Statistical Analysis Plan: I8F-MC-GPHN (Version 3)

Efficacy and Safety of Tirzepatide Once Weekly versus Placebo for Maintenance of Weight Loss in Participants without Type 2 Diabetes Who Have Obesity or Are Overweight with Weight-Related Comorbidities: A Randomized, Double-Blind, Placebo-Controlled Trial (SURMOUNT-4)

NCT04660643

Approval Date: 03-Feb-2023

Title Page

Protocol Title: Efficacy and Safety of Tirzepatide Once Weekly versus Placebo for Maintenance of Weight Loss in Participants without Type 2 Diabetes Who Have Obesity or Are Overweight with Weight-Related Comorbidities: A Randomized, Double-Blind, Placebo-Controlled Trial

Protocol Number: I8F-MC-GPHN

Compound Number: LY3298176

Short Title: Effect of Tirzepatide versus Placebo for Maintenance of Weight Loss (SURMOUNT-4)

Acronym: SURMOUNT-4

Sponsor Name: Eli Lilly and Company

Legal Registered Address: Indianapolis, Indiana USA 46285

Regulatory Agency Identifier Number(s): IND 139721

Confidential Information

Document ID: VV-CLIN-076401

List of Abbreviations

Term	Definition
ADA	anti-drug antibody
AE	adverse event
AESI	adverse event(s) of special interest
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
BG	blood glucose
BMI	body mass index
Bpm	beats per minute
CI	confidence interval
CN	conventional
CRF	case report form
CSR	clinical study report
C-SSRS	Columbia-Suicide Severity Rating Scale
CTR	Clinical Trial Registry
DBP	diastolic blood pressure
EAS	Efficacy Analysis Set
ECG	electrocardiogram
eCRF	electronic case report form
ED	early discontinuation
FAS	Full Analysis Set
GI	gastrointestinal
GIP	glucose-dependent insulinotropic polypeptide
GIPR	glucose-dependent insulinotropic polypeptide receptor
GLP-1	glucagon-like peptide-1
GLP-1R	glucagon-like peptide-1 receptor

Term	Definition
HDL	High-density lipoprotein
HLT	High Level Term
ICE	intercurrent event
ICH	International Council for Harmonisation
ISR	injection site reaction
IWQOL-Lite-CT	Impact of Weight on Quality of Life-Lite-Clinical Trials Version
IWRS	interactive web-response system
Lilly	Eli Lilly and Company
LLT	Lowest Level Term
LOCF	last observation carried forward
LSM	least squares mean
MACE	major adverse cardiovascular event(s)
MedDRA	Medical Dictionary for Regulatory Activities
mITT	modified intent-to-treat
MMRM	mixed model for repeated measures
MRD	minimum required dilution
MTD	maximum tolerated dose
NAb(+/-)	neutralizing antibody (positive/negative)
PG	plasma glucose
PGIS	Patient Global Impression of Status for Physical Activity
PHQ-9	Patient Health Questionnaire
PK	pharmacokinetic(s)
PT	Preferred Term
QTcF	Fridericia's corrected QT interval
QW	once weekly
REML	restricted maximum likelihood

Term	Definition
SAP	statistical analysis plan
SAE	serious adverse event
SBP	systolic blood pressure
SD	standard deviation
SI	Système International
SMQ	Standardized MedDRA Query
SOC	System Organ Class
SS	Safety Analysis Set
TE ADA (+/-)	treatment-emergent anti-drug antibody (positive/negative)
TEAE	treatment-emergent adverse event
T2DM	type 2 diabetes mellitus
UACR	urine albumin-to-creatinine ratio
ULN	upper limit of normal

Table of Contents

Table of Contents

Table **Page**

Table of Contents

Figure **Page**

Version history

This SAP for Study I8F-MC-GPHN (GPHN) is based on the protocol dated 17 November 2020.

SAP Version History Summary

SAP Version	Approval Date	Change	Rationale
1	26-Mar-2021	Not Applicable	Original version
2	12-Oct-2022	Details can be found in SAP Version History	Details can be found in SAP Version History
3	See date on Page 1	See below	See below

The key changes for GPHN SAP Version 3 and rationale are summarized below:

1. Clarified that the population-level summary for the "difference in response percentage between treatment conditions" is assessed by odds ratio per the US FDA feedback (see Table GPHN.1.2).
2. Clarified that placebo multiple imputation will be used for missing data imputation in Category 1 for treatment-regimen estimand if there are not enough retrieved dropouts to impute missing data in Category 2 (see Section 4.3.2.3).
3. Added sensitivity analyses to evaluate the robustness of the primary results with respect to different missing data imputation methods per the US FDA feedback (see Section 4.3.2.4).
4. Clarified the covariate in the analysis model (see Section 4.4):
    - For the analyses of endpoints that were compared to the time of "randomization (Week 36)"
        - Weight-related endpoints (ie, weight, BMI): stratification factors (except percentage weight loss at randomization [Week 36]) and corresponding value at randomization (Week 36) were included as covariates.
        - Other endpoints: all stratification factors and corresponding values at randomization (Week 36) were included as covariates.
    - For the analyses of endpoints that were compared to the time of "baseline (Week 0)"
        - Weight related endpoints (ie, weight, BMI): stratification factors (except percentage weight loss at randomization [Week 36]) and corresponding values at randomization (Week 36) and at baseline (Week 0) were included as covariates.
        - Other endpoints: all stratification factors and corresponding values at baseline (Week 0) were included as covariates.

1. Introduction

The main SAP focuses on describing the analysis plan for the double-blinded treatment period. Details of the analysis plan for the open-label lead-in period can be found in Appendix 4. Changes to the protocol-planned analyses are described in Section 4.9.

1.1. Objectives and Endpoints

Table GPHN.1.1. Objectives and Endpoints

Objectives	Endpoints
Primary	
To demonstrate that tirzepatide MTD is superior to placebo for percent change in body weight at Week 88	• Mean percent change in body weight from randomization (Week 36) to 88 weeks
Key Secondary (controlled for type I error)	
To demonstrate that tirzepatide MTD is superior to placebo in change from randomization (Week 36) for the following (measured at 88 weeks):	
• Body weight	• Mean change in body weight (kg)
• Waist circumference	• Mean change in waist circumference (cm)
• Maintaining body weight reduction achieved during the 36-week open-label period	• Percentage of participants who maintain ≥80% of the body weight lost during the 36 weeks of open-label period
To demonstrate that tirzepatide MTD is superior to placebo in change from Visit 2 (Week 0) for the following:	
• Body weight	• Percentage of study participants who achieve ≥5% body weight reduction at 88 weeks • Percentage of study participants who achieve ≥10% body weight reduction at 88 weeks • Percentage of study participants who achieve ≥15% body weight reduction at 88 weeks • Percentage of study participants who achieve ≥20% body weight reduction at 88 weeks
• Prevention of weight regain following the 36 weeks of open-label period	• Time (in weeks), during the 52-week double-blind treatment period, to first occurrence of participants returning to >95% baseline weight for those who have already lost ≥5% since Week 0

Objectives	Endpoints
To demonstrate that tirzepatide MTD is superior to placebo in change from randomization (Week 36) for percent change in body weight at Week 64	• Mean percent change in body weight from randomization (Week 36) to 64 weeks
Additional Secondary	
To demonstrate that tirzepatide MTD is superior to placebo in change from randomization (Week 36) for the following (measured at 88 weeks):	
• BMI	• Mean change in BMI (kg/m2)
• Glycemic control	• Mean change in: o Fasting glucose (mg/dL) o HbA1c (%)
• Insulin	• Mean change in fasting insulin (pmol/L)
• Lipid parameters	• Mean change in: o Total cholesterol (mg/dL) o LDL-cholesterol (mg/dL) o Non-HDL-cholesterol (mg/dL) o HDL-cholesterol (mg/dL) o VLDL-cholesterol (mg/dL) o Triglycerides (mg/dL) o Free Fatty acids (mg/dL)
• Blood pressure	• Mean change in: o systolic blood pressure (mmHg) o diastolic blood pressure (mmHg)
• Patient-Reported Outcomes	• Mean change in SF-36 v2 acute form in: o Physical Functioning domain (PF) o Role-Physical domain (RP) o Role-Emotional domain (RE) o Mental Health domain(MH) • Mean change in IWQOL-Lite-CT Physical Function composite score
To demonstrate that tirzepatide MTD is superior to placebo in change from Visit 2 (Week 0) for the following (measured at 88 weeks) for:	
• Body weight	• Mean change in body weight (kg) • Mean percent change in body weight (%) • Mean change in BMI (kg/m^2)
• Waist circumference	• Mean change in waist circumference (cm)
• Glycemic control	• Mean change in: o Fasting glucose (mg/dL)

Objectives	Endpoints
	○ HbA1c (%)
• Insulin	• Mean change in fasting insulin (pmol/L)
• Lipid parameters	• Mean change in: ○ Total cholesterol (mg/dL) ○ LDL-cholesterol (mg/dL) ○ Non-HDL-Cholesterol (mg/dL) ○ HDL-cholesterol (mg/dL) ○ VLDL-cholesterol (mg/dL) ○ Triglycerides (mg/dL) ○ Free fatty acids (mg/dL)
• Blood pressure	• Mean change in: ○ systolic blood pressure (mmHg) ○ diastolic blood pressure (mmHg)
• Patient-Reported Outcomes	• Mean change in SF-36 v2 acute form in Physical Functioning domain (PF). • Mean change in IWQOL-Lite-CT Physical Function composite score

Abbreviations: BMI = body mass index; HbA1c = hemoglobin A1c; HDL = high-density lipoprotein; IWQOL-Lite-CT = Impact of Weight on Quality of Life-Lite-Clinical Trials Version; LDL = low-density lipoprotein; MTD = maximum tolerated dose; SF-36 v2 acute form = Short Form-36 Version 2 Health Survey acute form; VLDL = very low-density lipoprotein.

1.2. Estimands

1.2.1. Primary Estimand

The primary clinical question of interest is: What is the impact of MTD of tirzepatide (10 mg or 15 mg) QW, compared with placebo, on the maintenance of weight loss after an initial 36-week open-label tirzepatide lead-in treatment period, in study participants who do not have T2DM, and have obesity (BMI ≥30 kg/m^2) or are overweight (BMI ≥27 kg/m^2) with at least 1 weight-related comorbid condition?

For primary disclosure of Study GPHN including submission to the US FDA

The estimand is described by the following attributes:

- Population: patients qualifying for study entry per protocol inclusion/exclusion criteria who reach and tolerate either tirzepatide 10-mg or 15-mg dose at Visit 11(Week 36).
- Endpoint: percent change in body weight from randomization (Week 36) to 88 weeks.
- Treatment condition of interest: tirzepatide MTD versus placebo, regardless of study drug adherence.
- Handling of intercurrent events (ICEs): The ICEs leading to treatment discontinuation for any reason are addressed by the treatment condition of interest attribute and handled by

treatment policy strategy as described in the International Council for Harmonisation (ICH) E9 (R1) Addendum (ICH 2019). Further details can be found in Section 4.3.2.3.

- Population-level summary: difference in mean between treatment conditions.

Rationale for estimand: the primary estimand was requested by US FDA; it aims at reflecting how participants are treated in clinical practice, irrespective to the compliance to planned course of treatment.

This *de facto estimand* is referred as the "treatment regimen" estimand in the latter sections of this document, which is equivalent to the "hybrid" estimand in the protocol.

For all other purposes

The estimand is described by the following attributes:

- Population: patients qualifying for study entry per protocol inclusion/exclusion criteria who reach and tolerate either tirzepatide 10-mg or 15-mg dose at Visit 11(Week 36).
- Endpoint: percent change in body weight from randomization (Week 36) to 88 weeks.
- Treatment condition of interest: tirzepatide MTD versus placebo, excluding data after discontinuation of study drug.
- Handling of ICEs: The ICEs leading to treatment discontinuation for any reason are addressed by the treatment condition of interest attribute and handled by the hypothetical strategy as described in the ICH E9 (R1) Addendum (ICH 2019).
- Population-level summary: difference in mean between treatment conditions.

Rationale for estimand: this estimand aims at reflecting treatment efficacy in an envisaged scenario in which the intercurrent events leading to treatment discontinuation would not occur.

We will refer to this estimand as "efficacy" estimand in the latter sections.

1.2.2. Estimands for Secondary Endpoints

Table GPHN.1.2. Attribute Description of Secondary Estimands

Secondary Estimand	Population	Endpoint	Treatment condition of interest	Handling of intercurrent events	Population-level summary
Key secondary related to mean change	Same as primary estimand	As described in Table GPHN.1.1	For primary disclosure of Study GPHN including submission to the US FDA, follow the same manner of the treatment regimen estimand in Section 1.2.1; for all other purposes, the same manner of the efficacy estimand in Section 1.2.1	For primary disclosure of Study GPHN including submission to the US FDA, follow the same manner of the treatment regimen estimand in Section 1.2.1; for all other purposes, follow the same manner of the efficacy estimand in Section 1.2.1	Difference in mean changes between treatment conditions
Key secondary related to percentage of participants meeting certain criteria					Difference in percentage of participants between treatment conditions as assessed by odds ratio
Key secondary related to time to meet certain criteria					Difference in time to response between treatment conditions
Additional secondary related to mean change	Same as primary estimand	As described in Table GPHN.1.1	Follow the same manner of the efficacy estimand in Section 1.2.1	Follow the same manner of the efficacy estimand in Section 1.2.1	Difference in mean changes between treatment conditions
Additional secondary related to percentage of participants meeting certain criteria					Difference in percentage of participants between treatment conditions as assessed by odds ratio

Abbreviation: FDA = Food and Drug Administration.

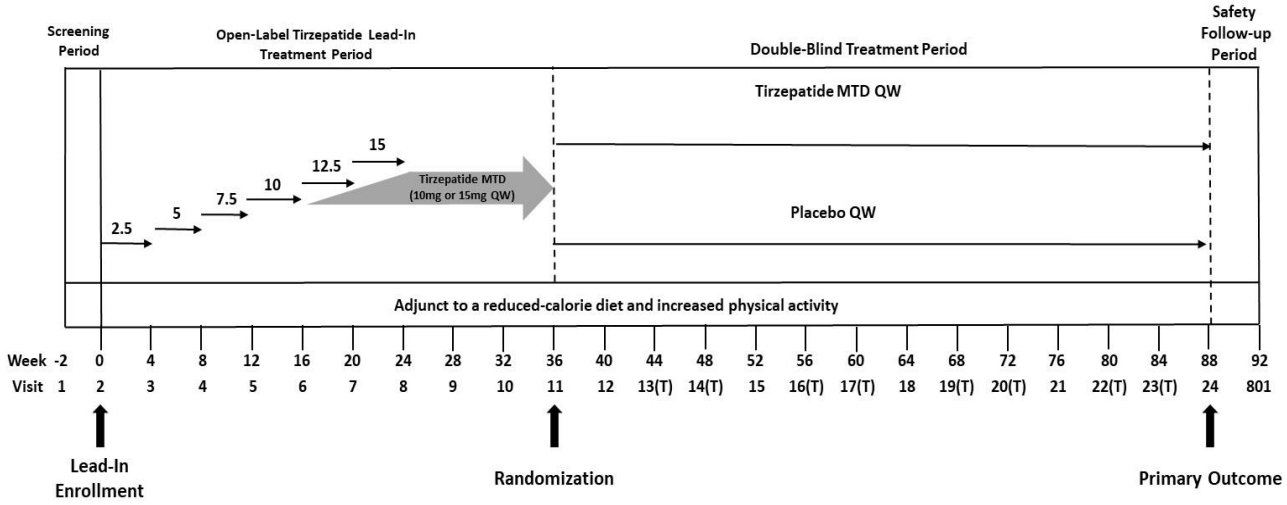
Unless otherwise specified, the attributes of estimands supporting exploratory objectives will follow the same manner of the efficacy estimand in Section 1.2.1, with the exception of endpoints and population-level summary, which will be objective specific.

1.3. Study Design

Study GPHN is a Phase 3, multicenter, randomized, parallel-arm, double-blind, placebo-controlled, 88-week study that will investigate the impact of MTD of tirzepatide (10 mg or 15 mg QW), compared with placebo, on the maintenance of weight loss after an initial 36-week open-label tirzepatide lead-in treatment period, in study participants who do not have T2DM, and have obesity (BMI ≥30 kg/m^2) or are overweight (BMI ≥27 kg/m^2) with at least 1 weight-related comorbid condition, excluding diabetes mellitus.

Study GPHN will consist of 4 periods: a 2-week screening period; a 36-week open-label tirzepatide lead-in treatment period (including a 20-week dose-escalation period); a 52-week double-blind, placebo-controlled treatment period; and a 4-week safety follow-up period. The study participants will be randomized in a double-blind 1:1 ratio (tirzepatide MTD QW or placebo) at the end of the open-label lead-in treatment period. Assignment to treatment groups will be determined by a computer-generated random sequence using the IWRS. An upper limit of 70% enrollment of women will be used to ensure a sufficiently large sample of men.

The randomization will be stratified by country, sex (female, male), tirzepatide MTD dose at 36 weeks (10 mg versus 15 mg), and percent weight loss at 36 weeks (<10% versus ≥10%). Where necessary to be included as a stratification factor, countries with fewer than 10 randomized participants will be pooled into 1 category (pooled country).



Abbreviations: MTD = maximum tolerated dose; QW = once weekly; (T) = telephone visit.

Figure GPHN.1.1. Illustration of study design for Clinical Protocol I8F-MC-GPHN.

The details about the overview of study periods and study visits can be found in Study GPHN protocol Section 4. Details of the unique visits not displayed in Figure GPHN.1.1 are provided below.

Visit 99

Visit 99 is only applicable to participants who discontinue the double-blind study treatment prematurely (after Week 36 and before Week 88) and decline to complete the remaining scheduled study visits. These participants will be asked to return for Visit 99 at 88 weeks ±7 days. This visit is critical to ensure complete data collection for the primary body weight endpoint.

Early Discontinuation of Treatment Visit

Participants unable or unwilling to continue the study treatment for any reason will perform an ED visit at the visit when the participant informs the site about the study treatment discontinuation. If the participant is discontinuing during an unscheduled visit or a scheduled visit, that visit should be performed as an ED visit.

2. Statistical Hypotheses

The primary objective is to demonstrate that tirzepatide MTD is superior to placebo for percent change from randomization in body weight at Week 88. Thus, the alternative hypothesis to be tested in relation to the primary estimand is as follows:

H1(wgt_pchg_wk36): QW tirzepatide MTD is superior to placebo for percent change from randomization (Week 36) in body weight at 88 weeks.

The alternative hypotheses for the key secondary objectives controlling for type 1 error rate are the following:

H2(wgt_chg_wk36): QW tirzepatide MTD is superior to placebo for change from randomization (Week 36) in body weight (kg) at 88 weeks.

H3(wc_chg_wk36): QW tirzepatide MTD is superior to placebo for change from randomization (Week 36) in waist circumference (cm) at 88 weeks.

H4(maintain80): QW tirzepatide MTD is superior to placebo for percentage of participants who maintain ≥80% of the body weight lost during the open-label lead-in period at 88 weeks.

H5(return95): QW tirzepatide MTD is superior to placebo for delaying time (in weeks), during the 52-week double-blind treatment period, to first occurrence of participants returning to >95% Visit 2 (Week 0) body weight among participants who lost ≥5% body weight during the open-label lead-in period.

H6(wgt5_wk0): QW tirzepatide MTD is superior to placebo for percentage of participants who achieve ≥5% body weight reduction from Visit 2 (Week 0) at 88 weeks.

H7(wgt10_wk0): QW tirzepatide MTD is superior to placebo for percentage of participants who achieve ≥10% body weight reduction from Visit 2 (Week 0) at 88 weeks.

H8(wgt_pchg_wk64): QW tirzepatide MTD is superior to placebo for percent change from randomization (Week 36) in body weight at 64 weeks.

H9(wgt15_wk0): QW tirzepatide MTD is superior to placebo for percentage of participants who achieve ≥15% body weight reduction from Visit 2 (Week 0) at 88 weeks.

H10(wgt20_wk0): QW tirzepatide MTD is superior to placebo for percentage of participants who achieve ≥20% body weight reduction from Visit 2 (Week 0) at 88 weeks.

2.1. Multiplicity Adjustment

The type 1 error rate control strategy for primary and key secondary efficacy objectives is documented in Figure GPHN.2.1. Both efficacy estimand and treatment regimen estimand will be used to assess those objectives. As they are intended for different purposes, no multiplicity adjustments will be made for conducting separate analyses relative to the efficacy and treatment regimen estimands. In addition, no multiplicity adjustments will be made for evaluating additional secondary and exploratory objectives and safety assessments.

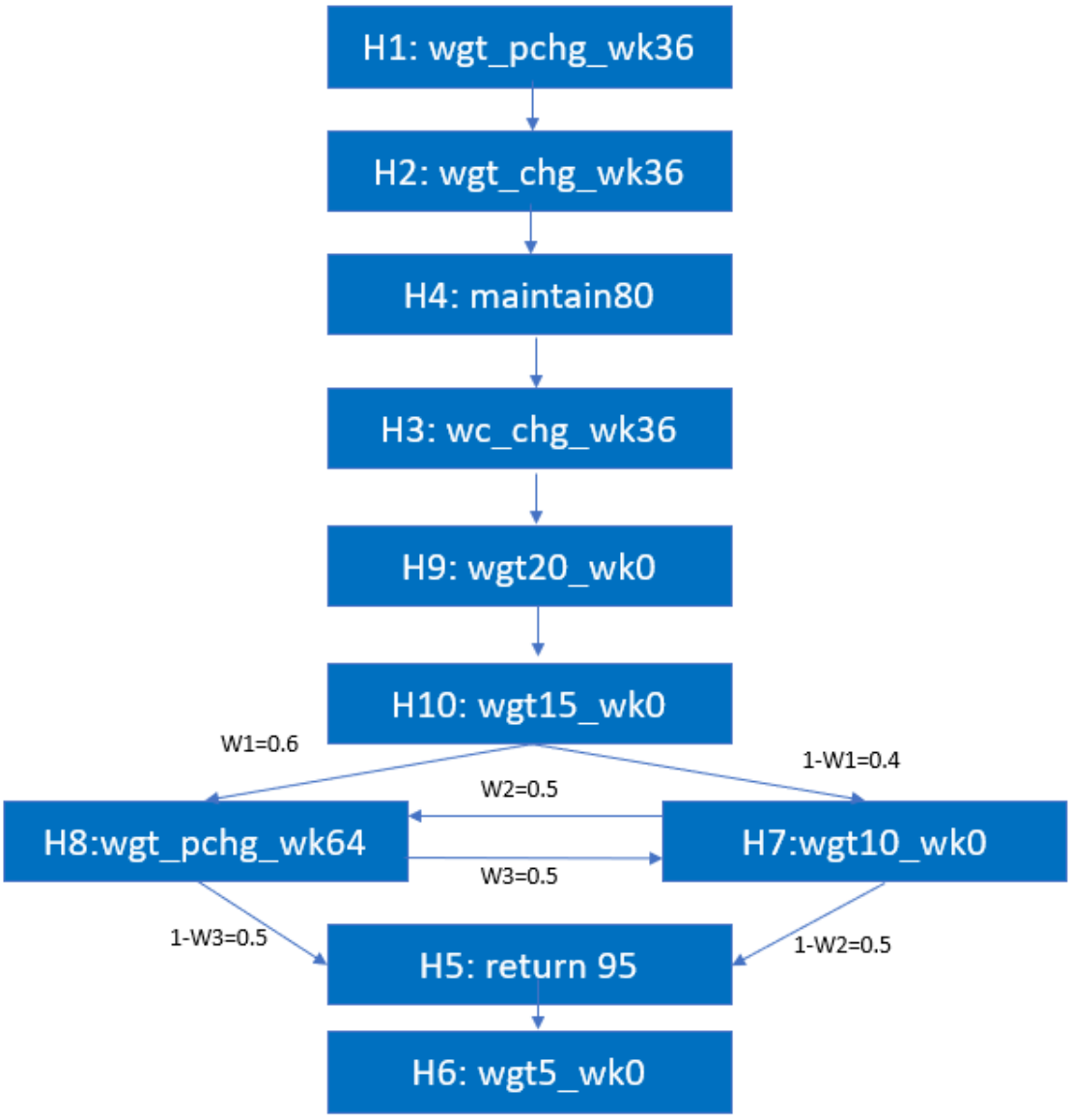


Figure GPHN.2.1. GPHN type 1 error control strategy for primary and key secondary efficacy endpoints.

3. Analysis Sets

For purposes of analyses, Table GPHN.3.1 defines the analysis population and analysis sets.

Table GPHN.3.1. Description of Analysis Population and Analysis Datasets

Population/Analysis Set	Description
Entered	All participants who sign informed consent.
Enrolled	All participants who are assigned to open-label tirzepatide treatment.
Randomized	All participants who are randomly assigned a study treatment (double-blind).
Modified Intent-to-Treat (mITT)	All enrolled participants who are exposed to at least 1 dose of study drug.
Efficacy Analysis Set (EAS)	Data obtained during double-blind treatment period from mITT, excluding data after discontinuation of study drug (last dose date + 7 days).
Full Analysis Set (FAS)	Data obtained during double-blind treatment period from mITT, regardless of adherence to study drug.
Safety Analysis Set (SS)	Data obtained during double-blind treatment period and safety follow-up period from mITT, regardless of adherence to study drug.

Unless otherwise specified, for analyses guided by the treatment regimen estimand, FAS will be used. For analyses guided by the efficacy estimand, EAS will be used, and for safety analysis, SS will be used.

For open-label period analysis, the analysis set is defined in Section 6.4.1.

For immunogenicity analysis, the analysis set is defined in Section 4.6.5.5.

4. Statistical Analyses

4.1. General Considerations

Statistical analysis of this study will be the responsibility of Lilly or its designee. Some analyses and summaries described in this analysis plan may not be conducted if not warranted by data (eg, few events to justify conducting an analysis). 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 CI will be calculated at 95% 2-sided.

Unless specified otherwise, efficacy and safety will be assessed using the mITT population, and data will be analyzed based on the assigned treatment (ie, not the actual treatment received by the participant). For submission to the US FDA to support the registration of tirzepatide for chronic weight management, the primary efficacy analysis will be conducted using FAS. For other purposes, the efficacy analysis of double-blind treatment period will be conducted using EAS. Safety analysis of double-blind treatment period will be conducted using SS.

Unless otherwise specified, for efficacy-related analyses at 88 weeks, country/pooled country, sex, tirzepatide MTD dose at randomization, percent weight loss at randomization and corresponding baseline value will be used as a covariate.

Summary descriptive statistics for continuous measures will include sample size, mean, SD, median, minimum, and maximum. The analysis model to make comparisons among treatment groups relative to continuous measurements assessed over time will be either ANCOVA or a 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, for analyses in the double-blind treatment period, baseline is defined as the last nonmissing data collected prior to or at randomization (Week 36); for analyses including data collected in open-label lead-in period, baseline is defined as the last nonmissing data collected prior to or at study enrollment (Week 0).

For some specific safety related parameters, the definition of baseline and postbaseline for double-blind treatment period are specified in Table GPHN.4.1.

The definition of baseline and postbaseline during the open-label lead-in treatment period can be found in Appendix 4.

Table GPHN.4.1. Baseline and Postbaseline Definitions for Safety Outcomes in Double-Blind treatment Period

Analysis Set	Analysis Type	Baseline	Postbaseline
SS	1.1) Treatment-Emergent Adverse Events	Ongoing at the time of the randomization to the double-blind treatment period.	Starts after randomization visit and ends after the whole study period (including off-drug follow-up visit).
SS	1.2) Treatment-Emergent Abnormal Labs[a], Vital Signs, and ECGs.	Refer to Table GPHN.6.1 for baseline definition.	Postbaseline will be defined as above (1.1). All scheduled and unscheduled measurements will be included.
SS	1.4) Change from Last 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 open-label lead-in period.	Postbaseline will be defined as above (1.1). Only scheduled visits will be included. The ED visits are considered scheduled visits.

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

[a] Immunogenicity-related analysis is specified in Section 4.6.5.5.

For the primary and key secondary efficacy endpoint analyses subject to type 1 error rate control, data for participants with missing values at the 88-week visit (or weight-related measurements at 64-week visit) will be imputed based on the method described in Section 4.3.2.3. Otherwise, missing values will not be explicitly imputed except for the parameters with only 1 postbaseline measure during the analysis period per schedule of activity, where LOCF approach will be applied to impute the endpoint when ED measure is available.

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 will be excluded from statistical analysis.

Statistical summaries and results of statistical analyses will be displayed in the following order: placebo, tirzepatide MTD.

Not all analyses described in this SAP will necessarily be included in the CSR. 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.

4.2. Participant Dispositions

A detailed description of participant disposition will be provided at the end of the study.

Frequency counts and percentages of all participants screened, enrolled, randomized, and receiving at least 1 dose of study drug will be summarized by treatment group, if applicable. Of the randomized population, frequency counts and percentages of participants who completed the study, prematurely discontinued the study (and/or study drug), including reason for premature discontinuation, will be summarized by treatment group.

The study completion status for the randomized population is defined as follows: participants with status reported via eCRF as "Completed" at follow-up visit (Visit 801) will be considered as completers; otherwise, they will be considered as non-completers. A Kaplan-Meier analysis of time from randomization to premature discontinuation from study and/or study treatment by treatment group will be provided.

A listing of study disposition for all randomized participants will be provided at the final database lock.

4.3. Primary Endpoint Analysis

For primary disclosure including submission to the US FDA to support the registration of tirzepatide for chronic weight management, the primary efficacy assessment will be guided by the treatment regimen estimand conducted using the FAS. Assessment of the primary objective will be conducted with hybrid imputation of missing data (see Section 4.3.2.3).

For other purposes, the assessment of primary efficacy objective will be guided by the efficacy estimand using the EAS.

4.3.1. Definition of Endpoint

The primary efficacy measure will be percent change in body weight from randomization. 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 (Week 36) to 88 weeks.

4.3.2. Main Analytical Approach

4.3.2.1. Analysis Related to the Treatment Regimen Estimand

The analysis related to the treatment regimen estimand will be conducted using 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 to Week 88. This model will include terms of treatment group, country/pooled country, sex, tirzepatide MTD dose at randomization, and body weight at randomization (Week 36) as covariates. The ANCOVA analysis will be conducted with multiple imputation of missing data guided by hybrid method at Week 88 (see Section 4.1 and Section 4.3.2.3 for details) and statistical inference over multiple imputation of missing data guided by Rubin (1987).

With the aid of the ANCOVA model, 2-sided 95% CI for mean percent change in body weight from baseline to the 88-week visit between tirzepatide MTD and placebo will be derived.

4.3.2.2. Analysis Related to the Efficacy Estimand

The analysis related to the 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 postbaseline visit.

For the MMRM model, the independent variables of analysis model are treatment group, visit, treatment-by-visit interaction, and stratification factors (country/pooled country, sex, and tirzepatide MTD dose at randomization) as fixed effects, and body weight at randomization (Week 36) 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 88-week visit for tirzepatide MTD compared to placebo will be derived and summarized. The resulting LSM estimates of mean percent change in body weight from baseline will be plotted by visit and by study treatment.

4.3.2.3. Hybrid Methods of Imputation

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

- **Category 1:** for missing data solely due to an exceptional circumstance, such as a pandemic or natural disaster (after other reasons for missing data are ruled out), considers the missing data as missing at random. The missing data will be imputed using all nonmissing data of the primary outcome measurement from the same treatment arm.
- **Category 2:** for missing data due to all other intercurrent events, it 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 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 (ie, placebo multiple imputation) will be used.

In cases where placebo multiple imputation method is used for missing data imputation in Category 2 because of not enough retrieved dropouts, the missing data in Category 1 will be imputed using all nonmissing data of the primary outcome measurement from the placebo group.

4.3.2.4. Sensitivity Analysis

For submission of Study GPHN to the US FDA tirzepatide application for chronic weight management, additional 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 88-week visit using an ANCOVA model. The terms of the model will be the same as specified in Section 4.3.2.1. These analyses are intended to assess the robustness of primary efficacy results using different missing data imputation methods:

- Retrieved dropout multiple imputation: Missing values of change in body weight at the 88-week visit will be imputed based on observed body weight change from randomization (Week 36) 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), the placebo multiple imputation method (described below) will be used.
- Placebo multiple imputation: Missing values of change in body weight at the 88-week visit will be imputed based on observed body weight change from randomization (Week 36) at the visit from participants in the placebo treatment group.
- 2-way tipping point analysis: The analysis will begin with the primary analysis aligned to treatment regimen estimand and adding positive and negative penalties simultaneously to both the tirzepatide MTD arm and the placebo arm, considering when results tip from superiority to inconclusive, and then considering the clinical plausibility of such scenarios.

4.4. Secondary Endpoint(s) Analysis

4.4.1. Main Analytical Approach for Categorical Outcome Related to Weight Measurement

For both treatment regimen and efficacy estimand, a logistic regression model will be used. For the proportion of patients achieving weight loss target since

- Baseline (Week 0), the following covariates will be included: treatment group, country/pooled country, sex, and tirzepatide MTD dose at randomization, and body weight at randomization (Week 36) and at baseline (Week 0).
- Randomization (Week 36), the following covariates will be included: treatment group, country/pooled country, sex, and tirzepatide MTD dose at randomization (Week 36), and body weight at randomization (Week 36).

For missing data imputation:

- For treatment regimen estimand: missing data will be handled by hybrid imputation of missing body weight at Week 88 (see Section 4.1 and Section 4.3.2.3 for details).

- For efficacy estimand: missing data will be handled by predicted value from MMRM analysis followed by dichotomization. When MMRM is used, variance-covariance structure for within-patient errors will be same as Section 4.3.2.2.

4.4.2. Key Secondary Efficacy Analyses Subject to Type 1 Error Rate Control

Table GPHN.4.2. Secondary Measures Controlled for Type 1 Error

Objectives	Relative to the efficacy measure:	Analysis conducted in a manner similar to	Additional Information
Tirzepatide MTD QW is superior to placebo:	Mean change in body weight (kg) from randomization (Week 36) to Week 88	For treatment regimen estimand: ANCOVA model described in Section 4.3.2.1. For efficacy estimand: MMRM model described in Section 4.3.2.2.	LSM estimates will be plotted by treatment through 88 weeks.
	Mean change in waist circumference (cm) from randomization (Week 36) to Week 88	For treatment regimen estimand: ANCOVA model with terms of treatment group, stratification factors, and corresponding value at randomization (Week 36) as covariates. For efficacy estimand: MMRM model with treatment group, visit, treatment-by-visit interaction, all stratification factors as fixed effects, and corresponding value at randomization (Week 36) as covariates.	LSM estimates will be plotted by treatment through 88 weeks. Refer to Sections 4.3.2.1 and 4.3.2.2 for analysis approach detail.
	Percentage of participants who maintain ≥80% of the body weight lost during the 36 weeks of open-label period	For both treatment regimen and efficacy estimand: Logistic regression model will be conducted.	Refer to Section 4.4.1 for analysis approach detail.
Tirzepatide MTD QW is superior to placebo:	Percentage of study participants who achieve ≥5/10/15/20% body weight reduction from Visit 2 (Week 0) to Week 88	For both treatment regimen and efficacy estimand: Logistic regression model will be used.	Refer to Section 4.4.1 for analysis approach detail.
	Time (in weeks), during the 52-week double-blind treatment period, to first occurrence of participants returning to >95% baseline weight for those who have already lost ≥5% since Week 0	For both treatment regimen and efficacy estimand: Cox-proportional hazards model will be used with the terms of treatment group (tirzepatide MTD and placebo), country/pooled country, sex, tirzepatide MTD dose at randomization, and weight at randomization (Week 36) and at baseline (Week 0) as covariates.	Only for participants with at least 5% body weight loss during the open-label lead-in period. For treatment regimen estimand: the participant without event will be censored at Week 88.

Objectives	Relative to the efficacy measure:	Analysis conducted in a manner similar to	Additional Information
			For efficacy estimand: the participant without event will be censored at treatment discontinuation time.
Tirzepatide MTD QW is superior to placebo:	Mean percent change in body weight from randomization (Week 36) to Week 64	For treatment regimen estimand: ANCOVA model will be conducted. For efficacy estimand: MMRM model will be conducted.	LSM estimates will be plotted by treatment through 64 weeks. Refer to Section 4.4.1 for analysis approach detail.

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

Decision will be guided by the 2-sided p-values in each objective subject to type 1 error rate control with details outlined in Section 2.1.

For submission of Study GPHN to the US FDA to support the registration of tirzepatide for chronic weight management, the key secondary efficacy assessment will be guided by the treatment regimen estimand conducted using the same population as for primary analysis. Assessment of key secondary objectives will be conducted with hybrid imputation of missing data (see Section 4.3.2.3).

For other purposes, the assessment of key secondary efficacy objectives will be guided by the efficacy estimand using the same population as for primary analysis.

4.4.3. Additional Secondary Endpoint(s)

Unless otherwise specified, other secondary efficacy analyses will be guided by the efficacy estimand using the same population as for primary analysis.

Analyses for labs including fasting glucose, hemoglobin A1c, fasting insulin, and lipid parameters will be performed for both SI units and CN units.

4.4.3.1. Additional Secondary Efficacy Analyses

Table GPHN.4.3. Secondary Efficacy Measures Not Controlled for Type 1 Error

Analysis variable	Analysis set	Analysis method	Additional Information
Mean change in measurements of lab tests and vital sign	Data from randomization (Week 36) to Week 88	MMRM model with treatment group, visit, treatment-by-visit interaction, all stratification factors as fixed effects, and the corresponding value at randomization (Week 36) as a covariate	Variance-covariance

Analysis variable	Analysis set	Analysis method	Additional Information
			structure for within-patient errors will be same as Section 4.3.2.2. LSM estimates may be plotted by treatment through 88 weeks. Log transformation will be adopted for fasting insulin and lipid parameters.
Mean change in SF-36 v2 acute form domain score in • Physical Functioning domain (PF) • Role-Physical domain (RP) • Role-Emotional domain (RE) • Mental Health domain (MH) Mean change in IWQOL-Lite-CT Physical Function composite score		ANCOVA model with terms of treatment group, all stratification factors, and the corresponding value at randomization (Week 36) as a covariate.	Missing data will be imputed using LOCF.
Mean change in measurements of lab tests and vital signs	Data from Week 0 to Week 88	MMRM model with treatment group, visit, treatment-by-visit interaction, all stratification factors, and the corresponding value at baseline (Week 0) as covariates.	Variance-covariance structure for within-patient errors will be same as Section 4.3.2.2. LSM estimates may be plotted by treatment through 88 weeks. Log transformation will be adopted for fasting insulin and lipid parameters.
Mean change in body weight, BMI, and percent change in body weight		MMRM model with treatment group, visit, treatment-by-visit interaction, country/pooled country and sex as fixed effects, and the corresponding value at baseline (Week 0) and at randomization (Week 36) as covariates.	Variance-covariance structure for within-patient errors will be same as Section 4.3.2.2. LSM estimates may be plotted by treatment through 88 weeks.

Analysis variable	Analysis set	Analysis method	Additional Information
Mean change in waist circumference		MMRM model with treatment group, visit, treatment-by-visit interaction, all stratification factors as fixed effects, and the corresponding value at baseline (Week 0) as covariates.	Variance-covariance structure for within-patient errors will be same as Section 4.3.2.2. LSM estimates may be plotted by treatment through 88 weeks.
Mean change in SF-36 v2 acute form in • Physical Functioning domain (PF) and IWQOL-Lite-CT in • Physical Function composite score		ANCOVA model with terms of treatment group, all stratification factors, and the corresponding value at baseline (Week 0) as covariates.	Missing data will be imputed using LOCF.

Abbreviations: ANCOVA = analysis of covariance; BMI = body mass index; IWQOL-Lite-CT = Impact of Weight on Quality of Life-Lite-Clinical Trials Version; LOCF = last observation carried forward; LSM = least squares mean; MMRM = mixed model for repeated measures; SF-36 v2 acute form = Short Form-36 Version 2 Health Survey acute form.

4.5. Exploratory Endpoint(s) Analysis

Unless otherwise specified, exploratory efficacy analyses will be guided by the efficacy estimand using the same population as for primary analysis.

Table GPHN.4.4. Exploratory Efficacy Analysis

Objective	Relative to the efficacy measure:	Analysis Conducted
Compare tirzepatide MTD QW with placebo from Visit 2 (Week 0) to Week 88	Percentage of participants achieving 25% body weight reduction from Week 0 to Week 88	Refer to Section 4.4.1 for detail.
Compare tirzepatide MTD QW with placebo from randomization to Week 88	Percentage of participants whose BMI shifts between clinically relevant categories, i.e., from randomization to Week 88 (<25, 25 to <30, 30 to < 35, 35 to < 40, ≥40)	Shift analysis will be conducted based on data from randomization to Week 88.
Compare tirzepatide MTD QW with placebo from randomization to Week 88	Percentage of participants whose lipid parameters shift between clinically relevant categories: • LDL-cholesterol (<70, 70 to <160, 160 to <190, ≥190mg/dL) • HDL-cholesterol (male: <40, ≥40mg/dL, female: <50, ≥50mg/dL) • Triglycerides (<150, 150 to <170, 170 to <500, ≥500mg/dL)	Shift analysis will be conducted based on data from randomization to Week 88.
Compare tirzepatide MTD QW with placebo from randomization to Week 88	Percentage of participants achieving ≥5% and ≥10% body weight reduction from randomization to Week 88	Refer to Section 4.4.1 for detail.

Objective	Relative to the efficacy measure:	Analysis Conducted
Visualize tirzepatide MTD QW and placebo's percent weight loss change from Week 0 up to safety follow-up after 88 weeks	Percent body weight loss measured from Week 0 to 88 weeks plus 4 weeks safety follow-up	Time-course plot will be generated. Only the participants who complete the 88 weeks treatment period and have the safety follow-up (Visit 801) will be included.
Compare tirzepatide MTD QW with placebo from randomization to Week 88	• SF-36v2 acute form Bodily Pain domain (BP) • SF-36v2 acute form General Health domain (GH) • SF-36v2 acute form Vitality domain • SF-36v2 acute form Social Functioning domain (SF) • SF-36v2 acute form Physical-Component Summary score • SF-36v2 acute form Mental-Component Summary score • 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	Data from randomization to Week 88 will be included. ANCOVA model with terms of treatment (tirzepatide MTD, placebo), all stratification factors, and corresponding score at randomization at Week 36 as a covariate. Missing data will be imputed using LOCF.

Abbreviations: ANCOVA = analysis of covariance; BMI = body mass index; HDL = high-density lipoprotein; IWQOL-Lite-CT = Impact of Weight on Quality of Life-Lite-Clinical Trials Version; LDL = low-density lipoprotein; MTD = maximum tolerated dose; QW = once weekly; SF-36 v2 acute form = Short Form-36 Version 2 Health Survey acute form; VAS = visual analog scale; VLDL = very low-density lipoprotein.

4.6. (Other) Safety Analyses

Unless specified otherwise, safety assessments for the double-blind treatment period will be based on the SS (Table GPHN.3.1). All data collected between randomization and the end date of study participation will be included, regardless of the adherence to study intervention.

Safety outcomes during the open-label lead-in treatment period will also be summarized for all enrolled participants. Detailed analyses can be found in Appendix 3.

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

The mean change from baseline differences among treatments at all scheduled visits will be assessed via an MMRM using REML. Unless specified otherwise, if the safety parameter is assessed at 88 weeks, then the model will include country/pooled country, sex, tirzepatide MTD dose at randomization, percent weight loss at randomization (<10% and ≥10%), 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 4.3.2.2 will be tested in order until convergence is met. If the data does not warrant the MMRM model, then ANCOVA model will be conducted.

For selected 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 weeks) from first dose or randomization (depends on the event of interest and analysis feature, eg, for time to study discontinuation, time from first dose might be used; for time to AE during double-blind treatment period, time from randomization might be used) 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 follows a negative binomial distribution and with treatment as a fixed effect. The logarithm of days during the double-blind treatment period will be adjusted as an offset to account for possible unequal treatment duration of follow-up between participants.

4.6.1. Extent of 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 in double-blind treatment period (defined as time in days from date of randomization to date of last dose of study treatment plus 7 days) will be provided for all randomized participants by treatment group using data from randomization to 88 weeks plus safety follow-up (Visit 801).

For the summary of duration on study and study treatment in double-blind treatment period, the frequency and percentage of participants in the following ranges will be summarized by planned treatment group as well:

- >0 weeks
- ≥12 weeks
- ≥24 weeks
- ≥36 weeks
- ≥52 weeks.

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

- 0 week
- ≥0 to <12 weeks
- ≥12 to <24 weeks
- ≥24 to <36 weeks
- ≥36 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.

4.6.2. Adverse Events

4.6.2.1. Treatment Emergent Adverse Events

A TEAE for the double-blind treatment period is defined as an event that first occurred or worsened in severity after the randomization to the double-blind treatment period. The definition of baseline and postbaseline for double-blind treatment period are specified in Table GPHN.4.1. The MedDRA LLT will be used in the treatment-emergent derivation. The maximum severity for each LLT during the baseline period including ongoing medical history will be used as baseline severity. For events with a missing severity during the baseline period, it will be treated as "mild" in severity for determining treatment-emergence. Events with a missing severity during the postbaseline period will be treated as "severe" and treatment-emergence will be determined by comparing to baseline severity.

The counts and percentages of randomized participants with TEAEs will be summarized by treatment using MedDRA PT nested within SOC. Statistical comparisons will be applied at both the SOC and PT levels. Events will be ordered by decreasing frequency within SOC. 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, SAE, death, discontinued from study treatment or study due to an AE, relationship to study drug will be summarized for all randomized participants by treatment.

The counts and percentages of randomized participants 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.

4.6.2.2. Common Adverse Events

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

4.6.3. Additional Safety Assessments

4.6.3.1. Deaths

Listing of all deaths will be provided by study period. The listing will include participant identification including the treatment, site number, date of death, age at the time of enrollment, sex, associated AE group identification, time from last dose of study drug to death (if participant had discontinued study drug), and primary cause of death.

Additional deaths that are reported outside of the study period will be obtained from the Lilly Safety System (LSS), which will be provided by Global Patient Safety.

4.6.3.2. Other Serious Adverse Events

The counts and percentages of randomized participants who experienced at least 1 SAE (including deaths and SAEs temporally associated with 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 (not limited to) treatment, participant identification including the site number, participant 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, and action taken related to study treatment.

4.6.3.3. Other Significant Adverse Events

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

4.6.3.4. Vital Signs

In the case that multiple measurements of an individual vital sign (eg, sitting SBP) are collected at the same visit, the mean of these measurements will be used for the analysis.

Vital signs (SBP, DBP, and pulse) will be summarized by treatment group at each planned visit. Change from baseline to postbaseline values for vital signs will be summarized for participants who have both a baseline and at least 1 postbaseline result. Treatment differences in mean change from baseline for vital signs will be assessed using analysis model described in Section 4.6. Only planned measurements will be included in the mean change analyses.

The counts and percentages of participants with treatment-emergent abnormal (high or low) vital signs (sitting SBP, DBP, and pulse) will be evaluated for patients randomized to the Tirzepatide treatment group and have both a baseline and at least 1 postbaseline result during the double-blind treatment period (including the off-drug follow-up time period). The baseline for treatment emergent abnormal vital signs is the same as the safety outcome baseline defined in open label period Table GPHN.6.1. 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. Both planned and unplanned measurements will be included in the analysis. The criteria for identifying participants with treatment-emergent vital signs abnormalities are listed in Table GPHN.4.5.

Table GPHN.4.5. Categorical Criteria for Treatment-Emergent Abnormal Blood Pressure and Pulse Measurement for Adults

Parameter	Low	High
Systolic BP (mmHg)	≤90 and decrease from baseline ≥20	≥140 and increase from baseline ≥20

(Supine or sitting – forearm at heart level)		≥129 and increase from baseline ≥20
Diastolic BP (mmHg) (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, for patients randomized to the Tirzepatide treatment group during the double-blind treatment period, the following analyses will be conducted, note that baseline for analyses listed below is the same as the safety outcome baseline defined in open label period Table GPHN.6.1:

- counts and percentages of participants who had resting pulse increases from baseline at 2 or more consecutive visits of more than 10 bpm and/or 20 bpm
- counts and percentages of participants who had resting pulse at any postbaseline visit increases from maximum baseline ≥20 bpm

In addition, the following analyses will be conducted by treatment group during the double-blind treatment period:

- counts and percentages of participants who had at least 1 resting pulse at any postbaseline visit exceeding 100 bpm
- counts and percentages of participants who had at least 1 resting pulse exceeding 100 bpm occurring at 2 or more consecutive study visits

4.6.3.5. Electrocardiograms

Summary statistics by treatment and by nominal visit will be provided for electrocardiogram (ECG) parameters (heart rate, PR interval, QRS, QT, and Fridericia's corrected QT interval (QTcF) [$QTcF = QT/RR^{0.333}$]). When the QRS is prolonged (eg, a complete bundle branch block), QT and corrected QT (QTc) should not be used to assess ventricular repolarization. Thus, for the analysis purposes, the QT and QTcF value will be set to missing when QRS is ≥120 msec for a particular ECG.

Change from baseline to postbaseline values for ECG parameters (heart rate and PR interval) will be summarized for participants who have both a baseline and at least 1 postbaseline result. Only scheduled measurements will be included in the mean change analyses.

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

- QT >500 msec
- QTcF >500 msec
- QTcF >480 msec

For patients randomized to Tirzepatide treatment group, the counts and percentages of participants who meet the following treatment-emergent criteria during the double-blind

treatment period (including the safety follow up time period) will be summarized. The baseline for treatment emergent abnormal ECG parameters is the same as the safety outcome baseline defined in open label period (Table GPHN.6.1):

- Treatment-emergent ECG abnormalities as listed in Table GPHN.4.6.
- Treatment-emergent increase from the maximum baseline in QTcF interval of >30 msec, 60 msec, or 75 msec. Maximum baseline (see Table GPHN.4.1 for details) will be the maximum non-missing observation in the baseline period. The maximum value during the double-blind treatment period will be analyzed. Scheduled and unscheduled measurements will be included.

Table GPHN.4.6. 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 = Fridericia's corrected QT interval.

4.6.3.6. Clinical Laboratory Evaluation

Descriptive summaries at each planned visit by treatment group will be provided for the baseline and postbaseline values and change from baseline values. The associated descriptive summaries will be presented in SI units and in CN units. Limits from the performing laboratory will be used to define low and high.

Observed and change from baseline to each visit will be displayed for participants who have both a baseline and at least 1 postbaseline planned measurement. Baseline is defined in Table GPHN.4.1.

A shift table will be provided including unscheduled 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 double-blind treatment period for the continuous measurements for selected lab tests.

4.6.4. Patient Narratives

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

- deaths
- SAEs
- permanent discontinuations of study treatment due to AEs, and
- pregnancy.

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

4.6.5. Special Safety Topics

For 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 patient-level data may be created using various formats such as a customized listing and/or a customized graphical patient profile. Adverse events of special interest are defined in each section of special safety topics, where applicable. Listing will be generated for each AESI.

4.6.5.1. Exocrine Pancreas Safety

4.6.5.1.1. Pancreatic Enzymes

Observed pancreatic enzyme data (p-amylase and lipase) will be summarized by treatment. The counts and percentages of participants with maximum postbaseline pancreatic enzyme value exceeding the following thresholds will be provided by baseline pancreatic enzyme value (≤ ULN, > ULN), and postbaseline: ≤1 × ULN, (>1 to ≤3) × ULN, (>3 to ≤5) × ULN, (>5 to ≤10) × ULN, >10 × ULN.

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

4.6.5.1.2. Pancreatitis Events

Summaries of adjudicated and investigator-reported pancreatic events will be provided by treatment. Detailed search criteria can be found in Appendix 6.

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

4.6.5.2. Gastrointestinal Adverse Events

4.6.5.2.1. 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 first onset of nausea, vomiting, and diarrhea will be plotted.

4.6.5.2.2. Severe Gastrointestinal Events

Severe GI AEs (GI SOC) will be captured with AE-CRF form and serious cases will be captured with the SAE form. The PTs in the GI SOC from the most recent MedDRA version at the time of database lock will be used to identify GI AEs, and only the PTs with severe/serious cases will be considered as AESIs.

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

4.6.5.3. Hepatobiliary Disorders

4.6.5.3.1. 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 7.

4.6.5.3.2. 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 7).

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

4.6.5.3.3. Liver Enzymes

Common analyses for laboratory analyte measurements described in Section 4.6.3.3 are applicable for the liver enzyme-related measurements. This section provides additional analyses for liver enzymes.

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

- The counts and percentages of participants with an ALT measurement ≥3 ×, 5 ×, and 10 × ULN for the central lab will be summarized for all participants with a postbaseline value and for subsets based on the following levels of baseline value:
    - participants whose nonmissing maximum baseline value is ≤1 × ULN
    - participants whose maximum baseline is >1 × ULN, and
    - participants whose baseline values are missing.
- The counts and percentages of participants with an aspartate aminotransferase) measurement ≥3 ×, 5 ×, and 10 × ULN for the central lab 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 measurement ≥2 × ULN for the central lab will be summarized for all participants with a postbaseline value, and for subsets based on the following levels of baseline values:
    - participants whose nonmissing maximum baseline value is ≤1 × ULN,
    - participants whose maximum baseline is >1 × ULN, but <2 × ULN,
    - participants whose maximum baseline value is ≥2 × ULN, and
    - participants whose baseline values are missing.
- The counts and percentages of participants with a serum alkaline phosphatase measurement ≥2 × ULN for the central lab will be summarized for all participants with a postbaseline value, and for subsets based on various levels of baseline, as described for total bilirubin.

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

4.6.5.4. Hypoglycemia

The following categories in accordance with the American Diabetes Association position statement on glycemic targets (ADA 2020) will be defined in the database.

Glucose Alert Value (Level 1):

- 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 PG 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 PG <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 PG <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 blood glucose (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 is 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 participants experiencing ≥1 episode) and the rate (episodes/patient/year) of Level 2 or Level 3 hypoglycemia, and Level 3 hypoglycemia will be reported.

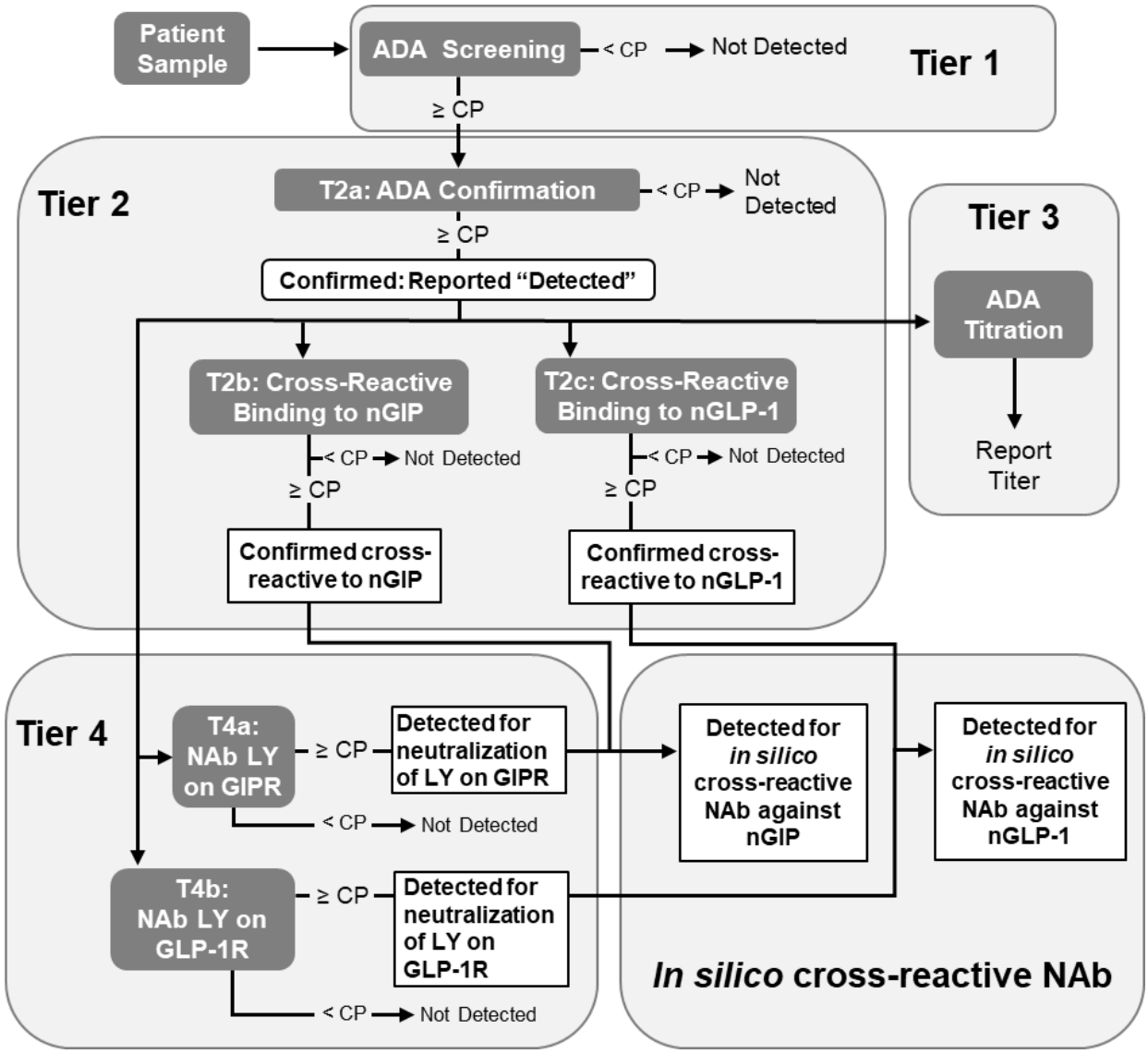
Severe/serious hypoglycemia will be considered as AESIs. The summaries of severe/serious hypoglycemia will be provided by treatment group. A listing of all events of severe/serious hypoglycemia may be provided, if deemed necessary. This listing will provide treatment, and clinical characteristics of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

4.6.5.5. Immunogenicity

4.6.5.5.1. 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 ADA assay result and potentially multiple cross-reactive antibodies assay results and multiple NAb 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 GPHN.4.1 details a flow chart that reflects the multitiered testing approach.



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

Figure GPHN.4.1. Flowchart of immunogenicity multitiered testing approach.

Table GPHN.4.7 outlines results as reported from Tier 2a of the multitiered testing approach. Tier 4 results are reported similarly.

Table GPHN.4.7. Sample 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 on other factors (see Table GPHN.4.8).
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 ADA 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 result is Not Detected (see Table GPHN.4.8).

Table GPHN.4.8. Sample Clinical 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 ≥ the assay's drug tolerance level, which may cause interference in the ADA detection method.

Abbreviation: ADA = anti-drug antibodies.

All ADA present samples will be evaluated for cross-reactivity to native GIP (Tier 2b), cross-reactivity to native GLP-1 (Tier 2c), NAb LY (tirzepatide) on GIPR (Tier 4a), and NAb LY (tirzepatide) on GLP-1R (Tier 4b). If cross-reactive ADA against native GIP is detected, the *in silico* assessment for cross-reactive NAb against native GIP is evaluated. If cross-reactive ADA against native GLP-1 is detected, the *in silico* assessment for cross-reactive NAb against native GLP-1 is evaluated (Figure GPHN.4.1).

Similar terminology to Table GPHN.4.8 applies for each type of cross-reactive and NAb assay. Importantly, each of these are distinct assays and, in general, have different assay operating characteristics.

The following are considered inconclusive for the NAb result:

- NAb LY on GIPR: if NAb result is not detected, and PK concentration is ≥ drug tolerance limit of the NAb LY on GIPR assay
- NAb LY on GLP-1R: if NAb result is not detected, and PK concentration is ≥ drug tolerance limit of the NAb LY on GLP-1R assay

To mitigate inconclusive cross-reactive neutralizing antibody interpretations against native GIP and GLP-1 due to potential tirzepatide concentrations above the drug tolerance limit of the NAb GIP on GIP-R (Tier 4c) and NAb GLP-1 on GLP-1R (Tier 4d) assays.

An *in silico* method utilizing results from Tiers 2b and 2c, Tiers 4a and 4b, and tirzepatide concentrations is used to determine cross-reactive NAb against native GIP and GLP-1. The *in silico* method is outlined in the following table:

Table GPHN.4.9. In Silico Classification for Cross-Reactive NAb

***In Silico* Classification**	**Cross-Reactive ADA Result**	**NAb Result**	**Circulating Tirzepatide Level (ng/mL)**	***In Silico* Cross-Reactive NAb Interpretation**
Cross-Reactive NAb to nGIP	Tier 2b: "Not Detected"	Tier 4a: "Not Detected" *or* Tier 4a: "Detected" or N/A or Missing	Any value or missing	Not Present
	Tier 2b: "Detected"	Tier 4a: "Not Detected"	< drug tolerance limit of Tier 4a assay	Not Present
	Tier 2b: "Detected"	Tier 4a: "Not Detected"	≥ drug tolerance limit of Tier 4a assay	Inconclusive
	Tier 2b: "Detected"	Tier 4a: "Detected"	< drug tolerance limit of Tier 4a assay	Present
	Tier 2b: "Detected"	Tier 4a: "Detected"	≥ drug tolerance limit of Tier 4a assay	Present
Cross-Reactive NAb to nGLP-1	Tier 2c: "Not Detected"	Tier 4b: "Not Detected" *or* Tier 4b: "Detected" or N/A or Missing	Any value or missing	Not Present
	Tier 2c: "Detected"	Tier 4b: "Not Detected"	< drug tolerance limit of Tier 4b assay	Not Detected
	Tier 2c: "Detected"	Tier 4b: "Not Detected"	≥ drug tolerance limit of Tier 4b assay	Inconclusive
	Tier 2c: "Detected"	Tier 4b: "Detected"	< drug tolerance limit of Tier 4b assay	Detected

In Silico Classification	Cross-Reactive ADA Result	NAb Result	Circulating Tirzepatide Level (ng/mL)	*In Silico* Cross-Reactive NAb Interpretation
	Tier 2c: "Detected"	Tier 4b: "Detected"	≥ drug tolerance limit of Tier 4b assay	Detected

Abbreviations: ADA = anti-drug antibodies; NAb = neutralizing antibody; nGIP = native glucose-dependent insulinotropic polypeptide; nGLP-1 = native glucagon-like peptide-1; Tier 2b = cross-reactive ADA to nGIP; Tier 4a = NAb LY (tirzepatide) on GIPR; Tier 4b = NAb LY (tirzepatide) on GLP-1R.

Note that in the case of an ADA Inconclusive sample, each of the NAb and Cross-Reactive NAb assay results is taken to be Inconclusive.

Note also that any reference to an assay cut point and/or drug tolerance is population specific, and is subject to modification to study-specific parameters per regulatory guidance.

4.6.5.5.2. Definitions of Immunogenicity Assessment Periods

Immunogenicity Baseline Observations: Baseline period for immunogenicity assessment for each participant includes all observations prior to first dose of study treatment. In instances where multiple baseline observations are collected, to determine participant ADA status the last nonmissing immunogenicity assessment prior to first administration of study drug 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 2 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

4.6.5.5.3. Definitions of Participant ADA Status

Participant evaluable for TE ADA: a participant has a nonmissing baseline ADA result, and at least 1 nonmissing postbaseline ADA result.

Treatment-emergent 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: A participant who is evaluable for TE ADA is TE ADA+ if either of the following holds:

- The participant has baseline status of ADA Not Present and at least 1 postbaseline status of ADA Present with titer ≥2× MRD of the ADA assay.
- The participant has 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 (B) status of ADA Present, with titer 1:B, and at least 1 postbaseline (P) status of ADA Present, with titer 1:P, with P/B ≥4.

As shown in Figure GPHN.4.1, 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 MRD (1:10), and a postbaseline sample with ADA detected and no titer is imputed to be 1 dilution above the MRD (1:20).

Treatment-emergent ADA inconclusive participant: A TE ADA-evaluable participant is TE ADA inconclusive if ≥20% of the participant's postbaseline samples are ADA inconclusive and all remaining postbaseline samples are ADA not present.

Treatment-emergent 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.

For each NAb assay, the following are defined:

NAb-positive (NAb+) participant: A participant who is TE ADA+ and has a NAb+ sample in the postbaseline period.

NAb-inconclusive participant: A participant who is TE ADA+, is not NAb+, and all samples that have TE ADA+ titer have a NAb-inconclusive sample result.

NAb-negative (NAb-) participant: A participant who is neither NAb+ nor NAb inconclusive.

Unless specified otherwise, the above-mentioned definitions of NAb are applicable to all NAb analyses, including cross-reactive NAb analyses, and cross-reactive antibodies.

4.6.5.5.4. Analyses to be Performed

Two analysis sets will be considered:

- mITT analysis set: refer to Section 3. Note that for mITT population, only time during tirzepatide exposure will be summarized.
- For participants who randomized to tirzepatide MTD arm.

The analysis period was defined in Section 4.6.5.5.2. The count and proportion of participants who are TE ADA+ will be tabulated, 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 and NAb. *In silico* classification for cross-reactive NAb will be summarized.

A summary will be provided of the count and percentage of tirzepatide-treated participants experiencing specific TEAE (see Table GPHN.4.10) by participant TE ADA status (TE ADA+, TE ADA-, TE ADA inconclusive) during the planned treatment period.

Table GPHN.4.10. Adverse Events for Analysis with Immunogenicity Results

TEAE category	Criteria
Hypersensitivity reactions	Anaphylactic reaction SMQ (narrow or algorithm)
	Hypersensitivity SMQ (narrow)
	Angioedema SMQ (narrow)
	Severe Cutaneous Adverse Reaction SMQ (narrow)
	Vasculitis 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 = Standardized MedDRA Query; TEAE = treatment-emergent adverse event.

A listing will be provided for all participants who had ADA Present at any time (including baseline) or had any specific TEAE (see Table GPHN.4.10). This listing includes a time course of ADA (clinical interpretation result plus flags for samples meeting TE ADA+ criteria and for samples with cross-reactive antibodies and NAb present) along with the TEAE.

Cases of TE ADA that are associated with AEs of either severe/serious hypersensitivity or severe/serious ISR will be classified as AESIs.

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

4.6.5.6. Hypersensitivity Reactions

Two main analyses are performed in support of assessment of potential immediate hypersensitivity, including anaphylaxis, and as well as potential non-immediate 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 eCRF, only date (no time) information is collected; if such events occurred on the same date as the study drug injection date, they 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 hypersensitivity TEAE will be summarized by PT with decreasing frequency by treatment.

Analyses for both time periods are based on the following:

- Narrow and algorithm terms in Anaphylactic reaction standardized MedDRA query (SMQ) (20000021) (Note that Anaphylactic reaction SMQ algorithm will only be summarized for Time Period A)
- Narrow terms in Angioedema SMQ (20000024)
- Narrow terms in Severe cutaneous adverse reactions SMQ (20000020)

- Narrow terms in Hypersensitivity SMQ (20000214), and
- Narrow terms in Vasculitis SMQ (20000174)

Additionally, for the Anaphylactic reaction SMQ in Time Period A, the algorithmic query (per the MedDRA Maintenance and Support Services Organization SMQ guide) will be performed. An algorithmic case must include either:

- A narrow term from the SMQ (Category A of the SMQ);
- Multiple terms from the SMQ, from the same administration of study drug, comprising terms from at least 2 of the following categories from the SMQ:
    - Category B - (Upper Airway/Respiratory)
    - Category C - (Angioedema/Urticaria/Pruritus/Flush)
    - Category D - (Cardiovascular/Hypotension)

Where algorithm is mentioned below, this applies only to 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 or algorithm term from any 1 of the 5 SMQs indicated above (ie, combined search across narrow and algorithmic portions of all 5 SMQs)
- Any narrow or algorithm term within each SMQ, separately (ie, narrow SMQ search).

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.

4.6.5.6.1. Severe/Serious Hypersensitivity Reactions

The severe/serious cases of hypersensitivity will be considered as AESIs. Summary with severe/serious hypersensitivity reactions may be provided, if deemed necessary.

4.6.5.7. Injection Site Reaction

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

Patient-based analysis and event-based analysis may be provided if necessary. The patient-based analysis summarizes all ISR questionnaire forms for an individual participant with a single statistic, typically an extreme value. This analysis allows each participant 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 participants. 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 participants 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 7.

The PT will be summarized within each category in decreasing order of incidence for tirzepatide-treated participants.

4.6.5.7.1. Severe/Serious Injection Site Reactions

The severe/serious treatment-emergent injection site reactions based on TEAE search criteria specified in Appendix 7 will be considered as AESIs. The counts and percentages of participants with severe/serious injection site reactions will be summarized by treatment. A listing of participants with treatment-emergent severe/serious ISRs may be provided, if deemed necessary.

4.6.5.8. Major Adverse Cardiovascular Events

The following positively adjudicated MACE will be considered as AESI:

- death due to cardiovascular AEs
- 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 percentages of participants with positively adjudicated MACE may be summarized by treatment.

A listing of participants reporting MACE events, either reported by investigator or identified by the Clinical Evaluation Committee, may be provided.

In addition, MACE reported by investigator may be summarized although a MACE reported by investigator and not positively adjudicated is not considered as AESI.

4.6.5.9. Major Depressive Disorder / Suicidal Ideation or Behavior

The severe/serious 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 7.

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 nested within SMQ. A listing of participants with major depressive disorder/suicidal ideation or behavior may be provided if deemed necessary.

In addition to spontaneously reported AEs assessed by the investigator, suicidal ideation and behavior and depression will be assessed through the use of the C-SSRS and the PHQ-9, respectively.

4.6.5.9.1. Patient Health Questionnaire

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

- None (not depressed) 0 – 4

- Mild 5 – 9
- Moderate 10 – 14
- Moderately Severe 15 – 19, and
- Severe 20 – 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.

4.6.5.9.2. Suicidal Ideation and Behavior Solicited Through C-SSRS

Suicide-related thoughts and behaviors occurring during the double-blind 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 postbaseline 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 the double-blind treatment period of the study to any 1 of the 5 suicidal ideation questions (Categories 1 – 5) on the C-SSRS
- **Suicidal behavior**: A "yes" answer at any time during double-blind treatment period of the study to any 1 of the 5 suicidal behavior questions (Categories 6 – 10) on the C-SSRS, and
- **Suicidal ideation or behavior**: A "yes" answer at any time during double-blind treatment period of the study to any 1 of the 10 suicidal ideation and behavior questions (Categories 1 – 10) on the C-SSRS.

A listing containing data for each participant with suicidal ideation, suicidal behavior, or non-suicidal self-injurious behavior during the double-blind treatment period of the study by treatment and visit may be provided. 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.

4.6.5.10. Malignancy

Malignancy will be considered as AESI. The counts and percentages of participants with treatment-emergent malignancy will be summarized by treatment and PT ordered by decreasing frequency. Detailed searching criteria can be found in Appendix 7.

4.6.5.11. Renal Safety

Laboratory measures related to renal safety will be analyzed as specified for laboratory measurements in Section 4.6.3.3.

Additionally, 2 shift tables examining renal function will be created. A min-to-min shift table of estimated glomerular filtration rate estimated by the Chronic Kidney Disease Epidemiology Collaboration 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 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 estimated glomerular filtration rate and UACR will be provided. Log transformation will be performed for UACR.

4.6.5.11.1. Acute Renal Events

Acute renal events, including those associated with chronic renal failure exacerbation, will also be captured.

Severe/serious renal events from the SMQ search below will be considered as AESI.

The counts and percentages of participants with 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)
- **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.

4.6.5.11.2. Dehydration

Dehydration events will be captured in the following SMQ and summarized. Severe/serious dehydration events will be considered as AESI:

- Narrow terms in Dehydration SMQ (20000232)

The counts and percentages of participants with dehydration will be summarized by treatment and PT and ordered by decreasing frequency. A listing of participants with treatment-emergent dehydration events may be provided, if deemed necessary.

4.6.5.12. Thyroid Safety Monitoring

4.6.5.12.1. Calcitonin

The number and proportion 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.

4.6.5.12.2. C-Cell Hyperplasia and Thyroid Malignancies

Thyroid malignancies and C-cell hyperplasia will be considered as AESI.

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.

4.6.5.13. Treatment-Emergent Arrhythmias and Cardiac Conduction Disorders

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

The treatment-emergent arrhythmias and cardiac conduction disorders events will be included using the MedDRA PTs. Detailed search criteria can be found in Appendix 7.

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.

4.6.5.14. Treatment of Overdose

Overdose is defined as taking more than 15 mg of tirzepatide in less than 72 hours. Overdosing of tirzepatide will be summarized by treatment group, and a listing of participants with tirzepatide overdosing may be provided.

In addition, a listing of participants reporting AEs related to overdosing of tirzepatide may be provided.

4.6.5.15. Abuse Potential

To identify AE terms suggestive of abuse liability potential, narrow terms from the SMQ Drug abuse and dependence (20000101) will be used. The counts and percentages of participants will be summarized by treatment group with decreasing frequency.

4.7. Other Analyses

4.7.1. 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 may be performed by Value Evidence Outcomes group at Lilly and documented in a separate analysis plan.

4.7.1.1. Patient Global Impression of Status for Physical Activity

The counts and percentages of participants endorsing each PGIS response category will be summarized by nominal visit and by treatment. A shift table from baseline to postbaseline of 5 PGIS response categories at each postbaseline visit by treatment will be created.

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

Per copyright owner, the QualityMetric Health Outcomes™ Scoring Software 4.5 will be used to derive the following domain and component scores:

- Mental Component Summary (MCS)
- Physical Component Summary (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).

For each above domain and component summary scores parameter, the raw scores will be transformed into the domain scores (t-scores) and the following analyses for the actual value and change from baseline value will be conducted:

- descriptive summaries by treatment group, and
- analysis described in Table GPHN.4.3 and Table GPHN.4.4.

The empirical cumulative distribution function curves of the change from randomization to Week 88 in SF-36 physical function domain score will be provided by treatment group.

If data allowed, analysis for SF-36 physical function domain score analysis described in Table GPHN.4.3 will be conducted to evaluate the treatment effect in participants who have limitations in physical function at baseline, which is defined as PGIS response at baseline of "moderately limited," "very much limited," or "extremely limited."

4.7.1.3. 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 nonmissing value; the total score will be calculated if a minimum of 75% of all 20 items has a nonmissing value.
- If the minimum required number of items is answered for a composite domain 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 patient 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 (i.e., 5)
  - S_{min} is the minimum possible raw score value (i.e., 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 following analyses for the actual value and change from baseline value will be conducted:

- descriptive summaries by treatment group, and
- ANCOVA analysis described in Table GPHN.4.3 and Table GPHN.4.4.

If data allows, analysis for IWQOL-Lite-CT physical function composite score analysis described in Table GPHN.4.3 and Table GPHN.4.4 will be conducted to evaluate the treatment effect in participants who have limitations in physical function at baseline (as defined in Section 4.7.1.3).

4.7.1.4. EQ-5D-5L

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

- descriptive summaries by treatment group, and
- analyses described in Table GPHN.4.4.

4.7.2. Subgroup Analyses

Efficacy subgroup analyses will be guided by the treatment-regimen estimand in FAS and the efficacy estimand in EAS.

Subgroup analyses may be done by country to support local regulatory registrations.

4.7.2.1. Subgroup Analysis of Body Weight Change

Subgroup analyses by the following baseline characteristics will be provided:

- age (<65 years and ≥65 years)
- sex (female and male)
- race

- ethnicity (Hispanic/Latino, non-Hispanic/non-Latino)
- BMI at study entry (<30, ≥30 and <35, ≥35 and <40, ≥40 kg/m^2)
- BMI at randomization (<25, ≥25 and <30, ≥30 and <35, ≥35 and <40, ≥40 kg/m^2)
- Percent body weight loss at 36 weeks (<10% and ≥10%), and
- Region of enrollment (US, outside of US).

The outcome measure for the subgroup analyses will include percent change in body weight from randomization at 88 weeks.

For the percentage change in body weight from randomization at 88 weeks, for each subgroup analysis aforementioned, the following analyses will be conducted:

For treatment regimen estimand: ANCOVA model will be used; for efficacy estimand: MMRM model will be used. The detail of analytical approach is documented in Section 4.4.1. For each estimand, 2 models will be considered:

- model on each stratum of the subgroup, and
- full model with use subgroup as a covariate.

4.8. Interim Analyses

No interim analyses are planned for this study.

A database lock will occur after all enrolled participants complete or discontinue the 36-week open-label lead-in treatment period. This is not considered an interim lock since the enrolled participants who complete the 36-week open-label lead-in treatment period will be randomized to a double-blind, placebo-controlled treatment period, and there is no plan to modify the study based on the results from this open-label lead-in database lock.

4.8.1. Data Monitoring Committee

Not Applicable.

4.9. Changes to Protocol-Planned Analyses

To provide data on efficacy of the investigational product that would be valuable to better inform clinical decisions in management of people living with obesity, protocol-planned objectives are changed as below.

Key secondary objectives:

- Moving percentage of study participants who achieve >20% body weight reduction at Week 88 from Visit 2 (Week 0) from exploratory endpoint to a key secondary endpoint.
- Moving percentage of study participants who achieve >15% body weight reduction at Week 88 from Visit 2 (Week 0) from additional secondary endpoint to a key secondary endpoint.

5. Sample Size Determination

Approximately 1000 participants will be screened and 750 participants enrolled into the 36-week open-label tirzepatide lead-in treatment period for approximately 600 participants to be randomized in a 1:1 ratio to tirzepatide MTD (300 participants) or placebo (300 participants).

The sample size determination assumes that evaluation of superiority of tirzepatide MTD to placebo will be conducted at a 2-sided significance level of 0.05 using a 2-sample t-test. Additionally, a difference of at least 6% mean body weight percentage change from randomization (36 weeks) to 88 weeks between tirzepatide MTD and placebo, a common SD of 8%, and a dropout rate of 25% are assumed for statistical power calculations. Under the assumptions above, randomizing 600 participants in a 1:1 ratio to tirzepatide MTD (300 participants) and placebo (300 participants) provides more than 90% power to demonstrate superiority of tirzepatide MTD to placebo.

6. Supporting Documentation

6.1. Appendix 1: Demographic and Other Characteristics of Study Population

6.1.1. Patient Characteristics

Summary tables will be generated for participant demographics and baseline disease characteristics for randomized population. Variables to be included (but not limited to) are: age, sex, race, ethnicity, weight, BMI, waist circumference, age groups, and BMI groups.

6.1.2. Historical Illnesses and Preexisting Conditions

The number and percentage of participants with historical illnesses and preexisting conditions will be summarized by treatment group using MedDRA PTs nested within SOC. System Organ Class will be in alphabetical order. Conditions (ie, PTs) will be ordered by decreasing frequency within SOC. A preexisting condition is defined as the condition/event recorded on the *Preexisting Conditions and Medical History* eCRF page or on the *Prespecified Medical History* eCRF page with a start date prior to the date of informed consent, and no end date (ie, the event is ongoing) or an end date on or after the date of informed consent. A historic illness is defined as an event with start date and stop date that are prior to the informed consent. The AEs occurring prior to first dose of study treatment will be included in the preexisting condition for reporting purpose.

6.1.3. Concomitant Medications

Summaries of Preferred Names of concomitant medications with number and percentage of patients sorted by decreasing frequency will be generated by treatment group.

Additionally, concomitant medications of interest (as defined below) will be summarized.

Concomitant medications of interest include the following:

- baseline antihypertensive therapy, by type/class
- baseline lipid lowering therapy, by type/class
- changes to baseline medication in post-randomization:
    - antihypertensive therapy, and
    - lipid lowering therapy
- utilization after randomization:
    - medicines that cause weight gain
    - antihyperglycemic medication for the treatment of diabetes for participants who develop T2DM during the study (antihyperglycemic medication for the treatment of prediabetes is not allowed pre protocol)
    - antidiarrheal medication, and
    - antiemetic medication.

The analysis of concomitant medication for the open-label period is provided in Appendix 4.

6.2. Appendix 2: Treatment Compliance

Treatment compliance will be defined as taking at least 75% of the required doses of study drug. Compliance at each treatment visit and over the corresponding 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 at the specific visit or over the corresponding study period, respectively. Treatment compliance will be summarized descriptively at each treatment visit and over the double-blind treatment period by treatment using the mITT population.

For the open-label period, the analysis of treatment compliance is provided in Appendix 4.

6.3. Appendix 3: Important Protocol Deviations

Important protocol deviations are identified in the Trial Issues Management Plan. A listing of all important protocol deviations will be provided at the end of the study. For the double-blind treatment period, a summary of important protocol deviations will be provided for all randomized participants by treatment.

6.4. Appendix 4: Analyses for the Open-Label Lead-in Period

6.4.1. General consideration

Unless specified otherwise, all analyses for open-label lead-in period will be based on all enrolled patients who are exposed to at least one dose of study drug.

Table GPHN.6.1. Baseline and Postbaseline Definitions in Open-Label Period

Population	Analysis Type	Baseline	Postbaseline
All enrolled participants	1.1) Treatment-Emergent Adverse Events	The baseline period is defined as the start of screening and ends prior to the first dose of tirzepatide treatment (typically at Week 0).	Starts after the first dose of tirzepatide treatment and ends prior to the randomization (typically at Week 36).
All enrolled participants	1.2) Treatment-Emergent Abnormal Labs, Vital Signs, and ECGs.	Baseline will include all scheduled and unscheduled measurements during the baseline period as defined above (1.1).	Postbaseline is defined as above (1.1). All scheduled and unscheduled measurements will be included.
All enrolled participants	1.3) Change from Last Baseline to Week *xx* and to Last Postbaseline for Labs, Vital Signs, and ECGs.	The last scheduled and unscheduled non-missing assessment recorded during the baseline period defined above (1.1).	Postbaseline is defined as above (1.1). Only scheduled visits will be included. The ED visits are considered scheduled visits.

Population	Analysis Type	Baseline	Postbaseline
All enrolled participants	Other outcomes (efficacy, PROs, questionnaires, etc.)	Baseline is defined as the last nonmissing data collected prior to or at study entry (Week 0)	Postbaseline is defined as data collected after the study entry (Week 0) and prior to or at the randomization (Week 36)

Abbreviations: ECG = electrocardiogram; ED = early discontinuation; PRO = patient-reported outcome.

For the open-label lead-in period, there will be no treatment comparison in summaries and analyses since all participants will be treated with tirzepatide.

6.4.2. Patient Characteristics

Summary tables will be generated for participant demographics and baseline disease characteristics for enrolled population. Variables to be included (but not limited to) are: age, sex, race, ethnicity, weight, BMI, waist circumference, age groups, and BMI groups.

6.4.3. Disposition

Summaries of study disposition and study drug disposition will be provided for the open-label lead-in period. Counts and percentages of all participants that are entered, enrolled, and discontinued from study or study drug will be summarized. A listing may also be provided if necessary.

The study completion status for the open-label lead-in period is defined as

- at the end of 36-week open-label lead-in period (when all enrolled participants complete or discontinue the open-label lead-in period). Participants with a nonmissing body weight measurement at 36 weeks (Visit 11) will be considered as completers; otherwise, will be considered as non-completers.

Summary of prematurely discontinued study treatment (including discontinuation reason) will be provided for all enrolled participants. A time-to-event analysis of premature study treatment discontinuation will also be conducted.

If data warrants, the counts and percentages of participants who follow the planned escalation scheme, have dose interruption, or have dose de-escalation/re-escalation, will be summarized for the open-label lead-in period. 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 visit during the dose escalation period.

6.4.4. Efficacy Analysis

Unless otherwise noted, statistical tests will be conducted at a 2-sided alpha level of 0.05, and the confidence interval will be calculated at 95% 2-sided.

Table GPHN.6.2. Efficacy Analysis for the Open-Label Period

Analysis Variable	Analysis Set	Analysis Method	Additional Information
Mean change in: -body weight -waist circumference -BMI	Data from Week 0 to Week 36	Change from baseline to postbaseline values at each scheduled visit will be summarized. Within treatment difference will be assessed by Wilcoxon signed-rank test	Missing data at Week 36 will be imputed using LOCF.
Mean change in: - Fasting glucose (mg/dL) - HbA1c (%) - fasting insulin (pmol/L) - Total cholesterol (mg/dL) - LDL-cholesterol (mg/dL) - Non-HDL-cholesterol (mg/dL) - HDL-cholesterol (mg/dL) - VLDL-cholesterol (mg/dL) - Triglycerides (mg/dL) - Free Fatty acids (mg/dL) -systolic blood pressure (mmHg) -diastolic blood pressure (mmHg)			
Mean change in health outcome measurements (Section 4.7.1)			
Percentage of study participants who achieve ≥5/10/15/20/25%body weight reduction		The counts and percentages will be summarized at each scheduled visit.	

Abbreviations: BMI = body mass index; HbA1c = hemoglobin A1c; HDL = high-density lipoprotein; LDL = low-density lipoprotein; LOCF = last observation carried forward; VLDL = very low-density lipoprotein.

6.4.5. Treatment Compliance

The definition of treatment compliance can be found in Section 6.2. It will be summarized descriptively at each treatment visit and over the open-label period for all enrolled participants.

6.4.6. Medical History, Preexisting condition, and Concomitant Medicines

Medical history, preexisting condition and concomitant medicines for all enrolled participants will be analyzed in a similar manner to Section 6.1 without performing treatment comparison.

6.4.7. Extent of Exposure

Summary of duration of open-label lead-in period (defined as time in days from date of Visit 2 to the date of randomization) and duration on study treatment (defined as time in days from date of first dose of study treatment on or after Visit 2 to date of last dose of study treatment during open-label period prior to randomization plus 7 days) will be provided for all enrolled participants.

For the summary of duration on study treatment, the frequency and percentage of participants in the following ranges will be summarized as well:

- >4 weeks
- ≥8 weeks
- ≥12 weeks
- ≥16 weeks
- ≥20 weeks
- ≥24 weeks, and
- ≥36 weeks.

In addition, the frequency and percentages of participants in the following study treatment exposure ranges may be summarized:

- 0 weeks
- >0 to <4 weeks
- ≥4 to <8 weeks
- ≥8 to <12 weeks
- ≥12 to <16 weeks
- ≥16 to <20 weeks
- ≥20 to <24 weeks
- ≥24 to <36 weeks, and
- ≥36 weeks.

6.4.8. Adverse Events

For the open-label lead-in period, a TEAE is defined as an event, based upon MedDRA LLT, that first occurred or worsened in severity after baseline.

For events occurring on the day of first taking study medication, the CRF-collected information (eg, 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.

Treatment-emergent adverse event, common TEAE, deaths, SAE, and other significant AE will be summarized in a similar way of Section 4.6.2 without performing treatment comparison.

6.4.9. Additional Safety Assessments

Vital signs, ECG parameters, and clinical laboratory results during the entire open label period (including the off-drug follow-up time period for participants who discontinued the study during the open label period) will be summarized in a similar way of Section 4.6.3 without performing treatment comparison.

6.4.10. Special Safety topics

Adverse events of special interest will be summarized in a similar way of Section 4.6.4 without performing treatment comparison.

6.4.11. Important Protocol Deviations

Important protocol deviations are identified in the Trial Issues Management Plan. A summary of important protocol deviations will be provided at the end of 36 weeks treatment (for all enrolled participants).

6.5. Appendix 5: Clinical Trial Registry Analyses

Additional analyses will be performed for the purpose of fulfilling the CTR requirements.

Analyses provided for the CTR requirements include the following:

- Summary tables of adverse events for the open-label period and the double-blind treatment period, provided as datasets which will be converted to XML files. For the open-label period, serious adverse events and 'other' non-serious adverse events are summarized by MedDRA preferred term. For the double-blind treatment period, they are summarized by treatment group and MedDRA preferred term.
- An adverse event is considered 'Serious' whether or not it is a TEAE.
- An AE 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, and
  - the number of events experienced
- 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, and
  - the total number of deaths causally related to treatment.

Demographic table including the following age ranges required by EudraCT: 18 to 65 years, 65 to 85 years, and 85 years and over.

6.6. Appendix 6: Exceptional Circumstances Impact

This section lists additional statistical analyses that may be performed at the final database lock to assess the impact of exceptional circumstances if the data warrant.

6.6.1. General Consideration

Percentage and count of randomized participants who followed the exceptional circumstances 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.

Percentage and count of randomized participants who completely missed at least 1 study visit due to exceptional circumstances may also be summarized.

Similar summaries during the open-label lead-in period may also be provided for all enrolled participants.

6.6.2. Exposure

A listing of enrolled participants who had study drug temporary interruption due to exceptional circumstances may be provided with information to indicate whether a participant is randomized (along with randomized treatment) and whether events occur post-randomization.

6.6.3. Protocol Deviation

Percentage and count of randomized participants having important protocol deviation related to exceptional circumstances will be summarized by treatment.

Percentage and count of randomized participants with protocol deviation and mitigation related to exceptional circumstances may also be summarized by treatment.

A listing of all randomized participants who had important protocol deviation and mitigation due to exceptional circumstances may be provided.

Similar summaries during the open-label lead-in period may also be provided for all enrolled participants.

6.6.4. Patient Disposition

A summary table for all randomized participants that discontinue study or study treatment due to exceptional circumstances will be provided by treatment. A similar summary during the open-label lead-in period may also be provided for all enrolled participants.

A listing of all enrolled participants who discontinued the study or study treatment due to exceptional circumstances will be provided with information to indicate whether a participant is randomized (along with randomized treatment).

6.6.5. Adverse Events

A listing of all enrolled participants who had COVID-19 infection, including death due to COVID-19, will be provided with information to indicate whether a participant is randomized (along with randomized treatment), and whether events occur post-randomization.

6.6.6. 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 subgroup impacted by exceptional circumstances (ie, participants without impact versus with impact) for SS group.

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

Similar summaries during the open-label lead-in period may also be provided for all enrolled participants.

6.6.7. Local Lab

Local lab due to exceptional circumstances will not be brought into the Lilly database at the time of primary database lock and final database lock per data collection system in Study GPHN, even though local laboratory is one of the options in exceptional circumstances. Therefore, this section is not applicable for analysis purposes.

6.6.8. Missing Data Due to Exceptional Circumstances

For the primary endpoints and key secondary endpoints, missing data due to exceptional circumstances will be handled as described in Section 4.1 and Section 4.3.2.3.

6.7. Appendix 7: Searching Criteria for Special Safety Topics

Pancreatitis Events

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

Hepatic Events

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

- 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-infectious 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)

Acute Gallbladder Disease

- 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).

Injection Site Reactions

Treatment-emergent injection site reaction will be identified using the MedDRA PT in any of the following:

- MedDRA High Level Term (HLT) of Injection site reaction
- HLT of Administration site reaction, and
- HLT of Infusion Site Reactions.

Major Depressive Disorder/Suicidal Ideation

The major depressive disorder/suicidal ideation events will be identified using the PTs from the Depression and suicide/self-injury SMQ as defined in MedDRA (SMQs narrow scope: 20000037 [Suicide/self-injury] and 20000167 [Depression (excl suicide and self injury)]).

Malignancies

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.

Arrhythmias and Cardiac Conduction Disorders

The treatment-emergent arrhythmias and cardiac conduction disorders events will be identified using the MedDRA PTs contained in any of the following SMQs:

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

7. References

[ADA]. Glycemic targets: standards of medical care in diabetes- 2020. *Diabetes Care.* 2020(suppl 1):S66–S76. https://doi.org/10.2337/dc20-S006

[ICH] International Council for Harmonisation. Addendum on estimands and sensitivity analysis in clinical trials to the guideline principles for clinical trials: E9(R1). Published 20 November 2019. Accessed August 31, 2022. https://www.gmp-compliance.org/files/guidemgr/E9-R1_Step4_Guideline_2019_1203.pdf

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.

Signature Page for VV-CLIN-076401 v1.0

Approval	PPD 03-Feb-2023 19:49:16 GMT+0000

Signature Page for VV-CLIN-076401 v1.0