Statistical Analysis Plan J2A-JE-GZPE (3)

A Phase 3, Long-term Safety Study of LY3502970 in Adult Participants with Type 2 Diabetes and Inadequate Glycemic Control with Diet and Exercise Alone or in Combination with Oral Antihyperglycemic Medications (ACHIEVE-J)

NCT06010004

Approval Date: 10-Jun-2025

Title Page

Protocol Title:

A Phase 3, Long-term Safety Study of LY3502970 in Adult Participants with Type 2 Diabetes and Inadequate Glycemic Control with Diet and Exercise Alone or in Combination with Oral Antihyperglycemic Medications (ACHIEVE-J)

Protocol Number: J2A-JE-GZPE

Compound Number: Orforglipron (LY3502970)

Short Title: A Long-term Safety Study of Orforglipron in Participants with Type 2 Diabetes

Acronym: ACHIEVE-J

Sponsor Name: Eli Lilly Japan K.K.

Legal Registered Address: Kobe, Hyogo, Japan

Regulatory Agency Identifier Number(s): Not applicable

Document ID: VV-CLIN-124254

Table of Contents

Table of Contents

Tables **Page**

Table of Contents

Figure **Page**

Version history

This Statistical Analysis Plan (SAP) for Study J2A-JE-GZPE is based on the GZPE(b) Harmonized Clinical Protocol dated 11 July 2024 and GZPE (1.1) Harmonized Clinical Protocol Addendum dated 11 July 2024. Please refer to Appendix 6, Section 6.6 for the statistical analysis for the Protocol GZPE addendum.

Table GZPE.1.1. SAP Version History Summary

SAP Version	Approval Date	Change	Rationale
1	05 September 2023	Not Applicable	Original version
2	29 October 2024	Converted the description and derivation for endpoints into a tabular format.	To align with protocol (b)
			To align with global study SAP
3	See date on Page 1	Aligned multiple imputation approach and MMRM model analysis for safety with other global study. Aligned with section 4.7.1 and added Appendix 2.	To align with global study SAP
		Updated analysis plan / data handling for SMBG, EBQ, DTSQ and Appendix 6. Deleted the description related to comparison with reference arm.	To be clarified.
		Added subgroup class in Section 4.7.1	Request from medical team.
		Added Appendix 7 (Section 6.7) for the integrated analysis with Japanese subset in Study J2A-MC-GZGT	For submission purposes.

Abbreviation: DTSQ = Diabetes Treatment Satisfaction Questionnaire; EBQ = Eating Behavior Questionnaire; MMRM = mixed model for repeated measures; SAP = statistical analysis plan; SMBG = self-monitored blood glucose.

1. Introduction

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

1.1. Objectives, Endpoints, and Estimands

Objectives	Endpoints
Primary	
To assess the safety and tolerability of orforglipron (3, 12, or 36 mg) administered QD during 52-week treatment as monotherapy or an add-on therapy to SU, biguanides, SGLT-2i, a-GI, TZD, or glinides in the treatment of T2D.	• Incidence of TEAEs
Secondary	
To assess the safety and tolerability of orforglipron (3, 12, or 36 mg) administered QD during 52-week treatment as monotherapy or an add-on therapy to SU, biguanides, SGLT-2i, a-GI, TZD, or glinides in the treatment of T2D.	Summary of safety data, including number and incidence of: • SAEs • discontinuations from study intervention or study due to AEs, and • AEs of special interest.

To assess the efficacy of orforglipron (3, 12, or 36 mg) administered QD during 52-week treatment as monotherapy or an add-on therapy to SU, biguanides, SGLT-2i, a-GI, TZD, or glinides in the treatment of T2D.	From baseline to Week 52 or at Week 52:
	• change from baseline in HbA1c
	• proportion of participants who achieve HbA1c
	  ○ <7.0% (<53 mmol/mol)
	  ○ ≤6.5% (≤48 mmol/mol)
	  ○ <5.7% (<39 mmol/mol)
	• change from baseline in fasting serum glucose (central laboratory)
	• change from baseline in body weight
	• percentage change from baseline in body weight
	• proportion of participants who achieve weight loss of ≥5%, ≥10%, and ≥15%
	• change from baseline in waist circumference
	• change from baseline in BMI, and
	• change from baseline in daily average 7-point SMBG.

Exploratory	

To assess the effects of orforglipron (3, 12, or 36 mg) on lipid parameters.	• Percentage change from baseline in lipid parameters to Week 52:
	  ○ total cholesterol
	  ○ HDL-cholesterol
	  ○ LDL-cholesterol
	  ○ VLDL-cholesterol
	  ○ non-HDL cholesterol, and
	  ○ triglycerides.

Objectives	Endpoints
To assess the effects of orforglipron (3, 12, or 36 mg) on hepatic parameters.	• Percentage change from baseline to Week 52: o ALT, and o AST. • Change from baseline to Week 52: o Fib-4 index, and o FLI
To assess the effects of orforglipron (3, 12, or 36 mg) on blood pressure.	• Change from baseline in blood pressure to Week 52: o systolic blood pressure, and o diastolic blood pressure.
To assess the effects of orforglipron (3, 12, or 36 mg) on markers of glucose metabolism and beta-cell function.	• Change from baseline to Week 52: o fasting insulin o fasting C-peptide o C-peptide index o fasting glucagon o HOMA-IR, and o HOMA-B.
To assess the effects of orforglipron (3, 12, or 36 mg) on patient-reported outcomes.	• DTSQc scores at Week 52 • DTSQs scores at baseline • Change from baseline to Week 52: o EQ-5D-5L health state utilities and VAS, and o Eating Behavior Questionnaire.
To assess the effect of orforglipron (3, 12, or 36 mg) on body composition. (Measured by InBody770 at particular study sites)	• Change from baseline to Week 52: o body weight o body fat mass, and o lean body mass.

Abbreviations: AE = adverse event; a-GI = alpha-glucosidase inhibitor; ALT = alanine aminotransferase; AST = aspartate aminotransferase; BMI = body mass index; DTSQ = Diabetes Treatment Satisfaction Questionnaire; Fib-4 = Fibrosis-4; FLI = fatty liver index; HbA1c= hemoglobin A1c; HDL = high density lipoprotein; HOMA-B = homeostasis model assessment of beta cell function; HOMA-IR = homeostatic model assessment for insulin resistance; LDL = low density lipoprotein; QD = once daily; SAE = serious adverse event; SGLT-2i = sodium-glucose co-transporter type 2 inhibitor; SMBG= self-monitored blood glucose; SU = sulfonylurea; T2D = type 2 diabetes; TEAE = treatment-emergent adverse event; TZD = thiazolidinedione; VAS = visual analog scale; VLDL = very-low-density lipoprotein.

Primary estimand

The primary clinical question of interest is:

What is the safety and tolerability of orforglipron during 52-week treatment as monotherapy or an add-on therapy to oral antihyperglycemic medication (OAM), in participants who meet the eligibility criteria irrespective of study treatment adherence or initiation of additional antihyperglycemic medications?

The estimand for the primary endpoint is described by these attributes:

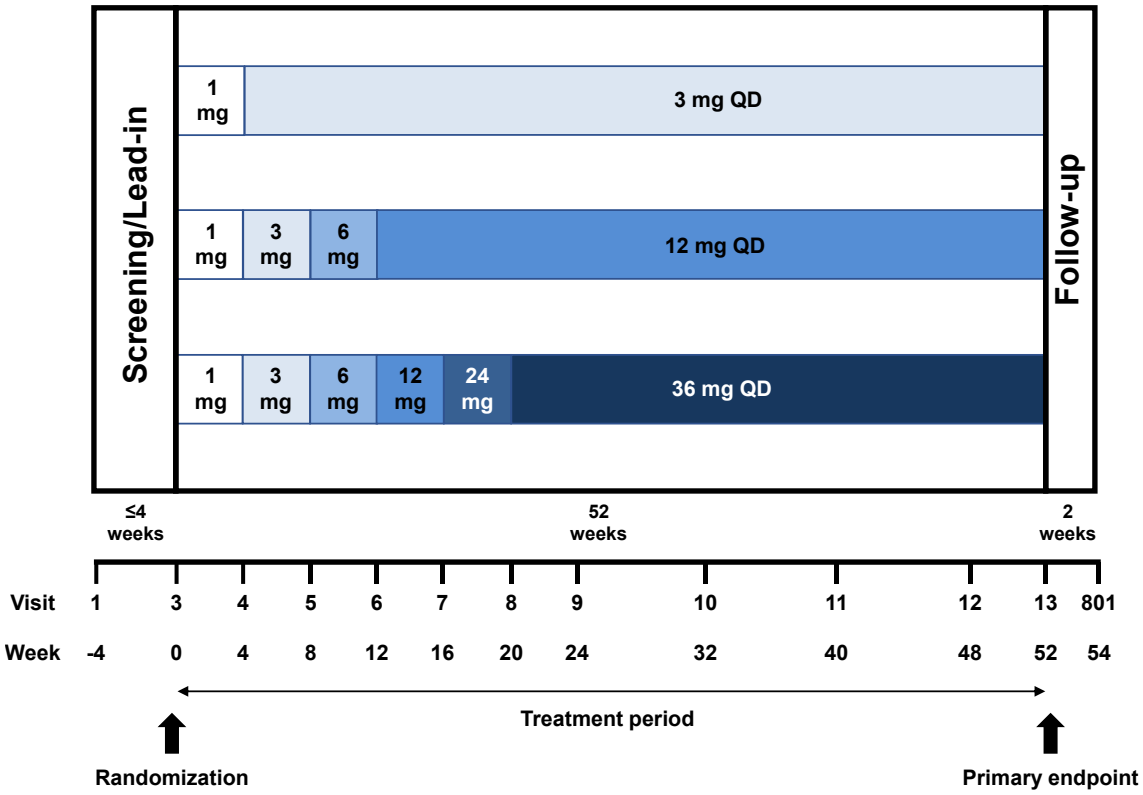
- ***Population*** – participants who meet the eligibility criteria.
- ***Primary endpoint*** – incidence of treatment-emergent adverse events (TEAEs).
- ***Treatment condition –*** the randomized treatment with allowance for potential dose interruptions and modifications, regardless of adherence to study intervention or initiation of additional antihyperglycemic medications.
- ***Intercurrent events (ICEs)*** – no ICEs are defined since treatment adherence and the initiation of additional antihyperglycemic medications are a part of the treatment condition.
- ***Population-level summary*** – frequency counts and percentages of TEAEs during the treatment period plus the safety follow-up period.
- ***Rationale for the estimand*** – this estimand aims to assess the safety and tolerability of orforglipron regardless of study treatment adherence.

1.2. Study Design

ACHIEVE-J (Study GZPE) is a Phase 3, multicenter, randomized, open-label study to assess the safety and efficacy of orforglipron (LY3502970) (3, 12, or 36 mg) in Japanese participants with type 2 diabetes (T2D) who have inadequate glycemic control with diet and exercise alone or who have been taking therapeutic doses of various OAMs; sulfonylureas (SU), biguanides, or sodium-glucose co-transporter type 2 inhibitors (SGLT-2i) alone, or alpha-glucosidase inhibitors (a-GI), thiazolidinedione (TZD), or glinides alone or in combination with metformin.

This study includes 3 periods as follows:

- screening and lead-in period: up to 4 weeks
- treatment period: 52 weeks, including 20 weeks of dose escalation, and
- safety follow-up period: 2 weeks.



Abbreviations: QD = once daily

Figure GZPE.1.1. Illustration of study design for Protocol GZPE.

Screening/Lead-in Period

Eligibility for Study GZPE will be determined at the initial screening visit (Visit 1). Screening procedures and participant training will be performed at Visits 1 and 2 (screening and lead-in periods). Participant training will include disease monitoring and management procedures, study eDiary, and study procedures.

Treatment Period

At Visit 3, participants will perform all required baseline study procedures (including the collection of all baseline laboratory measures) prior to random assignment and prior to taking the first dose of study intervention. Following random assignment, the participant will take the first dose of study intervention at the study site, according to the dose escalation regimen. The date and time of the first dose of study intervention should be recorded on the case report form (CRF).

Beginning at random assignment, all participants will receive study intervention according to the randomized treatment arm for the duration of the 52-week treatment period.

Safety Follow-up Period

A safety follow-up visit will occur approximately 2 weeks following the last dose of the study intervention.

2. Statistical Hypotheses

No formal hypothesis testing is planned for this study.

2.1. Multiplicity Adjustment

No multiplicity adjustments will be made because no formal hypothesis testing is planned.

3. Analysis Sets

The analysis sets are defined as

Participant Analysis Set	Description
Entered participants	All participants who sign informed consent[a].
Randomized participants	All participants who are randomly assigned a study intervention.
Safety participants	All participants who are randomly assigned a study intervention and who take at least 1 dose of study intervention.

[a] Refers to the informed consent for the study.

The data points sets are defined as

Data Points Sets	Description
Safety data points set	All data points obtained during the treatment period and the follow-up period defined as at or after baseline and up to the date of study withdrawal including the follow-up period, regardless of study intervention discontinuation or initiation of additional antihyperglycemic medications[a].
Treatment regimen estimand data points set	All data points obtained during the treatment period defined as at or after baseline and up to the last visit within the treatment period, regardless of study intervention discontinuation or initiation of additional antihyperglycemic medications[a].
Efficacy estimand data points set	All data points obtained during the treatment period defined as at or after baseline and up to the earliest date of discontinuation of study intervention or initiation of additional antihyperglycemic medications[a].

[a] "Additional antihyperglycemic medications" refers to any antihyperglycemic therapy that is initiated after baseline and used for more than 2 weeks (14 days).

4. Statistical Analyses

4.1. General Considerations

Statistical analysis will be the responsibility of Eli Lilly and Company (Lilly) or its designee. Some analyses and summaries described in this analysis plan may not be conducted if warranted by data (that is, too few events to justify conducting an analysis). Additional analyses of the data may be conducted as deemed appropriate without further changes made to the protocol or SAP, even after the final database lock.

Unless otherwise noted, all tests will be conducted at the 2-sided alpha level of 0.05, and the confidence intervals (CIs) will be calculated at 95%, 2-sided.

Participants will be analyzed according to the treatment group to which they were randomly assigned.

Change from baseline will be calculated as the value of interest at the visit minus the baseline value. In general, percent change from baseline will be calculated as the value of change from baseline divided by the baseline value in 100% scale. Some specific predefined parameters may be log-transformed before statistical analysis, if deemed necessary. If the baseline value is missing for a particular variable, then the change from baseline and percent change from baseline will not be calculated when deriving the summary statistics.

For continuous measures, summary statistics will include sample size, mean, standard deviation (SD), median, minimum, and maximum for the actual, the change from baseline and if applicable, the percent change from baseline measurements as well. Least squares (LS) means and standard errors derived from the analysis models will also be displayed for the actual value and the change or percent change from baseline measurements. For analysis of log-transformed parameters, model estimated means and standard errors on the original scale will be derived through back-transformation using the delta method from the LS means and standard errors on the natural log-scale. All baseline measures will be analyzed using an analysis of variance (ANOVA) model that has treatment group as the fixed effect.

For categorical measures, summary statistics will include sample size, frequency, and percentages. If applicable, Fisher's exact test or Pearson's chi-squared test will be used for treatment comparisons.

For time-to-event measures, participants without an event will be censored, and time-to-event will be the number of days between the start date and the date of the participant's end of follow-up + 1 day. For participants experiencing the event, "time-to-first-event" will be the time (in days) from start date to the first occurrence of the event + 1 day. For safety (if needed) and study intervention discontinuation analyses, the date of first dose of study intervention will be used as the start date. For other analyses (for example, study discontinuation), the randomization date will be used as the start date (in rare cases where randomization occurred prior to Visit 3 date, Visit 3 date will be used instead).

Data may exist at postbaseline visits where a variable is not scheduled to be collected. Data from unscheduled visits and from early discontinuation (ED) visits that do not correspond to a scheduled visit will be excluded from mixed-model for repeated measures (MMRM) analysis or analysis of covariance (ANCOVA), unless otherwise specified.

For stratification factors, a single strata variable will be created for hemoglobin A1c (HbA1c) at baseline (≤8.5%, >8.5%) and background therapy (Yes or No) (diet and exercise alone versus SU, biguanide, SGLT-2i, a-GI, TZD, or glinide). Unless otherwise specified, stratification factors will be used in the analysis models including MMRM and ANCOVA.

Details about the analyses regarding demographic and baseline characteristics, historical illnesses and preexisting conditions, treatment compliance, concomitant medications, and important protocol deviations can be found in Appendices 1 through 5 (Sections 6.1, 6.2, 6.3, 6.4, and 6.5), respectively.

The analysis model for continuous measures assessed over time will be an MMRM, with fixed effects of visit, and treatment, stratification factors, and baseline measurement all nested within visits. An unstructured covariance structure will model the relationship of within-patient errors. The analysis model for single time point measures will be an ANCOVA model with adjustment for baseline and stratification factors.

For efficacy and safety analyses, the overall population and each background therapy analysis for the selected outputs will be conducted separately, using the same statistical model. When MMRM or ANCOVA is used for each background therapy, the background therapy will be excluded from the model. If each background therapy analysis cannot be conducted by model analysis due to small sample size etc., the descriptive analysis or other simplified model may be used.

Detail about the analysis for the Study GZPE (1) Harmonized Clinical Protocol Addendum can be found in Appendix 6 (Section 6.6) and detail about the integrated analysis for Japanese participants can be found in Appendix 7 (Section 6.7).

4.1.1. Baseline Definition

Unless otherwise specified, the baseline for efficacy assessments is defined as the last available non-missing measurement to the first dose of study intervention. In most cases, this will be the measurement recorded at Week 0 (Visit 3). If there are no doses of study intervention administered, baseline will be defined as the last available non-missing measurement on or prior to randomization.

For the safety-related parameters, the definition of baseline and postbaseline are specified in Table GZPE.4.1.

In cases where the measurement is taken on the same day (where the time is not collected or not reliable) as the first dose, this measurement will be used as the baseline value for data analysis. For patient-reported outcome measures, data obtained at Visit 3, regardless of the timing relative to the first dose, will serve as the baseline.

Table GZPE.4.1. Baseline and Postbaseline Definitions for Safety Analyses

Analysis Type	Participant Analysis Set/Data Points Set	Baseline Period	Postbaseline Period
Treatment-emergent adverse events	Safety participants / safety data points set	Start of screening and ends prior to the first dose of study intervention (typically at Week 0).	Starts after the first dose of study intervention and ends at the end of the safety follow-up period or the date of study withdrawal (whichever is earlier)[a].
Treatment-emergent abnormal labs, vital signs, and ECGs.	Safety participants / safety data points set	Start of screening and ends prior to the first dose of study intervention (typically at Week 0). All scheduled and unscheduled measurements will be included.	Starts after the first dose and ends at the end of the safety follow-up period or the date of study withdrawal (whichever is earlier). All scheduled and unscheduled measurements will be included.
Change from last baseline to each postbaseline week and to last postbaseline for laboratory values, vital signs, and ECGs.	Safety participants / safety data points set	When the term "last baseline" is used, baseline starts from informed consent date and ends prior to the first dose (typically at Week 0).	Starts after the first dose and ends at the end of the safety follow-up period or the date of study withdrawal (whichever is earlier). Only scheduled visits will be included. The early discontinuation visits are considered scheduled visits.

Abbreviation: ECG = electrocardiogram.

[a] For events occurring on the day of first dose, the case report form collected information will be used to determine whether the event was pre- versus posttreatment if available. If the relevant information is not available, then the events will be counted as postbaseline.

4.1.2. Analysis Methods

The primary and secondary safety endpoints will be guided by an estimand comparing the safety of the orforglipron arms, irrespective of adherence to study intervention or initiation of additional antihyperglycemic medications.

The selected secondary and exploratory efficacy endpoints will be evaluated based on both the treatment regimen estimand and the efficacy estimand. The treatment regimen estimand will be the primary focus for endpoints analyzed with both. All other efficacy and exploratory endpoints will be conducted based on the efficacy estimand.

Unless otherwise specified, safety and tolerability assessments will be guided by an estimand comparing the safety of the orforglipron treatment arms for the entire study period (the treatment period plus safety follow-up), irrespective of adherence to study intervention or initiation of additional anti-hyperglycemic medications for all study population.

It is noted that for some safety endpoints, such as treatment discontinuation due to adverse events and laboratory values that are planned to be measured only during the treatment period, the analyses will be based on data observed during the treatment period only.

The estimands, participant analysis sets, analysis data points sets, and analysis models for efficacy and safety endpoints are summarized in Table GZPE.4.2.

Table GZPE.4.2. Estimands and Analysis Models for All Endpoints

	Treatment Regimen Estimand	Efficacy Estimand	Estimand Guiding Safety
Participants	Randomized	Randomized	Safety
Data points set	Treatment regimen	Efficacy estimand	Safety
Primary and secondary safety endpoints			X
Selected secondary efficacy endpoints	X	X	
Other secondary efficacy and exploratory endpoints		X	
Safety and tolerability assessments			X
Analysis model	• Continuous: ANCOVA with missing endpoints imputed[a]. • Binary: Descriptive statistics with/without Missing endpoints imputed[a].	• Continuous: MMRM (if there are multiple longitudinal postbaseline measurements) or ANCOVA (if there is only 1 postbaseline measurement) with missing endpoints imputed[b]. • Binary: Descriptive statistics with/without Missing endpoints imputed[b].	• Continuous: MMRM (if there are multiple longitudinal postbaseline measurements) or ANCOVA (if there is only 1 postbaseline measurement) with no imputation to be conducted. • Binary: Fisher's exact or Pearson's chi-squared test if applicable; no imputation will be conducted.

Abbreviations: ANCOVA = analysis of covariance; CI = confidence interval; MMRM = mixed model for repeated measures; X = apply.

[a] The treatment regimen estimand imputation method is described in Section 4.1.2.1.3.

[b] The efficacy estimand imputation method is described in Section 4.1.2.2.3.

4.1.2.1. Analysis Methods for Treatment Regimen Estimand

The treatment regimen estimand will be applied to the secondary efficacy measure, if necessary.

4.1.2.1.1. Method for Continuous Variables

For the assessment of continuous efficacy measures guided by the treatment regimen estimand, data for participants with missing values at Week 52 will be imputed based on methods described in Section 4.1.2.1.3 and then analyzed using an ANCOVA with a variance estimator that is robust to model misspecification and heteroscedasticity (Ye et al. 2022). The model will include

- treatment group as a factor variable
- strata variable as a factor variable
- baseline value (of the dependent variable)
- interaction between treatment group and strata variable, and
- interaction between treatment group and baseline value.

If the model fails to converge, the interaction terms will be removed.

Some parameters may be log-transformed before fitting the ANCOVA as specified in Sections 4.4, 4.5, and 4.6, with the associated log-transformed baseline value as a covariate. In these cases, LS means and standard error for each treatment group will be back-transformed and presented as mean percent change from baseline.

4.1.2.1.2. Method for Binary Variables

For binary efficacy measures derived from a continuous variable and guided by the treatment regimen estimand, the missing value in the underlying continuous variable will be imputed first based on the methods described in Section 4.1.2.1.3, and then the corresponding binary variable will be derived.

The binary efficacy measures will be summarized by treatment and visit. These will be estimated using two methods: descriptive statistics at each visit and descriptive statistics with the missing value imputed at Week 52.

4.1.2.1.3. Primary Multiple Imputation Strategy

Discontinuation is expected to be uncommon. Participants who discontinue the study intervention (that is, discontinue study treatment) will be encouraged to continue in Study GZPE for the treatment period and follow-up period. In this section, study discontinuation refers to treatment phase discontinuation as captured in the CRF.

If there are occurrences of missing data despite all reasonable precautions, missing data will be imputed in a manner consistent with what the values would likely have been, had they been collected.

In general, 5 scenarios are considered regarding the treatment journey of a trial participant relative to the primary time point (the Week 52 visit) shown in Figure GZPE.4.1.

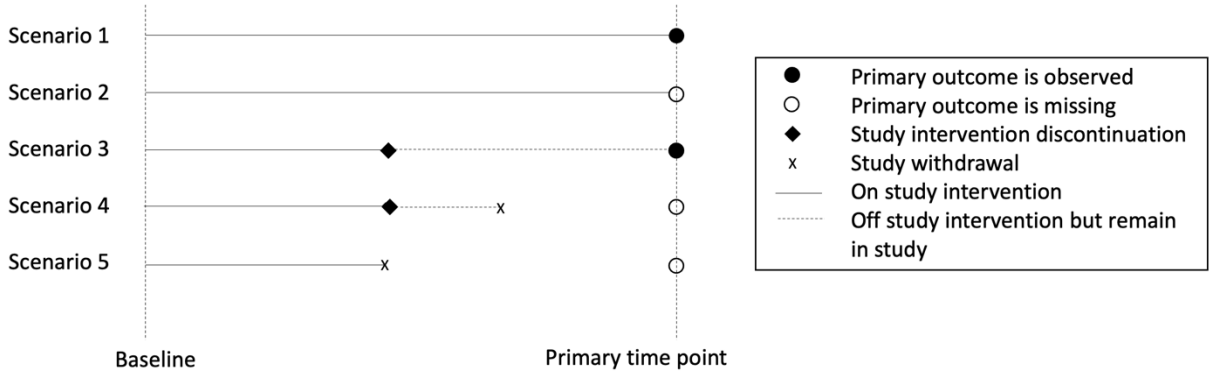


Figure GZPE.4.1. Treatment journey of a trial participant in relation to permanent discontinuation of study intervention or treatment period.

Scenario 5 covers situations where a participant discontinues from the study intervention and then comes back for a study discontinuation visit for the same discontinuation reason on or prior to a scheduled visit.

In principle, missing data due to permanent discontinuation of the study intervention (Scenarios 4 and 5A) will be imputed by treatment group using retrieved dropouts (MI-RD), that is, using multiple imputation based on data retrieved from participants who permanently discontinued the study intervention but continued in Study GZPE with non-missing measurements from the same treatment group. For Scenarios 4 and 5A, if there are not enough retrieved dropouts to provide a reliable imputation model (that is, the model implemented does not converge), an alternative multiple imputation method (return to baseline) will be used. Any missing values at baseline will be imputed in the same model when imputing the endpoint.

Table GZPE.4.3. Imputation Procedure

Scenario (Study Intervention/Study Discontinuation; Missingness at Endpoint)	Methods to Handle Missing Values at Endpoint
1. No study intervention or study discontinuation; no missing value	N/A
2. No study intervention or study discontinuation; with missing value	Use observed data, including baseline and all post-baseline visits, across all time points from Scenarios 1 and 2 to impute under the MAR assumption by treatment group. There should be very few such cases in a clinical trial. No additional covariates will be included in the imputation model.
3. Study intervention discontinuation with no study discontinuation; no missing value	N/A
4. Study intervention discontinuation with study discontinuation at a subsequent visit or completion of treatment period; with missing value	Missing values at the endpoint visit will be imputed through MCMC (predictive mean matching method), using baseline and endpoint visit data from Scenario 3 (MI-RD) and impute missing values by treatment group. No additional covariates will be included in the imputation model.
5. Study discontinuation resulting in study intervention discontinuation; with missing value	5A: If the study discontinuation is possibly related to study medication (see study discontinuation reasons classified as Category 5A in Table GZPE.4.4), missing values at the endpoint visit will be imputed in the same way as Scenario 4. 5B: If the study discontinuation is clearly due to administrative reasons not related to study intervention (see study discontinuation reasons classified as Category 5B in Table GZPE.4.4), missing values at the endpoint visit will be imputed in the same way as Scenario 2.

Abbreviations: MAR = missing at random; MCMC = Markov chain Monte Carlo; MI-RD = multiple imputation using retrieved dropouts; N/A = not applicable.

Table GZPE.4.4. Study Discontinuation Reasons

Disposition Reason	Associated Sub-categories	Category
Adverse event		5A
Death		5A
Protocol deviation		5A
Pregnancy		5A
Non-compliance with study drug		5A
Lack of efficacy		5A
Withdrawal by subject	Concern about study procedures/perceived risks	5A
	Scheduling conflicts	5A
	Subject is moving or has moved	5A
	Personal issue unrelated to trial	5A
	Due to epidemic/pandemic	5B
	Other (option to include a specify field)	5A
Physician decision	Concern about study procedures/perceived risks	5A
	Due to epidemic/pandemic	5B
	Other (option to include a specify field)	5A
Study terminated by sponsor		5B
Site terminated by sponsor		5B
Study terminated by IRB/ERB		5B
Lost to follow-up		5A
Other		5A

Abbreviations: ERB = ethical review board; IRB = institutional review board.

4.1.2.2. Analysis Methods for Efficacy Estimand

4.1.2.2.1. Method for Continuous Variables

For the assessment of continuous efficacy measures guided by the efficacy estimand, a restricted maximum likelihood-based MMRM (Wang and Du 2022) analysis will be used. All the longitudinal observations at each scheduled postbaseline visit will be included in the analysis. The MMRM model will include the following terms:

- treatment group as a factor variable
- visit as a factor variable
- strata variable as a factor variable
- baseline value (of the dependent variable)
- interaction between treatment group, visit, and strata variable, and
- interaction between treatment group, visit, and baseline value.

The LS means at the scheduled visits will be reported with the variability estimated using the robust inference as provided in Wang and Du (2022). An unstructured covariance matrix by

treatment groups will be used to model the within-participant errors, assuming heteroscedasticity (variant covariance structure across treatment groups) and the measurements for different participants are independent. If model fails to converge, the model will be simplified to include the following terms, removing 3-way interactions and the heteroscedasticity assumption:

- treatment group as a factor variable
- visit as a factor variable
- strata variable as a factor variable
- baseline value (of the dependent variable)
- interaction between visit and treatment group
- interaction between visit and strata variable, and
- interaction between visit and baseline value.

If the model still fails to converge, the following covariance structures will be tested in this order for the simplified model:

- unstructured covariance matrix
- heterogeneous Toeplitz
- heterogeneous autoregressive
- heterogeneous compound symmetry
- homogeneous Toeplitz
- homogeneous autoregressive, and
- homogeneous compound symmetry.

The first covariance structure that converges will be used.

Some parameters may be log-transformed before fitting the MMRM as specified in Sections 4.4, 4.5, and 4.6. In these cases, LS means and standard errors for each treatment group will be back-transformed and presented as mean percent change from baseline.

If the data does not warrant the use of an MMRM model, then an ANCOVA model will be conducted, as described in Section 4.1.2.1.1. Missing values will be imputed based on the methods described in Section 4.1.2.2.3.

4.1.2.2.2. Method for Binary Variables

For binary efficacy measures derived from a continuous variable and guided by the efficacy estimand, the missing value in the underlying continuous variable will be imputed, first based on methods described in Section 4.1.2.2.3, and then the corresponding binary variable will be derived.

The binary efficacy measures will be summarized by treatment and visit. These will be estimated using two methods: descriptive statistics at baseline and Week 52 and descriptive statistics with the missing value imputed at Week 52.

4.1.2.2.3. Handling of Missing Data

Missing data should be minimized at the best precaution. For continuous efficacy analyses guided by the efficacy estimand, the hypothetical strategy will be used to handle the ICEs. For the MMRM analysis, only data collected before the occurrence of any ICEs will be used. Through the MMRM, the potential efficacy measures (after the ICEs) will be implicitly imputed

as if participants did not have ICEs. For the ANCOVA analysis, missing values at the last visit will be imputed through multiple imputation using all non-missing data (excluding data collected after ICEs) from the same treatment group under the missing at random (MAR) assumption. The same method will also be used to impute the missing value in the underlying continuous variables of binary outcomes.

4.1.2.3. Analysis Methods for Safety

For categorical safety measures, treatment group differences of percentages will be conducted using Fisher's exact test or Pearson's chi-squared test, if applicable, unless otherwise specified.

For selected continuous safety measures, unless otherwise specified, mean change or mean percent change from baseline by treatment group at all scheduled visits will be assessed via an MMRM using restricted maximum likelihood (REML) which will include the following terms:

- treatment group as a factor variable
- visit as a factor variable
- baseline value (of the dependent variable), and
- interaction between visit and treatment group.

If the data does not warrant the MMRM model, then an ANCOVA model will be conducted as a fixed effect and the continuous baseline value as a covariate will be used.

No explicit imputation will be conducted for safety measures. Some parameters (such as, urine albumin-creatinine ratio [UACR], p-amylase, and lipase) may be log-transformed before fitting the MMRM as specified in Section 4.6. In these cases, LS means and standard errors for each treatment group will be back-transformed and presented as mean percent change from baseline.

For selected safety parameters, the Kaplan-Meier method will be used to estimate the cumulative event curve over time if required. Counts and proportions of participants who experience the event will be calculated by treatment group. Participants without an event will be censored at the end of Study GZPE participation. For participants experiencing the event, time-to-first-event will be the time (in days) from randomization to first occurrence of the event + 1 day.

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

Some safety analyses may be conducted after excluding data after the initiation of new anti-hyperglycemic medications.

4.2. Participant Dispositions

The participant dispositions for the screening period, the study intervention/treatment, the treatment period and/or the follow-up period will be collected in CRFs with the corresponding primary reason. The study completion for a participant is defined as the participant completing both the treatment period and the follow-up period, regardless of completion of study intervention.

The planned listings and summary tables for dispositions are provided in Table GZPE.4.5. No inferential analysis will be performed.

Table GZPE.4.5. Listings and Summary Tables Related to Dispositions

Analysis	Population/Period
Summary of disposition prior to randomization	Entered
Summary of study and study intervention disposition	Randomized/TP + FP (TP for study intervention disposition)
Participant allocation by center/site	Entered
Kaplan-Meier plot of time to study discontinuation	Randomized/TP + FP
Kaplan-Meier plot of time to study intervention discontinuation	Randomized/TP
Kaplan-Meier plot of time to study intervention discontinuation due to AE	Randomized/TP
Listing of randomization	Randomized
Listing of randomized participants who were discontinued from the study treatment due to inadvertent enrollment	Randomized
Listing of study and study intervention disposition	Randomized

Abbreviations: FP = follow-up period; TP = treatment period.

4.3. Primary Endpoint/Estimand Analysis

The primary endpoint is the incidence of TEAEs and will be evaluated using the safety participants and the safety data points set. The definition of endpoint and analytical approaches are described in Section 4.6.2.

4.4. Secondary Endpoints/Estimands Analysis

Table GZPE.4.6 includes the description and derivation of all the secondary endpoints for the efficacy measures.

Table GZPE.4.7 provides the detailed analyses including endpoint, estimand, data points set, method and imputation, population, and time point for the secondary endpoints.

The definition of endpoints and analytical approaches of secondary safety endpoints are described in Section 4.6.2.

The Kaplan-Meier method will be used to estimate the cumulative event curve over time for the selected endpoints.

For self-monitored blood glucose (SMBG), valid SMBG profiles will be used for analysis, defined as having non-missing values at ≥ 4 time points (including at least one post-meal point) among the 7 pre-specified time points and being collected within 2 weeks prior to the given visit. For each time point, the average of the corresponding SMBG values from the valid SMBG profiles will be used for analysis. In addition to the analyses of the mean of all 7-point measurements, similar analyses will be performed for each of the 7 points, the 2-hour morning, midday, and evening meal excursions, the mean of all meals' 2-hour excursion, the mean of all pre-meal measurements, and the mean of all 2-hour postprandial measurements. Additionally, if there is insufficient 7-point SMBG data available, but there is data from the 4-point SMBG form

with non-missing values at ≥ 4 time points, including at least one post-meal point, the data from the 4-point SMBG form will be used as 7-point data.

Table GZPE.4.6. Description and Derivation of Secondary Endpoints for Efficacy Measures

Measure	Description	Variable	Derivation/Comment	Handling Missing Components[a]
Glycemic control	HbA1c will be performed by the central laboratory and will be assayed by a Lilly-designated laboratory.	HbA1c (% and mmol/mol)	As provided	M1
		CHG HbA1c (% and mmol/mol)	Calculated as: postbaseline HbA1c – baseline HbA1c	M2
	Improvement in T2D is evaluated by HbA1c	T2D (HbA1c target value of ≤6.5%)	Response = yes if the postbaseline HbA1c value ≤6.5%	Missing if postbaseline value is missing
		HbA1c target value of <5.7%	Response = yes if the postbaseline HbA1c value <5.7%	Missing if postbaseline value is missing
		HbA1c target value of <7%	Response = yes if the postbaseline HbA1c value <7%	Missing if postbaseline value is missing
	Fasting glucose will be performed by the central laboratory and will be assayed by a Lilly-designated laboratory.	Fasting serum glucose (mg/dL and mmol/L)	As provided	M1
		CHG fasting serum glucose (mg/dL and mmol/L)	Calculated as: postbaseline fasting glucose – baseline fasting glucose	M2
Body weight	Per Protocol GZPE Section 10.10: • Body weight measurements should be done in a consistent manner using a calibrated electronic scale capable of measuring weight in kilograms to 1 decimal place. • All weights for a given participant should be measured using the same scale, whenever possible, at approximately the same time in the morning after evacuation of bladder contents.	Body weight (kg)	As measured	M1
		CHG in body weight (kg)	Calculated as: postbaseline body weight (kg) – baseline body weight (kg)	M2
		% CHG in body weight (%)	Calculated as: (postbaseline body weight [kg] – baseline body weight [kg]) / baseline body weight [kg] × 100 (%)	M2
		≥x% body weight reduction from baseline where, x = 5, 10, 15	Response = yes if at least a x% reduction in body weight from baseline, that is, percent change from baseline in body weight ≤ -x%	M2

Waist circumference	Per Protocol GZPE Section 10.10: • Waist circumference should be measured in the horizontal plane and at the umbilicus level according to JASSO guideline. • Measurements should be taken at the end of a normal expiration using a non-stretchable measuring tape. The tape should lie flat against the skin without compressing the soft tissue. The waist circumference should be measured twice, rounded to the nearest 0.5 cm. The measuring tape should be removed between the 2 measurements. Both measurements will be recorded in the CRF. If the difference between the 2 measurements exceeds 1 cm, this set of measurements should be discarded and the 2 measurements repeated.	Waist circumference (cm)	The average of 2 measures	If there is at least 1 measurement, take the average of all nonmissing measurements; otherwise, set it to be missing
		CHG in waist circumference (cm)	Calculated as: postbaseline waist circumference (cm) – baseline waist circumference (cm)	M2
BMI	BMI: Round to 1 decimal point. For example, a BMI of 26.6 kg/m^2 should **not** be rounded to 27.0 kg/m^2.	BMI (kg/m^2)	BMI will be calculated as: Weight(kg)/(Height[m])^2	Missing if weight or height is missing
		CHG in BMI (kg/m^2)	Calculated as: postbaseline BMI (kg/m^2) – baseline BMI (kg/m^2)	M2
SMBG	Participants must collect two 7-point SMBGs on 2 nonconsecutive days before assigned visits. A 7-point SMBG consists of measurements before and 2 hours after each of 3 main meals within the same day and at bedtime, that is: • Point 1: before morning meal fasting • Point 2: 2 hours after morning meal • Point 3: before midday meal • Point 4: 2 hours after midday meal • Point 5: before evening meal • Point 6: 2 hours after evening meal • Point 7: bedtime The participant needs to collect these	Point x SMBG (mg/dL) x = 1, 2, …,7	The average of 2 measures in the 2 nonconsecutive days at each time point.	If there is at least 1 measurement, take the average of all non-missing measurements; otherwise, set it to be missing
		CHG in point x SMBG (mg/dL)	Calculated as: postbaseline point x SMBG (mg/dL) – baseline point x SMBG (mg/dL)	M2
		Mean of all 7-point SMBG (mg/dL)	Calculated as: average of (average of point 1, average of point 2, …, average of point 7)	If there are at least 4 measurements, take the average of

	measurements within 14 days before the assigned visits. If more than two 7-point SMBG profiles are available, the 2 most recent nonconsecutive profiles should be used.			all nonmissing measurements; otherwise, set it to be missing
		CHG in mean of all 7-point SMBG (mg/dL)	Calculated as: postbaseline mean of all 7-point SMBG (mg/dL) – baseline mean of all 7-point SMBG (mg/dL)	M2
		Mean of pre-meal SMBG (mg/dL)	Calculated as: average of (average of point 1, average of point 3, average of point 5)	If there is at least 1 measurement, take the average of all nonmissing measurements; otherwise, set it to be missing
		CHG in mean of pre-meal SMBG (mg/dL)	Calculated as: postbaseline mean of pre-meal SMBG (mg/dL) – baseline mean of pre-meal SMBG (mg/dL)	M2
		Mean of 2-hour post-meal SMBG (mg/dL)	Calculated as: average of (average of point 2, average of point 4, average of point 6)	If there is at least 1 measurement, take the average of all non-missing measurements; otherwise, set it to be missing
		CHG in post-meal SMBG (mg/dL)	Calculated as: postbaseline mean of post-meal SMBG (mg/dL) – baseline mean of post-meal SMBG (mg/dL)	M2
		Morning pre-meal to 2-hour post-meal excursion	Calculated as: average of (point 2 - point 1)	If 1 excursion is available, then take that one.
		Midday pre-meal to 2-hour post-meal excursion	Calculated as: average of (point 4 - point 3)	If 1 excursion is available, then take that one.

		Evening pre-meal to 2-hour post-meal excursion	Calculated as: average of (point 6 - point 5)	If 1 excursion is available, then take that one.
		Pre-meal to 2-hour post-meal excursion daily mean	Calculated as: average of (average of [point 2 - point 1], average of [point 4 - point 3], average of [point 6 - point 5])	If there is at least 1 excursion average, take the average of all non-missing excursion averages; otherwise, set it to be missing
		Within-day coefficient of variation CV, and	Calculated as: average of CV. CV=100 × SD / MEAN. if > 3 out of 7 values are missing, then set both MEAN and SD as missing.	If only one CV is available, then take that one.
		Within-day SD	Calculated as: average of SD at each BG profile. If > 3 out of 7 values are missing, then set SD as missing.	if only one SD is available, then take that one.

Abbreviations: %CHG = Percent change from baseline; BMI = body mass index; CHG = Change from baseline; CRF = case report form; CV = cardiovascular; HbA1c = hemoglobin A1c; SMBG = self-monitored blood glucose; T2D = type 2 diabetes.

[a] M1 = Single item, missing if missing; M2 = Missing if baseline or postbaseline value is missing.

Table GZPE.4.7. Description of Secondary Endpoints for Efficacy Measures

Measure	Variable	Estimand Data Points Set / Population[a]	Analysis Method[bc]	Time Point[d]	Analysis Type
Glycemic control	CHG in HbA1c (% and mmol/mol)	TR	ANCOVA with PMI	Only Week 52	Secondary endpoint
		Efficacy	MMRM	Week 52	Secondary endpoint
	HbA1c target value of <7%	TR	Descriptive statistics with PMI	Only Week 52	Secondary endpoint
		Efficacy	Descriptive statistics with MMRM	Week 52	Secondary endpoint
	HbA1c target value of ≤6.5%)	TR	Descriptive statistics with PMI	Only Week 52	Secondary endpoint
		Efficacy	Descriptive statistics with MMRM	Week 52	Secondary endpoint
	HbA1c target value of <5.7%	TR	Descriptive statistics with PMI	Only Week 52	Secondary endpoint
		Efficacy	Descriptive statistics with MMRM	Week 52	Secondary endpoint
	CHG in fasting serum glucose (mg/dL and mmol/L)	TR	ANCOVA with PMI	Only Week 52	Secondary endpoint
		Efficacy	MMRM	Week 52	Secondary endpoint
Body weight	CHG in body weight (kg)	TR	ANCOVA with PMI	Only Week 52	Secondary endpoint
		Efficacy	MMRM	Week 52	Secondary endpoint
	%CHG in body weight (%)	TR	ANCOVA with PMI	Only Week 52	Secondary endpoint
		Efficacy	MMRM	Week 52	Secondary endpoint
	Percentage of participants with ≥ x% body weight reduction from baseline, where x = 5, 10, 15	TR	Descriptive statistics with PMI	Only Week 52	Secondary endpoint
		Efficacy	Descriptive statistics with MMRM	Week 52	Secondary endpoint
Waist circumference	CHG in waist circumference (cm)	Efficacy	MMRM	Week 52	Secondary endpoint
BMI	CHG in BMI (kg/m^2)	Efficacy	MMRM	Week 52	Secondary endpoint
SMBG	Point x SMBG (mg/dL) x = 1, 2, …,7	Efficacy	ANCOVA[c]	Week 52	Secondary endpoint
	CHG in point x SMBG (mg/dL)	Efficacy	ANCOVA[c]	Week 52	Secondary endpoint

Measure	Variable	Estimand Data Points Set / Population[a]	Analysis Method[bc]	Time Point[d]	Analysis Type
	Mean of all 7-point SMBG (mg/dL)	Efficacy	ANCOVA[c]	Week 52	Secondary endpoint
	CHG in mean of all 7-point SMBG (mg/dL)	Efficacy	ANCOVA[c]	Week 52	Secondary endpoint
	Mean of pre-meal SMBG (mg/dL)	Efficacy	ANCOVA[c]	Week 52	Secondary endpoint
	CHG in mean of pre-meal SMBG (mg/dL)	Efficacy	ANCOVA[c]	Week 52	Secondary endpoint
	Mean of 2-hour post-meal SMBG (mg/dL)	Efficacy	ANCOVA[c]	Week 52	Secondary endpoint
	CHG in post-meal SMBG (mg/dL)	Efficacy	ANCOVA[c]	Week 52	Secondary endpoint
	Pre-meal to 2-hour post-meal excursion daily mean	Efficacy	ANCOVA[c]	Week 52	Secondary endpoint
	Morning pre-meal to 2-hour post-meal excursion	Efficacy	ANCOVA[c]	Week 52	Secondary endpoint
	Midday pre-meal to 2-hour post-meal excursion	Efficacy	ANCOVA[c]	Week 52	Secondary endpoint
	Evening pre-meal to 2-hour post-meal excursion	Efficacy	ANCOVA[c]	Week 52	Secondary endpoint
	Within-day coefficient of variation CV, and	Efficacy	ANCOVA[c]	Week 52	Secondary endpoint
	Within-day SD	Efficacy	ANCOVA[c]	Week 52	Secondary endpoint

Abbreviations: %CHG = percent change from baseline; ANCOVA = analysis of covariance; BMI = body mass index; CHG = change from baseline; CV = coefficient of variation; HbA1c = hemoglobin A1c; MMRM = mixed model for repeated measures; PMI = primary multiple imputation; SD = standard deviation; SMBG = self-monitored blood glucose; TR = treatment regimen.

[a] Population is displayed if different from "Randomized Participants."

[b] PMI strategy will be performed after the log transformation.

[c] Only subjects with a nonmissing baseline value and at least one nonmissing postbaseline value for the response variable were included in the analysis.

[d] Assessments collected at multiple postbaseline visits will be analyzed at those scheduled visits using MMRM as supplemental analyses, in addition to the primary time point listed in the table.

4.5. Exploratory Endpoints Analysis

Table GZPE.4.8 includes the description and derivation of all exploratory endpoints.

Table GZPE.4.9 provides the detailed analyses including endpoint, estimand, data points set, method and imputation, population, and time point for the exploratory endpoints.

Table GZPE.4.8. Description and Derivation of Exploratory Endpoints

Measure	Description	Variable	Derivation/Comment	Handling Missing Components[a]
Lipid parameter	The tests will be performed by the central laboratory. HDL-cholesterol, triglycerides, and total cholesterol will be assayed by a Lilly designated laboratory. LDL-cholesterol, non-HDL-cholesterol, and VLDL-cholesterol will be calculated by a Lilly designated laboratory.	Total cholesterol	As provided. Log transformation before the analysis	M1
		%CHG in total cholesterol (%)	Calculated as: log(postbaseline total cholesterol) – log(baseline total cholesterol) then will transform back to percent change	M2
		HDL-cholesterol	As provided. Log transformation before the analysis	M1
		%CHG in HDL-cholesterol (%)	Calculated as: log(postbaseline HDL-cholesterol) – log(baseline HDL-cholesterol) then will transform back to percent change	M2
		LDL-cholesterol	As provided. If triglycerides are >400 mg/dL, the direct LDL will be directly measured Log transformation before the analysis.	M1
		%CHG in LDL-cholesterol (%)	Calculated as: log(postbaseline LDL-cholesterol) – log(baseline LDL-cholesterol) then will transform back to percent change	M2
		VLDL-cholesterol	As provided. Log transformation before the analysis	M1
		%CHG in VLDL-cholesterol (%)	Calculated as: log(postbaseline VLDL-cholesterol – log(baseline VLDL-cholesterol) then will transform back to percent change	M2
		Non-HDL-cholesterol	As provided. Log transformation before the analysis	M1
		%CHG in non-HDL-cholesterol (%)	Calculated as: log(postbaseline non-HDL-cholesterol) – log(baseline non-HDL-cholesterol) then will transform back to percent change	M2
		Triglycerides	As provided. Log transformation before the analysis	M1

		%CHG in triglycerides (%)	Calculated as: log(postbaseline triglyceride) – log (baseline triglyceride) then will transform back to percent change.	M2
Hepatic parameter	The tests will be performed by the central laboratory.	ALT	As provided. Log transformation before the analysis	M1
		%CHG in ALT (%)	Calculated as: log(postbaseline total cholesterol) – log(baseline total cholesterol) then will transform back to percent change	M2
		AST	As provided. Log transformation before the analysis	M1
		%CHG in AST (%)	Calculated as: log(postbaseline total cholesterol) – log(baseline total cholesterol) then will transform back to percent change	M2
		Fib-4 index	Calculated as:age (years) × AST (U/L)/(platelets [10^9/L] × ALT [U/L]1/2), with AST and ALT measured on the same day and platelets within ±30 days.	Missing if at least one of components is missing
		CHG in Fib-4	Calculated as: Postbaseline Fib-4 – baseline Fib-4	M2
		FLI	Calculated as: exp(y)/(1+exp(y)) × 100, where y=0.953 × ln (triglycerides) + 0.139 × BMI + 0.718 × ln (GGT) + 0.053 × (waist circumference) – 15.745	Missing if at least one of components is missing
		CHG in FLI	Calculated as: Postbaseline FLI – baseline FLI	M2

Measure	Description	Variable	Derivation/Comment	Handling Missing Components[a]
Blood pressure	Per Protocol GZPE Section 10.10: • Have the participant sit quietly for about 5 minutes before vital signs measurements are taken. • Take 3 measurements from the same arm, preferably the nondominant arm. • Measure the recordings at least 1 minute apart. • Blood pressure must be taken with an automated blood pressure instrument	SBP (mm Hg)	The average of 3 measures	If there is at least 1 measurement, take the average of all nonmissing measurements; otherwise, set it to be missing
		CHG in SBP (mm Hg)	Calculated as: postbaseline SBP (mm Hg) – baseline SBP (mm Hg)	M2
		DBP (mm Hg)	The average of 3 measures	If there is at least 1 measurement, take the average of all nonmissing measurements; otherwise, set it to be missing
		CHG in DBP (mm Hg)	Calculated as: postbaseline DBP (mm Hg) – baseline DBP (mm Hg)	M2
Insulin sensitivity	Fasting insulin will be performed by the central laboratory and will be assayed by a Lilly-designated laboratory.	Fasting insulin	As provided. Log transformation before the analysis	M1
		CHG in fasting insulin	Calculated as: log(postbaseline fasting insulin) – log(baseline fasting insulin)	M2
Glycemic control	Fasting C-peptide will be performed by the central laboratory and will be assayed by a Lilly-designated laboratory.	Fasting C-peptide (ng/mL)	As provided. Log transformation before the analysis	M1
		CHG fasting C-peptide (ng/mL)	Calculated as: log(postbaseline fasting C-peptide) – log(baseline fasting C-peptide)	M2
		CPI	Calculated as: Fasting C-peptide / fasting serum glucose x 100	Missing if fasting C-peptide or fasting serum glucose is missing
		CHG in CPI	Calculated as:	M2

			Postbaseline CPI – baseline CPI	
		Fasting glucagon	As provided.	M1
		CHG in fasting glucagon	Calculated as: Postbaseline fasting glucagon – baseline fasting glucagon	M2
HOMA	HOMA2 B estimates steady state beta cell function.	HOMA2 B (insulin)	Calculated from the link[b] using fasting glucose and insulin laboratory values	Missing if fasting glucose or insulin is missing
		Change from baseline in HOMA2 B (insulin)	Calculated as: postbaseline HOMA2 B (insulin) – baseline HOMA2 B (insulin)	M2
		Percent change from baseline in HOMA2 B (insulin)	Calculated as: [postbaseline HOMA2 B (insulin) / baseline HOMA2 B (insulin) – 1]×100%	M2
		HOMA2 B (C-peptide)	Calculated from the link[b] using fasting glucose and C-peptide laboratory values	Missing if fasting glucose or C-peptide is missing
		Change from baseline in HOMA2 B (C-peptide)	Calculated as: postbaseline HOMA2 B (C-peptide) – baseline HOMA2 B (C-peptide)	M2
		Percent change from baseline in HOMA2 B (C-peptide)	Calculated as: [postbaseline HOMA2 B (C-peptide) / baseline HOMA2 B (C-peptide) – 1] × 100%	M2
	HOMA2 IR estimates insulin resistance.	HOMA2 IR (insulin)	Calculated from the link[b] using fasting glucose and insulin laboratory values	Missing if fasting glucose or insulin is missing
		Change from baseline in HOMA2 IR (insulin)	Calculated as: Postbaseline HOMA2 IR (insulin) – baseline HOMA2 IR (insulin)	M2
		Percent change from baseline in HOMA2 IR (insulin)	Calculated as: [postbaseline HOMA2 IR (insulin) / baseline HOMA2 IR (insulin) – 1] × 100%	M2
		HOMA2 IR (C-peptide)	Calculated from the link[b] using fasting glucose and C-peptide laboratory values	Missing if fasting glucose or C-peptide is

				missing
		Change from baseline in HOMA2 IR (C-peptide)	Calculated as: postbaseline HOMA2 IR (C-peptide) – baseline HOMA2 IR (C-peptide)	M2
		Percent change from baseline in HOMA2 IR (C-peptide)	Calculated as: [postbaseline HOMA2 IR (C-peptide) / baseline HOMA2 IR (C-peptide) – 1] × 100%	M2
DTSQ	The DTSQ contains 8 items. The DTSQs is used to assess treatment satisfaction at baseline and the DTSQc is used to assess relative change in satisfaction from baseline at Week 52 or early termination.	DTSQs, DTSQc	As provided. perceived hyperglycemia item = item2 perceived hypoglycemia item = item 3 six-item overall satisfaction score (total score) = sum (item 1, and 4 through 8) 5-item treatment-specific satisfaction score = sum (item 1, 4, 5, 7 and 8) 2-item measure of convenience/flexibility = sum (item 4 and 5) 3-item measure of general satisfaction. = sum (item 1, 7 and 8)	M1

Measure	Description	**Variable**	**Derivation/Comment**	**Handling Missing Components** [c]
EQ-5D-5L and VAS	The EQ-5D-5L (EuroQol Research Foundation 2019) is a standardized 5-item, self-administered instrument for use as a measure of health outcome. It provides a simple descriptive profile and a single index value for health status that can be used in the clinical and economic evaluation of health care as well as population health surveys. The EQ-5D-5L assesses 5 dimensions of health: • mobility • self care • usual activities • pain/discomfort, and • anxiety/depression. Each dimension has 5 levels: no problems, slight problems, moderate problems, severe problems, and extreme problems. The VAS records the respondent's self-rated health on a vertical VAS where the endpoints are labeled as "best imaginable health state" (100) and "worst imaginable health state" (0).	EQ-5D-5L item scores	Five health profile dimensions, each dimension has 5 levels: 1 = no problems 2 = slight problems 3 = moderate problems 4 = severe problems 5 = extreme problems It should be noted that the Numerals 1-5 have no arithmetic properties and should not be used as a primary score.	Each dimension is M1. Note: Missing value can be coded as "9."
		EQ-5D-5L JP population-based utility index	EQ-5D-5L health states are converted into a single index utility score using a scoring algorithm based on public preferences. Uses the concatenation of the value of each EQ-5D-5L dimension score in the order: item 1; item 2; item 3; item 4; item 5. Derive EQ-5D-5L JP population-based index score according to the link[c] .	If any of the items is missing or coded as 9, the index score is missing.
		CHG in EQ-5D-5L utility index	Calculated as: postbaseline utility index score - baseline utility index score	M2
		EQ-5D VAS	Range from 0 = "worst imaginable health state" to 100 = "best imaginable health state" Note: Higher value indicates better health state.	M1
		CHG in EQ-5D-5L VAS	Calculated as: postbaseline VAS score – baseline VAS score	M2
EBQ	The EBQ is 55-item patient-reported questionnaire that assesses eating behavior (JASSO 2022). Each question within the questionnaire can be answered with one of the following four responses: 1. Not at all	EBQ	The total score and subtotals of the following seven major scales will be calculated following the guidance (JASSO 2022). • Recognition for weight and constitution • External eating behavior • Emotional eating behavior • Sense of hunger • Eating style	Missing, if all items are missing.

	2. Sometimes yes 3. Have a tendency; and 4. Absolutely yes.		• Food preference • Regularity of eating habits The total score and subtotals will be calculated by sex following JASSO 2022.	
		CHG in EBQ	Calculated as: postbaseline EBQ – baseline EBQ	M2
Body composition	Body composition will be measured at particular study sites. Change from baseline in body composition (body weight, body fat mass, lean body mass) to Week 52 will be evaluated. Total body water (L), protein (kg), minerals (kg), and body fat mass (kg) will be measured by InBody770 (Inbody Japan). The lean body mass (kg) will be calculated based on total body water, protein, and minerals. Body weight (kg) will be calculated based on total body water, protein, minerals, and body fat mass.	Total body mass (Body weight based on body composition)	As provided.	M1
		CHG in total body mass	Calculated as: postbaseline total body mass – baseline total body mass	M2
		Body fat mass	As provided.	M1
		CHG in body fat mass	Calculated as: postbaseline body fat mass – baseline body fat mass	M2
		Lean body mass	As provided.	M1
		CHG in lean body mass	Calculated as: postbaseline lean body mass – baseline lean body mass	M2

Abbreviations: %CHG = Percent change from baseline; ALT = alanine transaminase; AST = aspartate aminotransferase; BMI = body mass index; CHG = Change from baseline; CPI = C-peptide index; DBP = diastolic blood pressure; DTSQ = Diabetes Treatment Satisfaction Questionnaire; DTSQc = Diabetes Treatment Satisfaction Questionnaire – Change version; DTSQs = Diabetes Treatment Satisfaction Questionnaire – Status version; EBQ = Eating Behavior Questionnaire; Fib-4 = Fibrosis-4; FLI = Fatty Liver Index; GGT = gamma glutamyl transferase; HDL = high-density lipoprotein; HOMA = homeostasis model assessment; HOMA2 B = homeostasis model assessment estimates steady state beta cell function; HOMA2 IR = homeostasis model assessment estimates insulin resistance beta cell function; JP = Japan; LDL = low-density lipoprotein; SBP = systolic blood pressure; VAS = visual analog scale; VLDL = very low-density lipoprotein.

[a] M1 = Single item, missing if missing; M2 = Missing if baseline or postbaseline value is missing.

[b] Derive HOMA2 B and HOMA2 IR using the calculator at https://www.rdm.ox.ac.uk/about/our-clinical-facilities-and-mrc-units/DTU/software/homa

[c] Derive EQ-5D-5L JP population-based index score using the calculator at https://www.niph.go.jp/journal/data/64-1/201564010008.pdf

Table GZPE.4.9. Description of Exploratory Endpoints

Measure	Variable	Estimand Data Points Set / Population[a]	Analysis Method[bc]	Time Point[d]	Analysis Type
Blood pressure	CHG in SBP (mm Hg)	TR	ANCOVA with PMI	Only Week 52	Exploratory endpoint
		Efficacy	MMRM	Week 52	Exploratory endpoint
	CHG in DBP (mm Hg)	TR	ANCOVA with PMI	Only Week 52	Exploratory endpoint
		Efficacy	MMRM	Week 52	Exploratory endpoint
Markers of glucose metabolism and beta-cell function	CHG in Fasting insulin	Efficacy	MMRM	Week 52	Exploratory endpoint
	CHG in Fasting C-peptide (mg/dL)	Efficacy	MMRM	Week 52	Exploratory endpoint
	CHG in CPI	Efficacy	MMRM	Week 52	Exploratory endpoint
	CHG in Fasting glucagon	Efficacy	ANCOVA	Week 52	Exploratory endpoint
	CHG in HOMA2 B (insulin)	Efficacy	MMRM	Week 52	Exploratory endpoint
	CHG in HOMA2 IR (insulin)	Efficacy	MMRM	Week 52	Exploratory endpoint
	CHG in HOMA2 B (C-peptide)	Efficacy	MMRM	Week 52	Exploratory endpoint
	CHG in HOMA2 IR (C-peptide)	Efficacy	MMRM	Week 52	Exploratory endpoint
Patient-reported outcomes	CHG in DTSQs/DTSQc	Efficacy	Descriptive statistics	Week 52	Exploratory endpoint
	CHG in EQ-5D-5L utility index	Efficacy	ANCOVA	Week 52	Exploratory endpoint
	CHG in EQ-5D VAS	Efficacy	ANCOVA	Week 52	Exploratory endpoint
	CHG in EBQ	Efficacy	ANCOVA	Week 52	Exploratory endpoint
Body composition	CHG in body weight	Efficacy	ANCOVA	Week 52	Exploratory endpoint
	CHG in body fat mass	Efficacy	ANCOVA	Week 52	Exploratory endpoint
	CHG in lean body mass	Efficacy	ANCOVA	Week 52	Exploratory endpoint

Measure	Variable	Estimand Data Points Set / Population[a]	Analysis Method[bc]	Time Point[d]	Analysis Type
Lipid parameter	%CHG in total cholesterol (%)	TR	ANCOVA with PMI (log transformation)	Only Week 52	Exploratory endpoint
		Efficacy	MMRM (log transformation)	Week 52	Exploratory endpoint
	%CHG in HDL-cholesterol (%)	TR	ANCOVA with PMI (log transformation)	Only Week 52	Exploratory endpoint
		Efficacy	MMRM (log transformation)	Week 52	Exploratory endpoint
	%CHG in LDL-cholesterol (%)	TR	ANCOVA with PMI (log transformation)	Only Week 52	Exploratory endpoint
		Efficacy	MMRM (log transformation)	Week 52	Exploratory endpoint
	%CHG in VLDL-cholesterol (%)	TR	ANCOVA with PMI (log transformation)	Only Week 52	Exploratory endpoint
		Efficacy	MMRM (log transformation)	Week 52	Exploratory endpoint
	%CHG in non-HDL-cholesterol (%)	TR	ANCOVA with PMI (log transformation)	Only Week 52	Exploratory endpoint
		Efficacy	MMRM (log transformation)	Week 52	Exploratory endpoint
	%CHG in triglycerides (%)	TR	ANCOVA with PMI (log transformation)	Only Week 52	Exploratory endpoint
		Efficacy	MMRM (log transformation)	Week 52	Exploratory endpoint
Hepatic parameters	%CHG in ALT (%)	Efficacy	MMRM (log transformation)	Week 52	Exploratory endpoint
	%CHG in AST (%)	Efficacy	MMRM (log transformation)	Week 52	Exploratory endpoint
	CHG in Fib-4	Efficacy	MMRM (log transformation)	Week 52	Exploratory endpoint
	CHG in FLI	Efficacy	MMRM	Week 52	Exploratory endpoint
Blood pressure	CHG in SBP (mm Hg)	TR	ANCOVA with PMI	Only Week 52	Exploratory endpoint
		Efficacy	MMRM	Week 52	Exploratory endpoint
	CHG in DBP (mm Hg)	TR	ANCOVA with PMI	Only Week 52	Exploratory endpoint
		Efficacy	MMRM	Week 52	Exploratory endpoint
Markers of glucose metabolism and beta-cell function	CHG in Fasting insulin	Efficacy	MMRM	Week 52	Exploratory endpoint
	CHG in Fasting C-peptide	Efficacy	MMRM	Week 52	Exploratory endpoint
	CHG in CPI	Efficacy	MMRM	Week 52	Exploratory endpoint
	CHG in Fasting glucagon	Efficacy	ANCOVA	Week 52	Exploratory endpoint
	CHG in HOMA2 B (insulin)	Efficacy	MMRM	Week 52	Exploratory endpoint
	CHG in HOMA2 IR (insulin)	Efficacy	MMRM	Week 52	Exploratory endpoint
	CHG in HOMA2 B (C-peptide)	Efficacy	MMRM	Week 52	Exploratory endpoint

	CHG in HOMA2 IR (C-peptide)	Efficacy	MMRM	Week 52	Exploratory endpoint
Patient-reported outcomes	CHG in DTSQs/DTSQc	Efficacy	Descriptive statistics	Week 52	Exploratory endpoint
	CHG in EQ-5D-5L utility index	Efficacy	ANCOVA	Week 52	Exploratory endpoint
	CHG in EQ-5D VAS	Efficacy	ANCOVA	Week 52	Exploratory endpoint
	CHG in EBQ	Efficacy	ANCOVA	Week 52	Exploratory endpoint
Body composition	CHG in body weight	Efficacy	ANCOVA	Week 52	Exploratory endpoint
	CHG in body fat mass	Efficacy	ANCOVA	Week 52	Exploratory endpoint
	CHG in lean body mass	Efficacy	ANCOVA	Week 52	Exploratory endpoint
EBQ	CHG in total score	Efficacy	ANCOVA	Week 52	Exploratory endpoint
	CHG in sub total score	Efficacy	ANCOVA	Week 52	Exploratory endpoint

Abbreviations: %CHG = percent change from baseline; ALT = alanine transaminase; ANCOVA = analysis of covariance; AST = aspartate transferase; CHG = change from baseline; CPI = C-peptide index; DBP = diastolic blood pressure; DTSQc = Diabetes Treatment Satisfaction Questionnaire – Change version; DTSQs = Diabetes Treatment Satisfaction Questionnaire – Status version; EBQ = Eating Behavior Questionnaire; Fib-4 = Fibrosis-4; FLI = Fatty Liver Index ; HDL = high-density lipoprotein; HOMA = homeostasis model assessment; HOMA2 B = homeostasis model assessment estimates steady state beta cell function; HOMA2 IR = homeostasis model assessment estimates insulin resistance beta cell function; LDL = low-density lipoprotein; MMRM = mixed model for repeated measures; PMI = primary multiple imputation; SBP = systolic blood pressure; VAS = visual analog scale; VLDL = very low-density lipoprotein.

[a] Population is displayed if different from "Randomized Participants."

[b] PMI strategy will be performed after the log transformation.

[c] Only subjects with non-missing baseline value and at least one non-missing postbaseline value of the response variable were included in analysis.

[d] Assessments collected at multiple postbaseline visits will be analyzed at those scheduled visits using MMRM as supplemental analyses, in addition to the primary time point listed in the table.

4.6. Safety Analyses

The planned safety analyses are consistent with compound-level safety standards, which are based on various sources, including company standards, internal and external subject matter experts, publications from cross-industry initiatives (for example, PHUSE 2013, PHUSE 2015, PHUSE 2017, PHUSE 2018, PHUSE 2022), and publications from regulatory agencies (for example, FDA 2010, EMA 2014, CDER/BIRRS 2022, 2023). Descriptions of the safety analyses are provided in SAP GZPE, but some details are found in the compound-level safety standards.

Unless otherwise specified, safety analyses will be conducted using the safety participants and safety data points set (Table GZPE.4.1). It is noted that for some safety endpoints, such as study intervention discontinuation due to an AE and such, the analyses will be based on data observed during the treatment period only. The analytical approaches for safety analyses are specified in Section 4.1.2.3. Not all displays will necessarily be created as a static display. Some displays will be incorporated into interactive display tools instead of, or in addition to, a static display.

4.6.1. Extent of Exposure

Duration of exposure to study intervention will be summarized by treatment group for safety participants. Exposure will be summarized and calculated for the treatment period being considered as the date of last dose of study intervention minus the date of first dose of study intervention plus 1 day, using safety participants. Duration of study (date of end of study participation – date of randomization + 1) will also be summarized by treatment group.

Descriptive statistics (including n, mean, SD, minimum, first quartile, median, third quartile, and maximum) will be provided and the frequency of participants falling into different exposure ranges will be summarized. Overall exposure will be summarized in total participant-year (PY) of exposure, derived in the following manner:

total PY of exposure = sum of duration of exposure in days (for all participants in treatment group) / 365.25

The frequency and percentages of participants falling into the following ranges will be summarized by planned treatment group:

- >0
- ≥4 weeks
- ≥8 weeks
- ≥12 weeks
- ≥16 weeks
- ≥20 weeks
- ≥24 weeks
- ≥32 weeks
- ≥40 weeks,
- ≥48 weeks. and
- ≥52 weeks.

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

- 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 <32 weeks
- ≥32 to <40 weeks
- ≥40 to <48 weeks
- ≥48 to <52 weeks, and
- ≥52 weeks.

4.6.2. Adverse Events

Planned summary tables and listings related to the adverse events are provided in Table GZPE.4.10, and details are described in the compound-level safety standards.

Table GZPE.4.10. Listings and Summary Tables Related to Adverse Events

Analysis	Method	Population/Period
Overview of AEs, including • TEAE • SAE • death, and • permanent discontinuation from study intervention due to an AE.	Fisher's exact	Safety/TP + FP (TP for study intervention discontinuation due to an AE)
TEAEs by PT within SOC	Fisher's exact	Safety/TP + FP
Maximum Severity TEAEs		Safety/TP + FP
TEAEs with incidence ≥5% by PT	Fisher's exact	Safety/TP + FP
SAEs by PT within SOC	Fisher's exact	Safety/TP + FP
Primary AEs leading to permanent discontinuation of study intervention by PT within SOC	Fisher's exact	Safety/TP
Primary AEs leading to permanent discontinuation of study by PT within SOC	Fisher's exact	Safety/TP
AEs leading to study intervention interruption by PT within SOC	Fisher's exact	Safety/TP
AEs leading to study intervention reduction by PT within SOC	Fisher's exact	Safety/TP
Listing of SAEs		Safety/TP + FP
Listing of primary AEs leading to permanent discontinuation of study intervention		Safety/TP
Listing of primary AEs leading to permanent discontinuation of study		Safety/TP + FP
Listing of deaths		Safety/TP + FP
Listing of AEs related to suspected overdosing of orforglipron		Safety/TP
Narratives for participants with at least 1 notable event		Randomized/TP + FP

Abbreviations: AE = adverse event; eCRF = electronic case report form; FP = follow-up period; PT = preferred term; SAE = serious adverse event; SOC = system organ class; TEAE = treatment-emergent adverse event; TP = treatment period.

4.6.3. Safety Topics of Interest

This section includes safety topics of interest whether due to observed safety findings, potential findings based on drug class, or safety topics anticipated to be requested by a regulatory agency for any reason. In general, safety topics of interest will be identified by 1 or more standardized Medical Dictionary for Regulatory Activities (MedDRA) query(ies) (SMQs), system organ class (SOC), high-level term (HLT), FDA Medical Query (FMQ), or by a Lilly-defined MedDRA preferred term (PT) listing based upon the review of the most current MedDRA version, or by relevant laboratory changes. Search criteria are detailed in the compound-level safety standards.

The planned analyses for safety topics of interest are provided in Table GZPE.4.11 and are described more fully in the compound-level safety standards.

Table GZPE.4.11. Description and Analyses of Safety Topics of Interest

Safety Topic of Interest	Short Description	Analysis	Method	Population / Period
MACE	Death and nonfatal CV AEs will be adjudicated by a committee of physicians external to Lilly with cardiology expertise: a CEC.	Positively adjudicated MACE by category/subcategory and PT (if necessary).	Fisher's exact	Safety/TP + FP
		Listing of MACE reported by investigator (whether or not positively adjudicated) (if necessary).		Safety/TP + FP
Arrhythmias and cardiac conduction disorders	The TE arrhythmias and cardiac conduction disorders events will be derived using the MedDRA PTs contained in certain SMQs.	TE arrhythmias and cardiac conduction disorders by PT nested within SMQ and HLT[a].	Fisher's exact	Safety/TP + FP
Hypotension, orthostatic hypotension, and syncope	The AE database will be searched using predefined PTs.	TE hypotension, orthostatic hypotension and syncope disorders by PT[a].	Fisher's exact	Safety/TP + FP
Hypoglycemia	The 2023 American Diabetes Association position statement on glycemic targets[b] will be used to define: • Level 1 hypoglycemia • Level 2 hypoglycemia • Level 3 hypoglycemia, and (severe/serious hypoglycemia).	Incidence (percent of participants experiencing 1 or more episodes) and the rate (episodes/participant/year) of level 3 (severe/serious) hypoglycemia.		Primary analysis: Safety/TP + FP excluding events after initiation of new antihyperglycemic therapy (regardless of how many days of use). Supportive analysis: Safety/TP + FP regardless of initiation of new antihyperglycemic therapy.
		Incidence and rate of level 2 or level 3 hypoglycemia.		
		Incidence and rate of level 1 hypoglycemia (if warranted by data).		
		Listing of level 2 or level 3 hypoglycemia events.		Safety/TP + FP
Severe persistent hyperglycemia	Initiation of new antihyperglycemic therapy taken for severe persistent hyperglycemia.	Rescue therapy taken for severe persistent hyperglycemia (if warranted by data)[a].	Fisher's exact	Safety/TP + FP

Safety Topic of Interest	Short Description	Analysis	Method	Population / Period
GI adverse events	GI AEs using gastrointestinal disorders SOC will be captured.	Severe or serious TE gastrointestinal events by PT.	Fisher's exact	Safety/TP + FP
		Plots of prevalence and incidence over time for TE nausea, vomiting, diarrhea, constipation, and NVD by maximum severity.		Safety/TP + FP
		Summary of prevalence and incidence over time for TE nausea, vomiting, diarrhea, constipation and NVD.		Safety/TP + FP
		Plot of time to onset of TE nausea, vomiting, diarrhea, constipation and NVD.	Kaplan-Meier	Safety/TP + FP
		Study intervention discontinuation due to GI.	Fisher's exact	Safety/TP + FP
		Shift of minimum-to-minimum for eGFR as estimated by the CKD-EPI equation[c].		Safety/TP + FP
Renal safety	Laboratory measures related to renal safety will be analyzed. Renal events including acute renal failure and chronic renal failure exacerbation will be captured using SMQs. Dehydration events will also be identified using SMQs.	Shift of maximum-to-maximum for UACR.		Safety/TP + FP
		MMRM analyses for eGFR[c].	MMRM	Safety/TP + FP
		MMRM analyses for UACR (log-transformation).	MMRM	Safety/TP + FP
		TE renal events by PT nested within SMQs[a].	Fisher's exact	Safety/TP + FP
		TE dehydration events by PT[a].	Fisher's exact	Safety/TP + FP
		Shift of maximum-to-maximum for pancreatic enzyme.		Safety/TP + FP
Pancreatitis	The pancreatic enzyme data (p-amylase and lipase) will be observed through laboratory testing.	MMRM analysis for pancreatic enzymes: p-amylase and lipase (log-transformation).	MMRM	Safety/TP + FP
		Investigator-reported events and positively adjudicated pancreatic events, respectively	Fisher's exact	Safety/TP + FP
		TE pancreatic events by "Acute pancreatitis" SMQ and "Chronic pancreatitis" PT	Fisher's exact	Safety/TP + FP
		Listing of adjudicated and investigator-reported pancreatitis.		Safety/TP + FP

Safety Topic of Interest	Short Description	Analysis	Method	Population / Period
	All suspected cases of acute or chronic pancreatitis and AEs of severe or serious abdominal pain of unknown etiology will be sent for adjudication by an independent CEC.	TE thyroid C-cell hyperplasia and malignancies by PT.	Fisher's exact	Safety/TP + FP
Thyroid malignancies and C-cell hyperplasia	TE thyroid malignancies and C-cell hyperplasia will be identified using MedDRA HLTs and PTs. The purpose of calcitonin measurements is to assess the potential effect of orforglipron on thyroid C-cell function.	Shift of maximum-to-maximum for calcitonin value in the thresholds.		Safety/TP + FP
		MMRM analysis for calcitonin (log-transformation)	MMRM	Safety/TP + FP
		Postbaseline calcitonin value of ≥35 ng/mL that has increased at least 50% over baseline[a].	Fisher's exact	Safety/TP + FP
		TE malignancy by PT nested within SMQ[a].	Fisher's exact	Safety/TP + FP
Malignancies	Malignancy events will be derived using the MedDRA PTs contained in certain SMQs.	Abnormal postbaseline categories for hepatic safety parameters: ALT, AST, ALP, TBL, DBL, and GGT.		Safety/TP + FP
Hepatic safety	Hepatic labs include ALT, AST, serum ALP, TBL, DBL, INR, and GGT.	Hepatocellular drug-induced liver injury screening plot (TBL vs ALT or AST).		Safety/TP + FP
		Hepatocellular drug-induced liver injury screening table.		Safety/TP + FP
		Listing of participants with ALT or AST ≥3 × ULN.		Safety/TP + FP
		Cholestatic drug-induced liver injury screening plot (TBL vs ALP).		Safety/TP + FP
		Cholestatic drug-induced liver injury screening table.		Safety/TP + FP
		Listing of participants with ALP or TBL ≥2 × ULN.		Safety/TP + FP

Safety Topic of Interest	Short Description	Analysis	Method	Population / Period
	When criteria are met for hepatic evaluations, investigators will conduct close monitoring of hepatic symptoms and liver tests, perform a comprehensive evaluation for alternative causes of abnormal liver tests, and complete follow-up hepatic safety eCRFs. Potentially drug-related hepatic disorders will be identified using SMQs.	Participant profiles for participants meeting criteria for a comprehensive hepatic evaluation (as defined in Protocol GZPE).		Safety/TP + FP
		Shift of maximum-to-maximum for ALT, AST, ALP, TBL		Safety/TP + FP
		TE gallbladder and biliary tract disorders by PT nested within SMQ[a].	Fisher's exact	Safety/TP + FP
Gallbladder and biliary tract disorders	All events of TEAE biliary colic, cholecystitis, or other suspected events related to gallbladder disease will be identified by using certain SMQs.	TE allergic reactions and hypersensitivities by PT nested within applicable FMQ or SMQ[a].	Fisher's exact	Safety/TP + FP
Hypersensitivity reactions	AEs will be searched using MedDRA PTs from narrow scope FMQ or SMQs.	TE depression, suicidal ideation, and behavior events by PT nested within SMQ or FMQ[a].	Fisher's exact	Safety/TP + FP
Major depression, suicidal ideation, and behavior	AEs will be searched using MedDRA PTs that satisfies the search criteria.	TE abuse potential events by PT[a].	Fisher's exact	Safety/TP + FP
Abuse potential	AEs will be searched using a modified abuse potential FMQ.	Baseline and postbaseline diabetic retinopathy exam by treatment group and by eye (that is, any eye, left eye and right eye).		Safety/TP + FP
Diabetic retinopathy		TE diabetic retinopathy, diabetic macular edema and related complications by PTs.	Fisher's exact	Safety/TP + FP

Safety Topic of Interest	Short Description	Analysis	Method	Population / Period
	Diabetic retinopathy and macular edema will be monitored using retinal fundus photography and additional ophthalmological follow-up as warranted throughout the study for participants with T2D. Diabetic retinopathy complication will be searched using MedDRA PTs.	Shift of baseline diabetic retinopathy and macular edema status to most severe postbaseline diabetic retinopathy and macular edema status.		Safety/TP + FP

Abbreviations: AE = adverse event; ALP = alkaline phosphatase; ALT = alanine aminotransferase; AST = aspartate aminotransferase; CEC = clinical endpoint committee; CKD-EPI = Chronic Kidney Disease Epidemiology; CV = cardiovascular; DBL = direct bilirubin; eCRF = electronic case report form; eGFR = estimated glomerular filtration rate; FDA = United States Food and Drug Administration; FMQ = FDA Medical Query; FP = follow-up period; GGT = gamma-glutamyl transferase; GI = gastrointestinal; HLT = high-level term; INR = international normalized ratio; MACE = major adverse cardiovascular event; MedDRA = Medical Dictionary for Regulatory Activities; MMRM = mixed model for repeated measures; NVD = nausea, vomiting, diarrhea combined; PT = preferred term; SMQ = standardized MedDRA query; SOC = system organ class; T2D = type 2 diabetes; TBL = total bilirubin; TE = treatment-emergent; TEAE = treatment-emergent adverse event; TP = treatment period; UACR = urinary albumin to creatinine ratio.

[a] For these tables, if the number of events is less than 10, a listing will be provided instead.

[b] ElSayed et al. 2023.

[c] eGFR will be calculated by Chronic Kidney Disease-Epidemiology cystatin-c equation (Inker et al. 2012).

4.6.4. Additional Safety Assessments

Actual and change from baseline to postbaseline values of central laboratory measures, vital signs, and selected electrocardiogram (ECG) parameters will be summarized at each scheduled visit. Continuous variables, as well as the change from baseline for these variables, will be analyzed by MMRM or ANCOVA models as described in Section 4.1.2.3.

4.6.4.1. Central Laboratory Measures

All laboratory data will be reported in the International System of Units (SI). Selected laboratory measures will also be reported using conventional units.

The planned summaries for clinical laboratory evaluations are provided in Table GZPE.4.12 and are described more fully in the compound-level safety standards.

Table GZPE.4.12. Summary Tables Related to Clinical Laboratory Evaluations

Analysis	Method	Population/Period
Summary of participants with elevated or low values meeting specified levels	Descriptive statistics	Safety/TP + FP
Listing of abnormal laboratory findings		Safety/TP + FP

Abbreviations: CI = confidence interval; FP = follow-up period; SD = standard deviation; TP = treatment period.

4.6.4.2. Vital Signs

Measurements will be taken in triplicate at each visit. The average of the three measurements will be used as the analytical value for each visit. The analysis will use conducted using these analytical values.

The planned summaries for vital signs (systolic blood pressure, diastolic blood pressure, and pulse rate) are provided in Table GZPE.4.13, and are described more fully in the compound-level safety standards.

Table GZPE.4.13. Summary Tables Related to Vital Signs

Analysis	Method	Population/Period
Analysis of pulse rate and change from baseline	MMRM	Safety/TP + FP
Summary of participants meeting specific blood pressure and pulse rate levels	Descriptive statistics	Safety/TP + FP
Shift of maximum-to-maximum for pulse rate		Safety/TP + FP

Abbreviations: CI = confidence interval; FP = follow-up period; MMRM = mixed model for repeated measures; SD = standard deviation; TP = treatment period.

4.6.4.3. Electrocardiograms

Measurements taken in triplicate for a particular visit will be averaged before any analysis or summary outlined in this section. The planned summaries for ECG parameters are provided in Table GZPE.4.14 and are described more fully in the compound-level safety standards.

Table GZPE.4.14. Summary Tables Related to Electrocardiogram Parameters

Analysis	Method	Population/Period
Analysis of change from baseline in ECG parameters (heart rate, PR interval)	MMRM	Safety/TP + FP
Heart rate and PR interval in specified categories	Descriptive statistics	Safety/TP + FP

Abbreviations: CI = confidence interval; ECG = electrocardiogram; FP = follow-up period; MMRM = mixed model for repeated measures; PR = pulse rate; SD = standard deviation; TP = treatment period.

4.7. Other Analyses

4.7.1. Subgroup Analyses

All subgroup analyses will be considered as exploratory and should not be used to infer definitive treatment effects.

Efficacy subgroup analyses will be guided by the treatment regimen estimand. Subgroup analyses of mean change in HbA1c will be provided by the following baseline characteristics.

Subgroup analyses will also be conducted for each of the monotherapy or each add-on therapies (SUs, biguanides, SGLT-2is, a-GIs, TZD, or Glinides) by the following baseline characteristics（including but not limited to）if required. There are many strata for the subgroup analysis by monotherapy or each add-on therapies and the N in each stratum will be small. The efficacy subgroup analysis for the subgroup analysis by monotherapy or each add-on therapies will be descriptive statistics only, if required.

- age group (<65, ≥65 years)
- sex (female, male)
- baseline HbA1c (≤8.5%, >8.5%)
- baseline BMI (<25 or ≥25 kg/m^2)
- baseline BMI (<25, 25 to <30, ≥30 kg/m^2)
- duration of T2D (≤median, >median)
- beta-cell function such as C-peptide and CPI and so on. (≤median, >median)
- age and baseline BMI
    - <65 years and <25 kg/m^2
    - <65 years and ≥25 kg/m^2
    - ≥65 years and <25 kg/m^2, and
    - ≥65 years and ≥25 kg/m^2.

For the safety analysis, descriptive summaries will be provided as described in Section 4.6.2 and Section 1.1.

For efficacy analysis, the ANCOVA model specified in Section 4.1.2.1.1 will be fitted separately within each category of subgroup. LS means, SE, and 95% CI will be presented. The LS means and variance-covariance estimates will be used to test the treatment-by-subgroup interaction at a significance level of 0.10.

If any category within the subgroup (for example, status of yes or no) is <10% of the total population, only descriptive statistics will be provided for that category (that is, no inferential testing will be performed within the subgroup category).

For the efficacy analysis, a Bayesian shrinkage method will be used to provide adjusted estimates and inferences from subgroup analyses to potentially account for multiplicity. Following the paper by Wang et.al. (2024), for a given baseline characteristic with k subgroups, let Y_i $(i = 1, \ldots, k)$ be the observed sample estimate of the treatment effect in subgroup i.

The following hierarchical model will be used:

$$Y_i \sim N(\mu_i, \sigma_i^2),$$
$$\mu_i \sim N(\mu, \tau^2),$$
$$\mu \sim N(0, 5^2),$$
$$\tau \sim N^+(1),$$

where σ_i^2 is set to the observed variance for sample estimate, $N^+(1)$ is the half-normal distribution with a scaled parameter of 1. A weakly informative prior, as suggested by Röver et al. (2021) and specified above, will be applied to μ and τ. Of note, a standard deviation of 5% is chosen for the centrality parameter μ , so that its standard deviation is 2.5 times the unit information standard deviation (UISD). A scale parameter of 1 is chosen for the half-normal distribution such that the median heterogeneity in HbA1c reduction is approximately 0.67%, which is a reasonably large difference.

Additional subgroup evaluations including MMRM model and other efficacy and safety endpoints such as mean change in body weight, AEs and hypoglycemia may be conducted as exploratory analyses.

4.8. Interim Analyses

No interim analyses are planned for this study.

4.9. Changes to Protocol-Planned Analyses

Change from baseline to Week 52 in DTSQ was planned in protocol but there are no baseline for DTSQs data and only week 52 data for DTSQc. The descriptive statistics will be calculated for DTSQ endpoints.

5. Sample Size Determination

Since the primary objective is to assess the safety of long-term treatment of orforglipron and no formal hypothesis testing on the efficacy measures are planned, statistically derived power calculations were not considered for the determination of sample size. This sample size was decided based on the Guideline for Clinical Evaluation of Oral Hypoglycemic Agents (Pharmaceutical and Food Safety Bureau/Evaluation and Licensing Division Notification No. 0709-1. 2010). By taking account for the overall drop-out rate of approximately 10%, 39 participants with diet and exercise alone and 60 participants for combination with each OAM (SU, biguanides, SGLT-2i, a-GI, TZD, or glinides) are planned to be randomized; therefore, a total of 399 participants are planned to be randomly assigned. Up to 50% of participants who are taking metformin in combination with a-GI, TZD, or glinides will be enrolled. The sponsor or designee will inform the study sites when the upper limit is approaching.

6. Supporting Documentation

6.1. Appendix 1: Demographic and Baseline Characteristics

Demographics and other baseline characteristics will be summarized by treatment group for randomly assigned participants. The listing of basic demographic characteristics (that is, age, sex, ethnicity, race, country, baseline body weight, and so on) for randomly assigned participants will be provided.

Continuous variables will be summarized using descriptive statistics; categorical variables will be summarized using frequency counts and percentages. No inferential analysis for the comparability of baseline demographics and characteristics across treatment groups will be performed.

Table GZPE.6.1 describes the specific variables and how they will be summarized. Additional variables may also be included in the final study report.

Table GZPE.6.1. Demographics and Baseline Characteristics

Variable	Quantitative Summary	Categorical Summary
Age[a]	Yes	<65, ≥65 years <75, ≥75 years <65, ≥65 <75, ≥75 <85
Sex	No	Male, Female
Race	No	American Indian/Alaska Native, Asian, Black/African American, Native Hawaiian or other Pacific Islander, White, or Multiple
Height (cm)	Yes	
Baseline body weight (kg)	Yes	
Baseline BMI (kg/m^2)	Yes	<30, ≥30 to <35, ≥35 kg/m^2 <30, ≥30 kg/m^2 <25, ≥25 kg/m^2 <25, ≥25 to <30, ≥30 kg/m^2
Baseline waist circumference (cm)	Yes	
Baseline systolic blood pressure (mm Hg)	Yes	
Baseline diastolic blood pressure (mm Hg)	Yes	
Baseline pulse rate (beats/min)	Yes	
Baseline HbA1c (%)	Yes	≤8.0, >8.0 % ≤8.5, >8.5%
Baseline HbA1c (mmol/mol)	Yes	
Baseline fasting serum glucose (mg/dL)	Yes	
Baseline fasting serum glucose (mmol/L)	Yes	
Duration of T2D (years)	Yes	< Median, ≥ Median ≤5, >5 and ≤10, >10 years
prior use of antihyperglycemic medication and prior use of GLP-1 RAs	No	
Baseline eGFR (CKD-EPI) (mL/min/1.73m^2)	Yes	≥45 to <60, ≥60 to <90, ≥90 mL/min/1.73 m^2 <60, ≥60 mL/min/1.73 m^2
Baseline UACR (g/kg)	Yes	<30, ≥30 and ≤300, >300 g/kg
Tobacco use	No	N, Y

Abbreviations: BMI = body mass index; CKD-EPI = chronic kidney disease epidemiology; eCRF = electronic case report form; eGFR = estimated glomerular filtration rate; GLP-1 RA = glucagon-like peptide receptor agonist; HbA1c = hemoglobin A1c; T2D = type 2 diabetes; UACR = urinary albumin to creatinine ratio.

[a] Age in years will be calculated as length of the time interval from the imputed date of birth (01 July in the year of birth collected in the eCRF) to the informed consent date.

6.2. Appendix 2: Historical Illnesses and Preexisting Conditions

The count and percentages of participants with historical illnesses and preexisting conditions will be summarized by treatment group using the MedDRA preferred terms (PTs) nested within SOC. The SOCs will be listed in alphabetical order. Events will be ordered by decreasing frequency. Conditions (that is, PTs) will be ordered by decreasing frequency within SOC. This will be summarized for all randomized participants. Historical illnesses are illnesses that end prior to inform consent and preexisting conditions are conditions that are still ongoing at inform consent. No statistical comparisons between treatment groups will be performed.

Medical history will be also summarized by treatment group.

6.3. Appendix 3: Treatment Compliance

Treatment compliance is defined as taking at least 75% of scheduled study intervention during the treatment period. Compliance will be calculated by

([total number of doses dispensed – total number of doses returned]/total number of doses expected to be administered) × 100%.

If the data warrants, frequency counts and percentages of participants who are non-compliant to treatment and/or dose interruptions/modifications will be summarized by treatment group using the safety participants analysis set during the treatment period. Listings of such participants will be also provided.

6.4. Appendix 4: Prior/Concomitant Medications

Medications that start before or at the last date of the treatment period or follow-up period, and are ongoing or ended during the treatment period or follow-up period, will be classified as concomitant medication. Medications that start and end before first dose date of study intervention will be classified as prior therapy.

Baseline is defined as the corresponding medication taken on the day of the first dose of study intervention. If there are no doses of study intervention, the randomization date will be used instead of the first dose date.

The prespecified concomitant medications of interest will be summarized by treatment group at baseline. Additionally, medications of interest initiated after baseline and change to medications of interest used at baseline will be summarized during treatment period. The concomitant therapies will be mapped using the World Health Organization (WHO) Drug Dictionary in the clinical trial database and will be further classified using Anatomic-Therapeutic-Chemical (ATC) codes for reporting purposes. The concomitant medications of interest include the following groups of medication:

- prior glucagon-like peptide-1 use prior to study entry
- prior antihyperglycemic use prior to study entry
- baseline antihypertensive therapy, by type
- baseline lipid lowering therapy, by type
- utilization of the following medication during treatment period
    - antihypertensive therapy
    - lipid lowering therapy
- rescue therapy for severe persistent hyperglycemia
- initiation of additional antihyperglycemic medication during treatment period
- initiation of following medications during treatment or safety follow-up period
    - antidiarrheal medication
    - antiemetic medication, and
    - constipation medication.

Concomitant medication during the study will be summarized by ATC code by treatment group for randomized participants.

6.5. Appendix 5: Important Protocol Deviations

Important protocol deviations are defined in the Trial Issues Management Plan. A listing and a summary of important protocol deviations by treatment group will be provided at the end of study.

6.6. Appendix 6: Statistical Analysis for Addendum

This Appendix is based on the GZPE (1) Harmonized Clinical Protocol Addendum dated 27 May 2023.

6.6.1. Objectives and Endpoints

The table below shows the objectives and endpoints of the addendum, which are in addition to those outlined in Section 3 of the protocol for Study GZPE.

Objectives	Endpoints
Pharmacokinetics To evaluate the PK profile of orforglipron after multiple oral doses at Week 32 in Japanese participants with T2D	• AUC, C_{max}, and T_{max}
Exploratory To evaluate the effect of orforglipron on the PD profile and appetite after a standardized test meal at Week 32	• Graphical image of PD parameters and appetite for 4.5 hours • The change from baseline in serum glucose $AUC_{(0\text{-}4.5h)}$ • The change from baseline in C-peptide $AUC_{(0\text{-}4.5h)}$ • The change from baseline in insulin $AUC_{(0\text{-}4.5h)}$ • The change from baseline in glucagon $AUC_{(0\text{-}4.5h)}$, and • The change from baseline in appetite (fullness and hungry) in VAS.

Abbreviations: AUC = area under the concentration versus time curve;
$AUC_{(0\text{-}4.5h)}$ = area under the concentration versus time curve from time zero to 4.5 hours after dose;
C_{max} = maximum concentration; PD = pharmacodynamic; PK = pharmacokinetics; T2D = type 2 diabetes;
T_{max} = time of maximum observed concentration; VAS = visual analog scale.

6.6.2. Study Design

For the PK profile, PK samples will be collected at predose, and at 0.5, 1, 2.5, 4, 6, 8, 12, and 24 hours after administration of study intervention at Visit 10. See Section 5.2 of the Protocol GZPE addendum for the time schedule for PK and MTT.

At Visit 10, participants will come to the site in the morning, after at least 8 hours from the last administration of the study intervention. Participants may remain resident at the study site for PK sampling. When not resident at the study site, participants will return to the study site the next morning for PK sampling at the 24-hour time point in a fasting state. Study intervention will be administered after PK sampling at the 24-hour time point.

For the glucodynamic and pharmacodynamics (PD) evaluation, a meal tolerance test (MTT) will be performed at Visit 2 and Visit 10. Time points for blood sampling will be 0 (pre-meal), and 0.5, 1, 1.5, 2, 3, and 4.5 hours after the start of standardized test meal. Appetite (fullness and hungry) will also be assessed at 0 (pre-meal), 1, 2, 3, and 4.5 hours following the standardized test meal. Study intervention will be administered 30 minutes after the standardized test meal at Visit 10.

6.6.3. Sample Size Determination

A total of 36 participants will be randomized in a 1:1:1 ratio to 1 of 3 treatments. The sample size is required to evaluate safety, PK, and PD parameters, and is considered sufficient to evaluate the objectives of Study GZPE addendum.

6.6.4. Multiple Comparisons/Multiplicity

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

6.6.5. Participant Disposition

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

6.6.6. Demographic and Baseline Characteristics

All demographic and baseline clinical characteristics will be summarized by study treatment for all randomly assigned participants who participated in the Protocol GZPE addendum.

6.6.7. Pharmacokinetics

Pharmacokinetic parameter estimates for orforglipron will be calculated using standard noncompartmental methods of analysis.

The primary parameters to be estimated will be C_{max}, T_{max}, and $AUC_{(0\text{-}24h)}$ of orforglipron at Week 32. Other noncompartmental PK parameters may be reported.

PK analysis will be conducted using data from all randomized participants who received at least one dose of the investigational product and have evaluated PK data. All PK parameters will be summarized using descriptive statistics. Participants may be excluded from the PK summary statistics if a participant has an AE of vomiting that occurs at or before 2 times median T_{max}.

The average concentration profiles will be graphed using scheduled sampling times. Individual concentrations will be plotted utilizing actual sampling times.

6.6.8. Pharmacodynamics

The following measures of PD will be evaluated

- change from baseline in serum glucose (0.5, 1, 1.5, 2, 3, and 4.5 hours) at Week 32
- change from baseline in serum glucose $AUC_{(0\text{-}4.5h)}$ at Week 32
- change from baseline in C-peptide (0.5, 1, 1.5, 2, 3, and 4.5 hours) at Week 32

- change from baseline in C-peptide $AUC_{(0\text{-}4.5h)}$ at Week 32
- change from baseline in insulin (0.5, 1, 1.5, 2, 3, and 4.5 hours) at Week 32
- change from baseline in insulin $AUC_{(0\text{-}4.5h)}$ at Week 32
- change from baseline in glucagon (0.5, 1, 1.5, 2, 3, and 4.5 hours) at Week 32, and
- change from baseline in glucagon $AUC_{(0\text{-}4.5h)}$ at Week 32.

PD analyses will be conducted using data from all randomized participants who receive at least 1 dose of the investigational product and have evaluable PD parameter.

6.6.8.1. Pharmacodynamic Parameter Estimation

PD parameters will be calculated for glucose, C-peptide, insulin, and glucagon data following the standardized test meal at Week 2 and Week 32. The $AUC_{(0\text{-}4.5h)}$ will be calculated using the linear-trapezoidal method. The parameters $AUC_{(0\text{-}1h)}$, $AUC_{(0\text{-}2h)}$, $AUC_{(0\text{-}3h)}$, $AUC_{(0\text{-}4h)}$ and $AUC_{(0\text{-}5h)}$ may also be calculated.

All PD parameters will be summarized by treatment group and listed.

The average concentration profiles for glucose, C-peptide, insulin, and glucagon data will be graphed using scheduled sampling times.

6.6.8.2. Pharmacodynamic Statistical Inference

PD parameters may be transformed before statistical analyses, if deemed necessary. The change in PD parameters from baseline (Visit 2) to Week 32 (Visit 10) will be analyzed by an ANCOVA model with adjustment for baseline value and baseline HbA1c group (≤8.5%, >8.5%). The LS mean and standard error derived from the model will be displayed for change from baseline by treatment group.

6.6.9. Appetite after Standardized Test Meal

The following measures of appetite (fullness and hungry) using a VAS will be evaluated:

- change from baseline in appetite (1, 2, 3, and 4.5 hours) at Week 32, and
- change from baseline in appetite $AUC_{(0\text{-}4.5h)}$ at Week 32.

Appetite (fullness and hungry) using a VAS will be summarized by treatment and time after the standardized test meal (0 [pre-meal], 1, 2, 3, and 4.5 hours). The change in appetite parameters from baseline to Week 32 will be analyzed by an ANCOVA model with adjustment for baseline value and baseline HbA1c group (≤8.5%, >8.5%). The least squares mean and standard error derived from the model will be displayed for change from baseline by treatment group.

6.6.10. Changes to Protocol-Planned Analyses

There are no changes to the analyses described in the protocol.

6.7. Appendix 7: Integrated Analysis for Japanese participants

An integrated analysis will be conducted on the Japanese participants in orforglipron group (3 mg, 12 mg and 36 mg) until 52 weeks in Study GZGT cohort and the Japanese participants

with diet and exercise alone in the Study GZPE cohort, if required. The primary objective will be to evaluate the long-term safety of orforglipron monotherapy.

6.7.1. General considerations

The general consideration in the integrated analysis mainly follows on Program Safety Analysis Plan for Orforglipron (LY3502970) version 3 or later. If not applicable, the general consideration in the Study GZPE/GZGT statistical analysis plan may be followed.

The analysis populations used for tables, figures and listing is as below:

Population	Description
Safety Participants	All participants receiving at least 1 dose of the study intervention (include all study intervention arms)
Randomized Participants	All participants who are randomly assigned a study intervention

6.7.2. Sample Size Determination

Seventy-five Japanese participants (67 participants expected complete) with one-year's administration of orforglipron from Study GZGT and 39 Japanese participants (35 participants expected to complete) of the orforglipron monotherapy from Study GZPE will be enrolled and a total of around 100 Japanese participants will be evaluated for the integrated analysis to meet the Guideline for Clinical Evaluation of Oral Hypoglycemic Agents (Pharmaceutical and Food Safety Bureau/Evaluation and Licensing Division Notification No. 0709-1. 2010).

6.7.3. Multiple Comparisons/Multiplicity

No multiplicity adjustments will be made for evaluating objectives in the integrated analysis.

6.7.4. Endpoints

The endpoints for the integrated analysis are planned as below.

- participant disposition
- demographics and baseline characteristics
- AEs
- laboratory data, and
- vital signs.

Some efficacy outputs such as HbA1c, weight may be delivered if needed. The actual delivered outputs will be described in the list of analyses document for the integrated analysis.

7. References

Bradley C. Diabetes treatment satisfaction questionnaire. Change version for use alongside status version provides appropriate solution where ceiling effects occur. *Diabetes Care*. 1999;22(3):530-532. https://doi.org/10.2337/diacare.22.3.530

Bradley C, editor. Handbook of psychology and diabetes: a guide to psychological measurement in diabetes research and practice. Harwood Academic Publishers, 1994.

Bradley C, Lewis KS, Jennings AM, Ward JD. Scales to measure perceived control developed specifically for people with tablet-treated diabetes. *Diabet Med*. 1990;7(8):685-694. https://doi.org/10.1111/j.1464-5491.1990.tb01471.x

ElSayed NA, Aleppo G, Aroda VR, et al.; American Diabetes Association. 6. Glycemic targets: standards of care in diabetes-2023. Diabetes Care. 2023:46(suppl 1):S97-S110. https://doi.org/10.2337/dc23-S006

[FDA] United States Food and Drug Administration. Standard safety tables and figures: integrated guide. August 2022. Accessed July 31, 2023. https://downloads.regulations.gov/FDA-2022-N-1961-0046/attachment_1.pdf

Inker LA, Schmid CH, Tighiouart H, et al; CKD-EPI Investigators. Estimating glomerular filtration rate from serum creatinine and cystatin C. *N Engl J Med*. 2012;367(1):20-29. https://doi.org/10.1056/NEJMoa1114248

[JASSO] Japan Society for the Study of Obesity. Guideline for the management of obesity disease [in Japanese]. 2022. Tokyo: Life Science Publishing. 2022.

[PMDA] Pharmaceuticals and Medical Devices Agency. Guideline for clinical evaluation of oral hypoglycemic agents. Pharmaceutical and Food Safety Bureau/Evaluation and Licensing Division Notification No. 0709-1. Tokyo, Japan, 2010. www.pmda.go.jp/files/000153859.pdf. Accessed 27 March 2023.

Röver C, Bender R, Dias S, et al. On weakly informative prior distributions for the heterogeneity parameter in Bayesian random-effects meta-analysis. *Res Synth Methods*. 2021;12(4):448-474. https://doi.org/10.1002/jrsm.1475

Wang B, Du Y. Robustly leveraging post-randomization information to improve precision in randomized trials; 2022. https://doi.org/10.48550/arXiv.2110.09645

Wang Y, Tu W, Koh W, et al. Bayesian hierarchical models for subgroup analysis. *Pharm Stat*. 2024; 23(6):1065-1083. https://doi.org/10.1002/pst.2424

Ye T, Shao J, Yi Y, Zhao Q. Toward better practice of covariate adjustment in analyzing randomized clinical trials. *J Am Stat Assoc*. 2022. https://doi.org/10.1080/01621459.2022.2049278

Signature Page for VV-CLIN-124254 v3.0

Approval	PPD Statistician 10-Jun-2025 07:52:36 GMT+0000

Signature Page for VV-CLIN-124254 v3.0

Approved on 10 Jun 2025 GMT