compare tirzepatide 15 mg or MTD (10 mg or 15 mg) to semaglutide 2.4 mg or MTD (1.7 mg or 2.4 mg) at Week 72			
Additional lipid parameters	Mean change in LDL (mg/dL),	MMRM model in Section 4.3.2	Use corresponding baseline lipid parameters as a covariate. Data will be log-transformed prior to fitting the MMRM model.
	Mean change in HDL (mg/dL) from baseline	MMRM model in Section 4.3.2	Use corresponding baseline lipid parameters as a covariate. Data will be log-transformed prior to fitting the MMRM model.
Blood pressure parameters	Mean change in SBP (mm Hg) from baseline,	MMRM model in Section 4.3.2	Use SBP baseline as a covariate.
	Mean change DBP (mm Hg) from baseline	MMRM model in Section 4.3.2	Use DBP baseline as a covariate.

Objective	Relative to the efficacy measure:	Analysis conducted in a manner similar to:	Additional Information
SF-36	Change from baseline in SF-36 v2 norm-based scores: Physical component summary and Mental component summary, Physical Functioning domain score, and each additional individual domain: Role-Physical, Bodily Pain, General Health, Vitality, Social Functioning, Role-Emotional, Mental Health	MMRM model in section 4.3.2	Use baseline of each corresponding norm-based score. In addition to EAS, the FAS will be used as a sensitivity analysis.
PGIS	Percent of participants with improved categorical shift in Patient Global Impression of Status (PGIS) for Physical Activity from Baseline	Shift tables will be provided within each baseline category versus each postbaseline category	Counts and percentages of participants will be displayed.

4.6. Safety Analyses

Unless specified otherwise, safety assessments will be conducted based on the Safety Analysis Set (see Section 3) irrespective of adherence to study intervention or initiation of anti-obesity medication or bariatric surgery, unless indicated otherwise.

Adverse events will be coded from the actual term using the Medical Dictionary for Regulatory Activities and reported with preferred terms and system organ class. Selected notable AEs of interest may be reported using high-level terms or Standardized Medical Dictionary for Regulatory Activities Queries. Summary statistics will be provided for incidence of TEAEs, SAEs, and study discontinuation due to AEs, study intervention discontinuation due to AEs, or deaths from the time of first dose through the end of last study visit. Counts and proportions of participants experiencing AEs will be reported for each treatment group and Fisher's exact test will be used to compare the treatment groups.

4.6.1. Adverse Events

Adverse events will be coded from the actual term using the Medical Dictionary for Regulatory Activities and reported with preferred terms and system organ class.

A TEAE is defined as medical occurrence that emerges during a defined treatment period, having been absent pre-treatment, or worsens relative to the pre-treatment state, and does not necessarily have to have a causal relationship with this treatment. The maximum severity for each low-level term (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) 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 percentages of patients with TEAEs will be summarized by treatment using MedDRA preferred term (PT) nested within system organ class (SOC) and Fisher's exact test will be used to compare the treatment groups at both the SOC and PT levels. Events will be ordered by decreasing frequency within SOC. For events that are sex-specific, the denominator and computation of the percentage will include only patients from the given sex.

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

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

The percentages of patients with TEAEs by maximum severity will be summarized by treatment using MedDRA PT. For each patient and TEAE, the maximum severity for the MedDRA PT is the maximum postbaseline severity observed from all associated LLTs mapping to the MedDRA PT. The maximum severity will be determined based on the non-missing severities. If all severities are missing for the defined postbaseline period of interest, it will show as missing in the table. Only counts and percentages will be included for the TEAEs by maximum severity.

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

- Deaths
- SAEs
- Pregnancy
- Permanent Discontinuations of study treatment due to AEs

4.6.1.1. Death

A listing of all deaths will be provided. The listing will include patient identification including the treatment, site number, date of death, age at the time of enrollment, gender, MedDRA PT of associated AE, time from first dose of study drug to death, time from last dose of study drug to death (if patient had discontinued study drug), cause of death as reported by investigator, cause of death as adjudicated by Clinical Endpoint Committee (CEC).

4.6.1.2. Serious Adverse Events

The number and percentage of patients who experienced an SAE (including deaths and SAEs temporally associated or preceding deaths) will be summarized by treatment using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency within SOC. A listing of all SAEs will be provided. Listing will include but not limited to treatment, patient identification including the site number, treatment group, date of event, age at the time of enrollment, gender, MedDRA SOC and PT, severity, action taken, outcome, relationship to study drug, time from first dose of study drug to the event, and event duration.

4.6.1.3. Discontinuation from Study Due to Adverse Event

The number and percentage of patients who prematurely discontinue the study due to an AE will be summarized by treatment using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency within SOC. A listing of all the discontinuation from Study due to Adverse Event will be provided.

4.6.1.4. Discontinuation from Study Treatment Due to Adverse Event

The number and percentage of patients who prematurely discontinue study drug due to an AE will be summarized by treatment using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency within SOC. A listing of all the discontinuation from Study Treatment due to Adverse Event will be provided.

4.6.2. 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.

4.6.2.1. Hypoglycemia

The following categories of hypoglycemic events will be defined in the database:

Level 1 hypoglycemia

Glucose <70 mg/dL (3.9mmol/L) and ≥ 54 mg/dL (3.0 mmol/L): Level 1 hypoglycemia can alert a person to take action such as treatment with fast-acting carbohydrates. Providers should continue to counsel patients to treat hypoglycemia at this glucose alert value.

Level 2 hypoglycemia

Glucose <54 mg/dL (3.0 mmol/L): Level 2 hypoglycemia is also referred to as documented or blood glucose-confirmed hypoglycemia with glucose <54 mg/dL (3.0 mmol/L). This glucose threshold is clinically relevant regardless of the presence or absence of symptoms of hypoglycemia.

Level 3 hypoglycemia

Severe hypoglycemia: A severe event characterized by altered mental and/or physical status requiring assistance for treatment of hypoglycemia. For example, participants had altered mental status, and could not assist in their own care, or were semiconscious or unconscious, or experienced coma with or without seizures, and the assistance of another person was needed to actively administer carbohydrate, glucagon, or other resuscitative actions. Glucose measurements may not be available during such an event, but neurological recovery attributable to the restoration of glucose concentration to normal is considered sufficient evidence that the event was induced by a low glucose concentration.

Nocturnal hypoglycemia

Nocturnal hypoglycemia is a hypoglycemia event, including severe hypoglycemia, that occurs at night and presumably during sleep.

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

To avoid duplicate reporting, all consecutive blood glucose values occurring within a 1-hour period may be considered a single hypoglycemic event and the one with the lowest blood glucose value can be selected to be the representative. If two records with the same lowest values within an hour, the earlier occurrence will be selected.

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

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.

4.6.2.2. Pancreatitis

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 with 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 with treatment group, visit, and treatment-by-visit interaction, prediabetes status at randomization, sex, and BMI category as fixed effects, and baseline of the measurement as a covariate will be used to analyze each pancreatic enzyme with a log transformed response variable.

Pancreatitis Events:

Summaries of adjudicated and investigator-reported pancreatic events will be provided by treatment. Detailed searching criteria can be found in Section 6.4: Appendix 4.

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

4.6.2.3. Thyroid Malignancies and C-Cell Hyperplasia

Treatment-emergent thyroid malignancies and C-Cell hyperplasia will be considered AESIs. Detailed searching criteria can be found in Section 6.4: Appendix 4.

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.

4.6.2.4. Calcitonin Measurements

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 maximum 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.

4.6.2.5. Major Adverse Cardiovascular Events

Deaths due to cardiovascular events and nonfatal cardiovascular AEs will be adjudicated by a committee of physicians external to Lilly with cardiology expertise. This committee will be blinded to treatment assignment. The nonfatal cardiovascular AEs to be adjudicated include:

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

The counts and proportion of participants with positively adjudicated Major Adverse Cardiovascular Events may be summarized by treatment if deemed necessary.

In addition, MACE reported by investigator may also be summarized by treatment groups if number allows it, 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) will be provided. The dates of randomization, event, first dose and last dose of study intervention, and time from randomization to event will be listed.

4.6.2.6. Arrhythmias and Cardiac Conduction Disorders

Severe/serious treatment-emergent arrhythmias and cardiac conduction disorders will be considered AESIs.

The treatment-emergent arrhythmias and cardiac conduction disorders events will be identified using the MedDRA PTs. Detailed searching criteria can be found in Section 6.4: Appendix 4.

The counts and percentages of participants with treatment emergent ventricular 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 ventricular arrhythmias and cardiac conduction disorders may be provided if deemed necessary.

4.6.2.7. 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 are 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 Section 6.4: Appendix 4. 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.

The serious/severe 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.

4.6.2.8. Injection Site Reactions

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

Additionally, potential injection site reactions will be searched by pre-defined MedDRA HLTs of injection site reactions, administration site reactions, and infusion related reactions. Detailed searching criteria for injection site reaction events can be found in Section 6.4: Appendix 4. The PT will be used for summary by treatment within each HLT category.

Severe/Serious Injection Site Reactions:

The severe/serious 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 severe/serious treatment-emergent ISRs will be summarized by treatment. A listing of participants with treatment-emergent severe/serious ISRs may be provided, if deemed necessary.

4.6.2.9. Hepatobiliary Disorders

4.6.2.9.1. *Hepatic Events*

Severe/serious hepatic events will be considered as AESI and summarized separately. The counts and percentages of participants with treatment-emergent potentially drug-related hepatic events will be summarized by treatment using the MedDRA PTs. Detailed searching criteria can be found in Appendix 4.

4.6.2.9.2. *Acute Gallbladder Disease*

Events related to acute gallbladder disease will also be summarized by treatment groups by PT with decreasing frequency. Detailed searching criteria can be found in Appendix 4).

Severe/serious acute gallbladder disease will be considered as AESIs and summarized separately.

4.6.2.10. Liver Enzymes

Analyses for laboratory analyte measurements are described in Section 4.6.4. 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 will be summarized between treatment groups:

- A shift table of maximum to maximum alanine aminotransferase (ALT) measurement from baseline (≤1 × ULN, > 1 × ULN, missing) to postbaseline with the following categories: ≤1 × ULN, >1 to <3 × ULN, ≥3 to <5 × ULN, ≥5 to <10 × ULN, ≥10 × ULN, missing.
- A shift table of maximum to maximum aspartate transaminase (AST) measurement from baseline (≤1 × ULN, > 1 × ULN, missing) to postbaseline with the following categories: ≤1 × ULN, >1 to <3 × ULN, ≥3 to <5 × ULN, ≥5 to <10 × ULN, ≥10 × ULN, missing.
- Shift tables of maximum to maximum total bilirubin and direct bilirubin from baseline to postbaseline with the following categories: ≤1 × ULN, > 1 to <2 × ULN, ≥2 × ULN, missing.
- Shift tables of maximum to maximum serum alkaline phosphatase from baseline to postbaseline with the following categories: ≤1 × ULN, >1 to <2 × ULN, ≥2 × ULN, missing.

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

4.6.2.11. Gastrointestinal Adverse Events

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.

Severe gastrointestinal (GI) adverse events will be captured with the 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.

4.6.2.12. 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.

Severe/serious renal events from the following SMQ search will be considered 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.

An MMRM using REML model will be used to analyze each eGFR parameter at all scheduled postbaseline visits. The model will include prediabetes status at randomization (yes/no), sex (female/male), BMI group (<35 kg/m^2, ≥ 35 kg/m^2), treatment group, visit, and treatment-by-visit interaction as fixed effects, and baseline value of the dependent variable as a covariate.

Shift tables including unplanned measurements from minimum baseline to minimum postbaseline for all eGFR parameters will be provided. Shift tables will include the number and percentage of patients within each baseline category (<30, >=30 to <45, >=45 to <60, >=60 to <90, >=90, or missing) versus each postbaseline category ((<30, >=30 to <45, >=45 to <60, >=60 to <90, >=90, or missing) by treatment.

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, if deemed necessary.

4.6.2.13. Depression, Suicidal Ideation, or Behavior Monitoring

The severe/serious 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 Section 6.4: Appendix 4.

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).

4.6.2.13.1. 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 (that is, worsening of depression): includes participants in the none, mild, moderate, or moderately severe category during baseline and with at least 1 postbaseline measurement; and
- any 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
- any 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.

4.6.2.13.2. Suicidal Ideation and Behavior Solicited through C-SSRS

The counts and percentages of participants who experienced each of the C-SSRS categories as well as the following composite measures will be presented. The participants with at least 1 post-baseline C-SSRS assessment will be 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.

Approved on 23 Oct 2024 GMT

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

4.6.2.14. Events Related to Potential 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 Section 6.4: Appendix 4.

4.6.2.15. 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 Section 6: Appendix 4.

4.6.2.16. Overdose

A listing of participants reporting AEs related to overdosing of tirzepatide and semaglutide will be provided.

4.6.3. Vital Signs

Descriptive summaries by treatment and by nominal visit will be provided for baseline and postbaseline values as well as change from baseline values. If 2 or more records are taken at the same visit, they will be averaged prior to being used for data summaries and analyses.

An MMRM using REML model will be used to analyze the changes from baseline in pulse at all scheduled postbaseline visits. The model will include prediabetes status at randomization (yes/no), sex (female/male), BMI group (<35 kg/m^2, ≥ 35 kg/m^2), treatment group, visit, and treatment-by-visit interaction as fixed effects, and baseline value of the dependent variable as a covariate.

Counts and percentages of patients with treatment-emergent abnormal sitting systolic blood pressure (BP), sitting diastolic BP, and pulse will be presented by treatment. The criteria for identifying patients with treatment-emergent vital sign abnormalities are stated in Table GPHJ.4.1.

Table GPHJ.4.1. Categorical Criteria for Abnormal Treatment-Emergent Blood Pressure and Pulse Measurements

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

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

4.6.4. Clinical Laboratory Evaluation

All laboratory data will be reported in the International System of Units and Conventional Units. Values that are outside of reference ranges will be flagged as high (H) or low (L) in the listings. Descriptive summaries by treatment and by nominal visit will be provided for the baseline and postbaseline values as well as the change from baseline values for selected measurements.

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

Shift tables will be produced for selected measurements. A shift table will include unplanned measurements. The shift table will include the number and percentage of patients within each baseline category (low, normal, high, or missing) versus each postbaseline category (low, normal, high, or missing) by treatment. The proportion of patients shifted will be compared between treatments.

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

4.6.5. Product Complaints

A product complaint is any written, electronic, or oral communication that alleges deficiencies related to the identity, quality, durability, reliability, safety, effectiveness, or performance of a study intervention. When the ability to use the study intervention safely is impacted.

Product Complaints related to study interventions used in clinical studies are collected in order to ensure the safety of participants, monitor quality, and to facilitate process and product improvements.

Investigators will instruct participants to contact the site as soon as possible if he or she has a product complaint or problem with the study intervention so that the situation can be assessed.

An event may meet the definition of both a product complaint and an AE/SAE. In such cases, it should be reported as both a product complaint and as an AE/SAE.

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

4.7. Other Analyses

4.7.1. Health Outcomes

4.7.1.1. 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 Section 4.5.1.2.

4.7.1.2. Patient Global Impression of Status for Physical Activity

For PGIS Physical Activity, impaired physical function at baseline will be defined as a response of:

- moderately limited
- very much limited, or
- extremely limited

Descriptive summaries and analyses as described in Section 4.5.1.2 will be conducted.

4.7.2. Subgroup analyses

Subgroup analyses of the primary endpoint, percent change from baseline in weight at week 72 will be made based on the efficacy estimand using the EAS to assess consistency of the intervention effect. The MMRM model specified in Section 4.3.2 with the addition of the subgroup term (if not already included) and the treatment arm by subgroup interaction across the subgroups below:

- age group: < 65 years, ≥ 65 years

- sex: female, male
- baseline BMI: <35, ≥35 kg/m2
- race: white vs black vs other
- ethnicity: Hispanic vs not Hispanic, and
- prediabetes status at randomization (yes/no).

For each level of the subgroups above, the MMRM model specified in Section 4.3.2 will be used with the following changes:

- For the sex subgroups, sex will be removed from the model.
- For the baseline BMI subgroups, BMI subgroups will be removed from the model.
- For the prediabetes status at randomization subgroup, prediabetes status at randomization will be removed from the model.

4.8. Interim Analyses

No interim analyses are planned for this study. If an unplanned interim analysis is deemed necessary for reasons other than a safety concern, the protocol must be amended. Final analysis will be performed when all participants complete the study.

4.9. Changes to Protocol-Planned Analyses

The following changes to the protocol planned analyses have been modified in this SAP:

In section 4.3.2, the covariance structures, and the order in which they will be tested if the unstructured covariance structure does not converge was updated.

5. Sample Size Determination

Approximately 700 participants (350/group) will be randomly assigned in a 1:1 ratio to each study arm.

For the treatment-regimen estimand, this sample size provides approximately 90% power to show that tirzepatide 15 mg or MTD is superior to semaglutide 2.4 mg in the mean percent change from baseline at Week 72 using the following assumptions:

- a 2-sample t-test with a 2-sided significance level of 0.05
- a study drug discontinuation rate of 20% resulting in a common SD of 12%, and
- a 3% treatment difference.

For the efficacy estimand, this sample size also provides at least 90% power to show that tirzepatide 15 mg or MTD is superior to semaglutide 2.4 mg in the mean percent change from baseline at Week 72 using the following assumptions:

- a 2-sample t-test with a 2-sided significance level of 0.05
- a common SD of 10%
- a 3% treatment difference, and
- study drug discontinuation rate of 20%.

6. Supporting Documentation

6.1. Appendix 1: Demographic and Baseline 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 participants in the modified intent-to-treat (mITT) populations. Baseline demographic and clinical characteristics of special interest include but are not limited to: age (years), sex (female, male), race, ethnicity (if applicable), height (cm), weight (kg), BMI (kg/m2), waist circumference (cm), age group (<65 years, ≥65 years), BMI group (<35, ≥35), country, weight-related comorbidities. In addition, similar demographics and baseline characteristics might be provided by study treatment by prediabetes status(yes/no)at randomization.

6.2. Appendix 2: Treatment Compliance

Treatment compliance for each visit interval is defined as taking at least 75% of the required doses of study drug. Similarly, a participant will be considered significantly noncompliant if he or she is judged by the investigator to have intentionally or repeatedly taken more than the prescribed amount of medication (more than 125%). 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.

In addition to the assessment of a participant's compliance with the study drug administration, other aspects of compliance with the study treatments will be assessed at each visit based on the participant's adherence to the visit schedule, completion of study diaries, and any other parameters the investigator considers necessary.

6.3. Appendix 3: Clinical Trial Registry Analyses

Additional analyses will be performed for the purpose of fulfilling the Clinical Trial Registry (CTR) requirements.

Analyses provided for the CTR requirements include the following:

- Summary of adverse events, provided as a dataset which will be converted to an XML file. Both Serious Adverse Events and 'Other' Non-Serious Adverse Events are summarized: by treatment group, by MedDRA preferred term.
    - An adverse event is considered 'Serious' whether or not it is a treatment emergent adverse event (TEAE).
    - An adverse event is considered in the 'Other' category if it is both a TEAE and is not serious. For each Serious AE and 'Other' AE, for each term and treatment group, the following are provided:
        - the number of participants at risk of an event
        - the number of participants who experienced each event term
        - the number of events experienced.

Approved on 23 Oct 2024 GMT

- For each Serious AE, these additional terms are provided for EudraCT:
  - the total number of occurrences causally related to treatment
  - the total number of deaths
  - the total number of deaths causally related to treatment.
- Consistent with www.ClinicalTrials.gov requirements, 'Other' AEs that occur in fewer than 5% of patients/subjects in every treatment group may be excluded if a 5% threshold is chosen. Allowable thresholds include 0% (all events), 1%, 2%, 3%, 4% and 5%.
- AE reporting is consistent with other document disclosures for example, the CSR, manuscripts, and so forth.

Demographic table including the following age ranges required by EudraCT: adults (18-64 years), 65-85 years, and 85 years and over.

6.4. Appendix 4: Searching Criteria for Adverse Events of Special Interest

The adverse events of special interest (AESI) analyses are detailed in Section 4.6.2. The search criteria for each AESI are stored in CLUWE:

\\statsclstr\prd\diabetes\incretin\aesi\data\aesi_lookup.xlsx.

6.5. Appendix 5: Concomitant Therapy and Procedures of Interest

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

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

Concomitant medications of interest include the following: Use of the following medications after baseline:

- antidiarrheal medication
- antiemetic medication.

The following list of standardized generic names (CMDECOD) and corresponding standardized codes (TRTTERMCD) will be used to define the intercurrent event of initiation of other anti-obesity medication:

Standardized Generic Name	Code
LIRAGLUTIDE	05745801001
ORLISTAT	01215601001
SIBUTRAMINE	01356801001
PHENYLPROPANOLAMINE	00103801001
MAZINDOL	00309701001
PHENTERMINE	00131701001
LORCASERIN	07224601001
PHENDIMETRAZINE	00401901001
PHENTERMINE;TOPIRAMATE	12942601001
BUPROPION;NALTREXONE	12725401001
SEMAGLUTIDE	06507301001

Standardized Generic Name	Code
CELLULOSE MICROCRYSTALLINE;CITRIC ACID	15738901002
TIRZEPATIDE	15438301001

The following list of standardized codes (MHDECOD) will be used to define the intercurrent event of having bariatric surgery or other weight loss procedures:

Standardized Code
Bariatric gastric balloon insertion
Endoscopic sleeve gastroplasty
Metabolic surgery
Gastric banding

7. References

Alosh M, Bretz F, Huque M. Advanced multiplicity adjustment methods in clinical trials. *Stat Med*. 2014;33(4):693-713. https://doi.org/10.1002/sim.5974

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

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

[FDA] United States Food and Drug Administration. Adjusting for Covariates in Randomized Clinical Trials for Drugs and Biological Products. May 2023. Accessed Oct 11, 2024. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/adjusting-covariates-randomized-clinical-trials-drugs-and-biological-products

Ye T, Bannick M, Yi Y, Shao J. Robust variance estimation for covariate-adjusted unconditional treatment effect in randomized clinical trials with binary outcomes. *Stat Theory Relat Fields*. 2023;7(2):159-163. https://doi.org/10.1080/24754269.2023.2205802Rubin DB. Multiple imputation for nonresponse in surveys. New York: John Wiley & Sons Inc.; 1987.

Signature Page for VV-CLIN-131848 v2.0

Approval	PPD 23-Oct-2024 18:29:53 GMT+0000

Signature Page for VV-CLIN-131848 v2.0