Statistical Analysis Plan: J1I-MC-GZBD (Version 3)

A Phase 2 Study of Once-Weekly LY3437943 Compared with Placebo and Dulaglutide in Participants with Type 2 Diabetes

NCT04867785

Approval Date: 13-Dec-2022

Statistical Analysis Plan (J1I-MC-GZBD): A Phase 2 Study of Once-Weekly LY3437943 Compared with Placebo and Dulaglutide in Participants with Type 2 Diabetes

Protocol Title: A Phase 2 Study of Once-Weekly LY3437943 Compared with Placebo and Dulaglutide in Participants with Type 2 Diabetes

Protocol Number: J1I-MC-GZBD

Amendment Number: a

Compound Number: LY3437943

Short Title: Effect of LY3437943 Versus Placebo and Dulaglutide in Participants with Type 2 Diabetes

Sponsor Name: Eli Lilly and Company

Legal Registered Address: Indianapolis, Indiana USA 46285

Regulatory Agency Identifier Number(s)

Registry	ID
IND	145825

Document ID: VV-CLIN-074016

Approved on 13 Dec 2022 GMT

Table of Contents

List of Abbreviations and Terms

Term	Definition
ABPM	ambulatory blood pressure monitoring
AC	assessment committee
ADA	anti-drug antibody
AE	adverse event
AESI	adverse event of special interest
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANHECOVA	analysis of heterogeneous covariance
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
BMI	body mass index
bpm	beats per minute
CI	confidence interval
CEC	clinical endpoint committee
CN	conventional
COVID-19	Coronavirus disease 2019
CRF	case report form
CSR	clinical study report
CV	cardiovascular
DBP	diastolic blood pressure
DILI	Drug-Induced Liver Injury
DXA	Dual-Energy X-Ray Absorptiomery

EAS	Efficacy Analysis Set
ECG	electrocardiogram
eCRF	electronic case report form
eGFR	estimated glomerular filtration rate
FAS	Full Analysis Set
GI	gastrointestinal
HbA1c	hemoglobin A1c
HLT	High Level Term
ICE	intercurrent event
ICH	International Council for Harmonisation
IWRS	interactive web response system
LLT	Lowest Level Term
LS	least squares
MACE	major adverse cardiovascular events
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	mixed-effect model repeated measures
PD	pharmacodynamic
PK	pharmacokinetic
PT	Preferred Term
QTcF	Fredericia's corrected QT interval
QW	once weekly
SAE	serious adverse event
SAP	statistical analysis plan
SBP	systolic blood pressure

SC	subcutaneous
SD	standard deviation
SI	System International
SMQ	Standardised Medical Query
SOC	System Organ Class
SS	Safety Analysis Set
T2D	type 2 diabetes
TBL	total bilirubin
TE ADA	treatment-emergent anti-drug antibodies
TEAE	treatment-emergent adverse event
UACR	urine albumin-to-creatinine ratio
ULN	upper limit of normal

Version history

This SAP for Study J1I-MC-GZBD (GZBD) is based on the protocol dated 10 May 2021.

SAP Version	Approval Date	Change	Rationale
1.0	03-Dec-2021	Not Applicable	Original version
2.0	07-Apr-2022	• In Section 3 Analysis Set, added Safety Analysis Set. • Updated baseline definition in Section 4: "Unless otherwise specified, baseline is defined as the last nonmissing measurement prior to first dosing of the study drug, or prior to the randomization date for the subjects who are randomized but never dosed". • Made clear that ANHECOVA will use EAS and hypothetical strategy only to deal with ICE for efficacy parameters defined in Section 4.6.6, Section 4.7.1.1, Section 4.7.1.2 and Section 7.7 • Added metabolic acidosis and skin burning sensation and related adverse events as AESI in Section 4.6.	• Add Safety Analysis Set to better distinguish from Full Analysis Set. It is clearer than using the Full Analysis Set with different time periods. • This definition covers all the scenarios expected. • ANHECOVA was defined for hybrid estimand only. So the text was added to use ANHECOVA appropriately for efficacy estimand where the parameter has only one scheduled postbaseline timepoint. • Per Global Patient Safety recommendation.
3	See date on Page 1	• The details and rationale of the third interim	• The third interim analysis was requested by Lilly senior management to help support end of phase 2 interactions

		analysis was added to Section 4.8. • Make clear the analysis set for analysis regarding safety parameters in Section 7.7.2. • Make clear the baseline HbA1c stratum to be 8.5% or less, greater than 8.5%. • Make clear the baseline BMI stratum to be less than 30 kg/m^2, 30 kg/m^2 and greater. • Make clear the exploratory endpoint for glucose control to be percentage of participants reaching HbA1c of <5.7%, ≤6.5%, and <7.0% at 24 and/or 36 weeks. • Added subgroup analyses based on quartiles of baseline BMI in Section 4.7.2. • Add the analysis of non-HDL cholesterol to exploratory objectives in Section 1.1 and Section 4.5. • Make clear that the dose-pooling strategy would be used in Kaplan-Meier estimation in Section 4.1.	with regulatory agencies and needs to be documented in the SAP. • Make clear to use safety analysis set for analysis regarding safety parameters. • The definition of baseline HbA1c stratum was not consistent across the SAP. • The definition of baseline BMI stratum was not consistent across the SAP. • The definition was not consistent across the SAP. • Explore the effect of QW LY3437943 in population with different levels of baseline BMI. • Explore the effect of QW LY3437943 on lipids more comprehensively. • The text was added to clarify how the number of events and the total number of subjects at risk were calculated in Kaplan-Meier estimation. • Model assumption was not appropriate for the analyses.

		• Delete the full model for subgroup analyses in Section 4.7.2.	Also, the interest of the study was the intervention effect of QW LY3437943 within each subgroup instead of the difference between each subgroup.

1. Introduction

1.1. Objectives, Endpoints, and Estimands

Objectives	Endpoints
Primary	
To demonstrate superiority of QW LY3437943 (0.5, 4.0, 8.0, or 12.0 mg) in change from baseline for HbA1c relative to placebo, in participants with T2D inadequately controlled with diet and exercise with or without a stable dose of metformin	Change in HbA1c (%) from baseline to 24 weeks
Secondary	
To compare the effect of QW LY3437943 versus placebo and versus dulaglutide on:	
Glucose control	• Change in HbA1c (%) from baseline to 24 (for dulaglutide comparison only) and 36 weeks • Percentage of participants reaching HbA1c of <7.0% at 24 and/or 36 weeks • Change in FBG from baseline to 24 and 36 weeks (mg/dL, mmol/L)
Body weight control	• Change in body weight (kg) from baseline to 24 and 36 weeks
To assess safety and tolerability of study intervention	• Adverse events overall • Adverse events of special interest • Laboratory parameters • Electrocardiogram • Vital signs • Number of participants testing positive for anti-LY3437943 antibodies
To assess the PK of LY3437943 and potential participant factors that may influence its PK	• LY3437943 plasma concentrations

Objectives	Endpoints
Tertiary/Exploratory	
To assess the relationship between LY3437943 dose and/or exposure and key efficacy and safety measures and potential participant factors that may influence these relationships	• Dose-response and concentration-response analyses for key efficacy and safety parameters
To compare the effect of QW LY3437943 versus placebo and versus dulaglutide on:	
Glucose control	• Change in SMBG profile from baseline to 24 and 36 weeks • Percentage of participants reaching HbA1c of <5.7%, ≤6.5%, and <7.0% at 24 and/or 36 weeks
Body weight control	• Percentage of participants with 5%, 10%, or 15% greater body weight loss from baseline to 24 and 36 weeks
Appetite VAS	• Change from baseline to 4, 8, 12, 16, 24, and 36 weeks
Lipids	• Change from baseline to 16, 24, and 36 weeks for total cholesterol, triglycerides, HDL cholesterol, LDL cholesterol, VLDL cholesterol, and non-HDL cholesterol
Biomarkers	• Change from baseline to Week 16, 24, and 36 weeks for mechanistic biomarkers (see a detailed list of parameters in Section 4.5)

Patient-reported outcomes	• Mean change from baseline to Weeks 16, 24, and 36 for:
	  ○ SF-36v2 acute form domain scores
	  ○ Eating Inventory domain scores

Abbreviations: FBG = fasting blood glucose; HbA1c = hemoglobin A1c; HDL = high-density lipoprotein; LDL = low-density lipoprotein; PK = pharmacokinetics; QW = once-weekly; SF-36v2 = Short Form-36 version 2 Health Survey; SMBG = self-monitoring of blood glucose; T2D = type 2 diabetes; VAS = Visual Analog Scale; VLDL = very low-density lipoprotein.

Primary Estimand

The primary clinical question of interest is: What is the treatment difference in HbA1c change from baseline after 24 weeks of treatment assuming all participants who meet the inclusion criteria would have completed the treatment period without additional anti-hyperglycemic rescue medication?

The "efficacy estimand" is described by the following attributes:

- Population: participants who meet the inclusion criteria. Further details can be found in Sections 5 and 9 of protocol J1I-MC-GZBD (b).
- Endpoint: change from baseline in HbA1c at Week 24.
- Treatment condition: the randomized treatment with allowance for dose modification based on GI tolerability.
- The 2 ICEs "permanent discontinuation of study drug" and "initiation of rescue medication" are handled by the hypothetical strategy and the potential outcome of interest is the response in the efficacy measurement if participants had adhered to the randomized treatment without using additional antihyperglycemic rescue medication. There are no other defined ICEs. Dose modification will not be considered as ICEs for the definition of estimand in this study.
- Population-level summary: mean changes in HbA1c at Week 24.

Treatment effect of interest is thus defined as the difference in mean changes in HbA1c at Week 24 between QW LY3437943 and placebo.

Rationale for "efficacy estimand": This Phase 2 study aims to study the efficacy of LY3437943 under the ideal condition that all participants adhere to the randomized treatment without using additional anti-hyperglycemic rescue medication.

Estimand(s) for Secondary Objectives

The same estimand for the primary objective will be used for the following efficacy endpoints for the secondary objectives:

- change in HbA1c from baseline to 36 weeks
- change in HbA1c from baseline to 24 weeks (for dulaglutide comparison only)
- percentage of participants reaching HbA1c of less than 7.0% at 24 and/or 36 weeks
- change in fasting blood glucose (mg/dL, mmol/L) from baseline to 24 and 36 weeks, and
- change in body weight (kg) from baseline to 24 and 36 weeks.

Unless specified otherwise, safety and tolerability assessments will be guided by an estimand comparing the safety of LY3437943 doses with placebo and dulaglutide for the whole study period (the treatment period plus safety follow-up period) irrespective of adherence to study intervention for all study populations.

Supplemental Estimand(s) for Primary Efficacy Endpoint

An alternative estimand, called "hybrid estimand", is the mean treatment difference in HbA1c change from baseline at Week 24 between LY3437943 and placebo in participants who meet the inclusion criteria with ICEs handled separately according to the reasons for the events:

- Category 1: The ICEs of permanent discontinuation of study drug due to reasons unlikely related to the efficacy/safety outcomes will be handled by the hypothetical strategy.
- Category 2: The ICEs of permanent discontinuation of study drug due to lack of efficacy or use of rescue medication before study treatment discontinuation will be handled by the hypothetical strategy.
- Category 3: All other ICEs will be handled by the treatment policy strategy. With treatment policy strategy, the occurrences of "permanent discontinuation of study drug" and "initiation of rescue medication" are considered irrelevant in defining the treatment effect of interest. The value for the response variable is used regardless of whether these ICEs occur.

Population-level summary: mean change in HbA1c at Week 24.

Treatment effect of interest is thus defined as the difference in mean changes in HbA1c at Week 24 between QW LY3437943 and placebo.

Rationale for "hybrid estimand": following ICH E9 (R1) guidance on estimand, the study will collect informative treatment disposition reasons and ICEs will be handled separately according to the reasons of ICEs for this supplemental estimand. This supplemental estimand will also be applied to the efficacy endpoints for the secondary objectives listed in Estimand(s) for Secondary Objectives.

Further details can be found in Section 4.3.3.

Another alternative estimand, called "adherer estimand," is the treatment difference in HbA1c change from baseline at Week 24 between LY3437943 and placebo in participants who would adhere to both LY3437943 and placebo.

The "adherer estimand" is described by the following attributes:

- Population: a principal stratum including participants who meet the inclusion criteria and would adhere to both LY3437943 and placebo (note that this is a hypothetical population). We define that a participant adheres to the treatment if he/she has taken at least 75% of the study drug without permanent study drug discontinuation.
- Endpoint: change from baseline in HbA1c to 24 weeks.
- Treatment condition: the randomized treatment with allowance for dose modification based on GI tolerability.
- Another intercurrent event – "initiation of rescue medication" is handled by the hypothetical strategy and the potential outcome of interest is the response in the efficacy measurement if participants had adhered to both LY3437943 and placebo without using additional anti-hyperglycemic rescue medication. There are no other defined ICEs. Dose modification will not be considered as ICEs for the definition of estimand in this study.
- Population-level summary: mean changes in HbA1c at Week 24.

Treatment effect of interest is thus defined as the difference in mean changes in HbA1c at Week 24 between QW LY3437943 and placebo among the defined population.

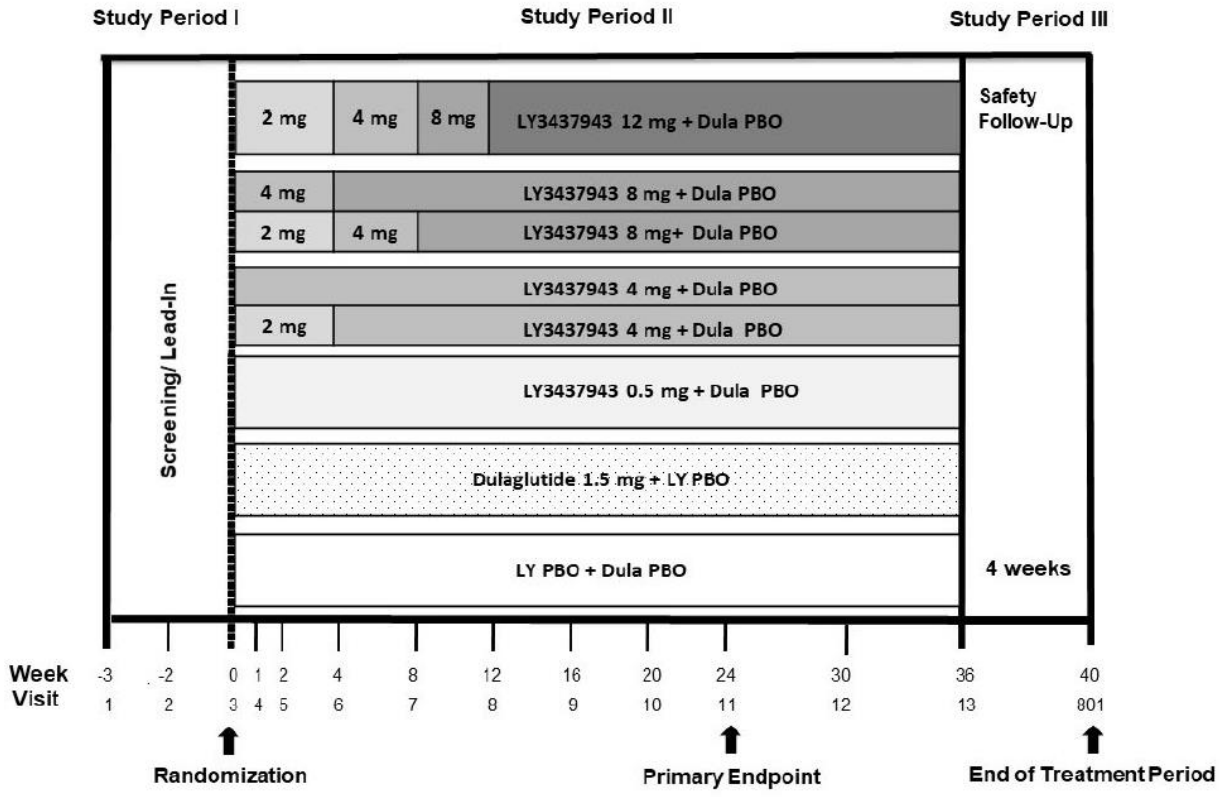
Rationale for "adhere estimand": This Phase 2 study aims to study the efficacy of LY3437943 versus placebo among those participants who could adhere to both treatments.

1.2. Study Design

Study GZBD is a Phase 2, multicenter, randomized, double-blinded, parallel, placebo- and active comparator-controlled 36-week study, with the primary outcome at 24 weeks, to investigate the safety and efficacy of LY3437943 in participants with T2D who failed to achieve adequate glycemic control on diet and exercise alone or on a stable dose of metformin (1000 mg/day or more for at least 3 months prior to screening).

Four maintenance doses of LY3437943 will be evaluated in the trial, 0.5, 4, 8, and 12 mg. Dose escalation to reduce the risk of tolerability issues will occur in certain treatment groups up to Week 12 by increasing the volume of administered study drug (or placebo). For maintenance doses 4 mg or higher, the initial dose will be 2 or 4 mg followed by additional escalation steps as appropriate. Study participants will be randomized in a 2:2:2:1:1:1:1:2 ratio to either placebo, dulaglutide 1.5 mg, LY3437943 0.5 mg, LY3437943 4 mg (with starting dose at 2 mg; slow escalation), LY3437943 4 mg (with starting dose at 4 mg; no escalation), LY3437943 8 mg (with starting dose at 2 mg; slow escalation), LY3437943 8 mg (with starting dose at 4 mg; rapid escalation), or LY3437943 12 mg (with starting dose at 2 mg; slow escalation).

All participants will undergo a 3-week screening period, a 36-week treatment period, followed by a 4-week safety follow-up period.



Abbreviations: Dula = dulaglutide; LY = LY3437943; PBO = Placebo.

Figure GZBD.1.1. Illustration of study design for clinical protocol J1I-MC-GZBD.

2. Statistical Hypothesis

The primary objective is to demonstrate that LY3437943 0.5, 4, 8, or 12 mg administered SC QW is superior to placebo with respect to HbA1c change from baseline to Week 24, in participants with T2D who failed to achieve adequate glycemic control on diet and exercise alone or on a stable dose of metformin. Thus, the null hypothesis to be tested in relation to the primary estimand is as follows:

- null hypothesis: each dose level of LY3437943 is not different from placebo in change from baseline for HbA1c at Week 24, in participants with T2D inadequately controlled with diet and exercise alone or treated with a stable dose of metformin.

2.1. Multiplicity Adjustment

No adjustment for multiplicity will be performed.

3. Analysis Sets

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

Participant Analysis Set	Description
Entered	All participants who sign the informed consent form.
Randomized	All participants who are randomly assigned to a treatment arm.
Efficacy Analysis Set (EAS)	Data obtained during the treatment period from all randomly assigned participants. Excludes participants discontinuing study drug due to inadvertent enrollment and data after permanent discontinuation of study drug or initiation of rescue medication. Participants will be included in the treatment group to which they were randomly assigned.
Full Analysis Set (FAS)	Data obtained during treatment period from all randomly assigned participants who take at least 1 dose of double-blind study treatment, excluding participants discontinuing study drug due to inadvertent enrollment, regardless of adherence to study drug or initiation of rescue medication. Participants will be analyzed according to the treatment group to which they were randomly assigned.
Safety Analysis Set (SS)	Data obtained during treatment period plus safety follow-up period from all randomly assigned participants who take at least 1 dose of double-blind study treatment, regardless of adherence to study drug or initiation of rescue medication. Participants will be analyzed according to the treatment group to which they were randomly assigned.

4. Statistical Analyses

4.1. General Considerations

Statistical analysis of this study will be the responsibility of Eli Lilly and Company (Lilly) or its designee. Some analyses and summaries described in this analysis plan may not be conducted if not warranted by data (e.g., few events to justify conducting an analysis). Additional analyses of the data may be conducted as deemed appropriate.

Unless otherwise noted, tests of treatment effects will be conducted at a 2-sided alpha level of 0.05, and the CI will be calculated at 95%, 2-sided. All tests of interactions between treatment groups and other factors will be conducted at a 2-sided alpha level of 0.10.

Unless stated otherwise, statistical summaries and analyses will be conducted based on planned randomized treatment group: placebo, dulaglutide 1.5 mg, LY 0.5 mg, LY 4.0 mg (2.0 mg), LY 4.0 mg (4.0 mg), LY 8.0 mg (2.0 mg), LY 8.0 mg (4.0 mg), and LY 12.0 mg (2.0 mg), regardless of the actual treatment(s) received by the participant due to any dose modification. The evaluation of the efficacy and safety endpoints for LY 4.0 mg and LY 8.0 mg compared with placebo or dulaglutide will be made by pooling 2 dose escalation regimens, i.e., combine LY 4.0 mg (2.0 mg) and LY 4.0 mg (4.0 mg) as LY 4.0 mg (pooled), and combine LY 8.0 mg (2.0 mg) and LY 8.0 mg (4.0 mg) as LY 8.0 mg (pooled). Therefore, the statistical comparisons between treatment groups covers the following: 1) between each of LY3437943 treatment groups: LY 0.5 mg, LY 4.0 mg (2.0 mg), LY 4.0 mg (4.0 mg), LY 4.0 mg (pooled), LY 8.0 mg (2.0 mg), LY 8.0 mg (4.0 mg), LY 8.0 mg (pooled), LY 12.0 mg (2.0 mg) and placebo, 2) between each of LY3437943 treatment groups: LY 0.5 mg, LY 4.0 mg (2.0 mg), LY 4.0 mg (4.0 mg), LY 4.0 mg (pooled), LY 8.0 mg (2.0 mg), LY 8.0 mg (4.0 mg), LY 8.0 mg (pooled), LY 12.0 mg (2.0 mg), and the dulaglutide treatment group.

The primary estimand (a precise definition of the treatment effect to be estimated) of interest in comparing the efficacy of LY3437943 doses with placebo is the “efficacy estimand” (Section 1.1). The primary efficacy assessment, guided by the “efficacy estimand” will be conducted using the EAS (Section 3). A restricted maximum likelihood-based, dose-pooling, MMRM analysis will be used to analyze continuous longitudinal variables. All the longitudinal observations at each scheduled postbaseline visit will be included in the analysis. This dose-pooling variant of MMRM replaced the fixed effect of the treatment group with that of the preplanned time-varying treatment doses at each time point in an effort to derive a more efficient estimator when the same treatment regimen over time is shared across the treatment groups due to preplanned titration (Qu et al. 2021). For example, in LY 4.0 mg (2.0 mg) and LY 8.0 mg (2.0 mg) arms, the doses administered for the first 4 weeks are both 2.0 mg. Therefore, pooling all the patients in these 2 groups provides a more efficient estimation in the effect of 2.0 mg which is also anticipated to increase the estimation precision for the later administered doses. The model for the analysis of the primary efficacy endpoint of change from baseline in HbA1c will include the fixed effects of treatment doses planned in the titration scheme – Placebo, Dulaglutide 1.5 mg, LY 0.5 mg, LY 2.0 mg (2.0 mg), LY 4.0 mg (2.0 mg), LY 4.0 mg (4.0 mg), LY 8.0 mg (2.0 mg), LY 8.0 mg (4.0 mg), and LY 12.0 mg (2.0 mg), stratification factors including baseline HbA1c stratum [8.5% or less, greater than 8.5%] and baseline BMI stratum [less than 30, 30 or greater], and continuous, fixed covariate of the baseline value, all nested

within visits. An unstructured covariance structure will be used to model the within-participant errors. Significance tests will be based on LS means and robust "Huber-White" standard errors. If this analysis fails to converge, the following covariance structures will be tested in order:

- autoregressive, and
- compound symmetry.

The first covariance structure that converges will be used.

Unless specified otherwise, safety assessments will be guided by an estimand comparing the safety of LY3437943 doses with placebo and dulaglutide irrespective of adherence to the study drug. Thus, safety analyses will be conducted using the SS.

For continuous measures, summary statistics will include sample size, mean, SD, median, minimum, and maximum for both the actual and the change from baseline measurements. LS means and standard errors derived from the analysis models will also be displayed for the change from baseline measurements. Treatment comparisons will be displayed showing the treatment difference LS means and the 95% CIs for the treatment differences, along with the p-values for the treatment comparisons. All baseline measures will be analyzed using an analysis of variance model that has treatment group as the model terms.

For categorical measures, summary statistics will include sample size, frequency, and percentages. A logistic regression model will be used to examine the treatment difference in binary efficacy outcomes with missing endpoints imputed. Fisher's exact test or Pearson's chi-square test will be used for treatment comparisons in other categorical outcomes.

For laboratory values, both CN and SI units will be presented. Therefore, both percent and mmol/mol will be presented for HbA1c and both mg/dL and mmol/L will be presented for glucose measurements. For body weight, a kilogram (kg) will be presented. Summary statistics for discrete count measures will include sample size, mean, SD, median, minimum, and maximum.

Kaplan-Meier method will be used for estimation of cumulative event-free survival rates over time, and Cox proportional hazards regression analysis will be used to compare hazards rates among treatments. In Kaplan-Meier estimation, the dose-pooling strategy will be used: subjects in the pre-specified analysis set with the same pre-planned titration history will be pooled together to calculate the number of events and the total number of subjects at risk.

Unless otherwise specified, baseline is defined as the last nonmissing measurement prior to first dosing of the study drug, or prior to the randomization date for the subjects who are randomized but never dosed.

End of study participation for a participant will be the earliest of date of death, date of withdrawal from further participation in the study, or date of safety follow-up visit (Visit 801). For participants considered to be lost-to-follow-up, end of study participation will be the date of lost-to-follow-up reported by the investigator. Participant data included in the database after the last date of study participation (including the safety follow-up period) will be excluded from statistical analysis.

Statistical treatment comparisons will only be performed between LY3437943 and placebo and between LY3437943 and dulaglutide. Since the trial is not adequately powered to detect the

difference among LY3437943 doses, comparisons among LY3437943 doses will not be performed unless otherwise specified.

Not all analyses described in this SAP will necessarily be included in the CSRs. Any analysis described in this SAP and not provided in the CSR would be available upon request.

4.2. Participant Dispositions

A listing and summary of study disposition for all randomized participants will be provided at the primary database lock and final database lock, respectively. Frequency counts and percentages of all participants screened, randomized, and receiving at least 1 dose of the study drug will be presented by treatment groups and dose escalation subgroups. A listing and summary of randomized participants not receiving study drug will be provided. All participants who discontinue the study and/or study drug will be identified, and the extent of their participation in the study will be reported. If known, a reason for their discontinuation will be given. The primary reasons for discontinuation will be listed and summarized by treatment groups and dose escalation subgroups. Kaplan-Meier plots of time to premature study discontinuation, premature study treatment discontinuation, premature study treatment discontinuation due to AE, and initiation of rescue medication will be provided based on all randomized populations. Time-to-event analyses of premature study treatment discontinuation, study drug discontinuation due to AE, and study discontinuation may be conducted.

Details about the demographic and baseline characteristics, historical illnesses and preexisting conditions, concomitant medications, treatment compliance, and important protocol deviations can be found in Appendix 1 (Section 7.1) through Appendix 5 (Section 7.5), respectively.

4.3. Primary Estimand Analysis

The primary efficacy assessment, guided by the “efficacy estimand”, will be conducted using the EAS.

For the “efficacy estimand”, the hypothetical strategy is used to handle the ICEs (permanent discontinuation of study drug or initiation of rescue medication), so only data collected before the occurrence of any such ICEs will be used in the MMRM estimation (Section 4.1). Through the MMRM, the potential efficacy measures (after the ICEs) had participants not experienced ICEs will be implicitly imputed.

To confirm the efficacy of LY3437943 with adequate statistical power, the evaluation of the efficacy and safety endpoints for LY 4.0 mg and LY 8.0 mg compared with placebo will be made by pooling 2 dose escalation regimens, i.e., combine LY 4.0 mg (2.0 mg) and LY 4.0 mg (4.0 mg) as LY 4.0 mg (pooled), and combine LY 8.0 mg (2.0 mg) and LY 8.0 mg (4.0 mg) as LY 8.0 mg (pooled).

The primary efficacy comparison will be based on the contrast between each treatment group of LY3437943 and placebo at Week 24 (Visit 11) from the MMRM analysis of change from baseline in HbA1c using the EAS (Section 4.1). The analysis model and selection of covariance structure is described in Section 4.1.

Treatment comparisons will be performed for the primary objective at the full significance level of 0.05.

4.3.1. Definition of Endpoint

The primary efficacy measure will be HbA1c (%) change from baseline (Week 0) at Week 24. The change from baseline in HbA1c for each participant at each nominal visit is defined as: postbaseline – baseline.

4.3.2. Main Analytical Approach

Change from baseline in HbA1c will be analyzed using the MMRM model for the "efficacy estimand" as described in Section 4.1.

4.3.3. Analyses for Supplemental Estimand(s)

The analysis related to a supplemental estimand, "hybrid estimand" (Section 1.1), will be conducted using data in the FAS.

Hybrid estimand is defined as the treatment difference in the mean change in HbA1c from baseline at Week 24 between LY3437943 and placebo for the study target population with ICEs handled differently according to the reasons of the events as follows:

- Category 1: The ICEs of permanent discontinuation of study drug due to reasons unlikely related to the efficacy/safety outcomes will be handled by the hypothetical strategy.
- Category 2: The ICEs of permanent discontinuation of study drug due to lack of efficacy before study treatment discontinuation will be handled by the hypothetical strategy.
- Category 3: All other ICEs will be handled by the treatment policy strategy.

In this study, following ICH E9 (R1) guidance, a plan was made to collect informative treatment disposition reasons, through eCRF, for why data intended for collection are missing and classify them into categories 1 through 3 as shown in Table GZBD.4.1 (ICH 2019).

Table GZBD.4.1. Treatment Disposition Reasons

Disposition Reason	Associated Sub-Categories	Category
Adverse event		3
Protocol deviation	Due to epidemic/pandemic	1
	Other	3
Pregnancy		3
Lack of efficacy		2
Other		3
Withdrawal by subject	Concern about study procedures/perceived risks	3
	Scheduling conflicts	1
	Subject is moving or has moved	1
	Personal issue unrelated to trial	1
	Due to epidemic/pandemic	1
	Other (option to include a specify field)	3
Physician decision	Concern about study procedures/perceived risks	3
	Scheduling conflicts	1
	Subject is moving or has moved	1
	Due to epidemic/pandemic	1
	Other (option to include a specify field)	3
Study terminated by sponsor		1
Site terminated by sponsor		1
Study terminated by IRB/ERB		1

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

To estimate the "hybrid estimand", multiple imputations will be used to impute the corresponding missing potential outcome according to the missingness patterns (Table GZBD.4.2) with ICEs handled differently according to the reasons of the events. Note that if a participant had both early treatment discontinuation and rescue therapy during the study, then missingness will be handled according to the first ICE.

When participants have missing values without ICEs, the missing values will be imputed using data from participants who do not have ICE or missing values.

Change from baseline in HbA1c will then be analyzed using ANHECOVA with terms of treatment group: placebo, dulaglutide 1.5 mg, LY 0.5 mg, LY 4.0 mg (2.0 mg), LY 4.0 mg (4.0 mg), LY 8.0 mg (2.0 mg), LY 8.0 mg (4.0 mg) and LY 12.0 mg (2.0 mg), stratification factors including baseline HbA1c stratum (8.5% or less, greater than 8.5%) and baseline BMI stratum (less than 30, 30 or greater), and continuous covariate of baseline value, along with the interaction terms between treatment group and the stratification factors and that between treatment group and the baseline value. The heteroscedasticity is also considered by introducing treatment specific variance estimates. This formulation has been shown to provide the optimal efficiency gain regardless of randomization schemes within the class of linear model adjustment (Ye et al. 2009).

The evaluation for LY 4.0 mg and LY 8.0 mg compared with placebo or dulaglutide will be made by pooling 2 dose escalation regimens, i.e., combine LY 4.0 mg (2.0 mg) and LY 4.0 mg

(4.0 mg) as LY 4.0 mg (pooled), and combine LY 8.0 mg (2.0 mg) and LY 8.0 mg (4.0 mg) as LY 8.0 mg (pooled).

Table GZBD.4.2. Strategy to Handle ICE and Missingness for Hybrid Estimand

ICE	Strategy to Handle ICE	Assumptions for Missingness	Methods to Handle Missing Values at Endpoint
Category 1: Treatment discontinuation due to reasons unlikely related to efficacy/safety outcome	Hypothetical	MAR	Data collected after the ICE will be set to missing. Missing values will be imputed using all nonmissing data (excluding data collected after ICEs) from the same treatment arm.
Category 2: Treatment discontinuation due to lack of efficacy or initiation of rescue therapy	Hypothetical	MAR	Data collected after the ICE will be set to missing. Missing values will be imputed using all nonmissing data (excluding data collected after ICEs) from the same treatment arm.
Category 3: All other treatment discontinuations	Treatment policy	MNAR Considers that these participants could not adhere to their assigned treatment and may not benefit from the assigned treatment.	Missing values will be imputed using participants in the same treatment arm with similar ICE but non-missing values (retrieved dropout imputation). In cases where there are not enough retrieved dropouts to provide a reliable imputation model, will impute the missing data using the jump-to-reference (placebo) imputation approach. For Study GZBD, the likelihood of sufficient retrieved dropouts is assumed low.

Abbreviations: ICE = intercurrent events; MAR = missing at random; MNAR = missing not at random

The analysis related to “adherer estimand” will be conducted using data in the EAS. “Adherer estimand” for the primary efficacy endpoint is the treatment difference in HbA1c change from baseline at Week 24 between LY3437943 and placebo in participants who would adhere to both LY3437943 and placebo. We define that a participant adheres to the treatment if he/she has taken at least 75% of the study drug without permanent study drug discontinuation. A multiple imputation approach proposed by Luo et al. (2021) will be used. The data used for the imputation may include the baseline covariates, intermediate outcomes/variables, adherence status, and the primary outcome of change in HbA1c from baseline at Week 24. Such data for each treatment group is used to impute the potential outcomes (including intermediate and primary outcomes, and adherence status) for participants assigned to other treatment groups. After imputation, for

each participant the potential outcomes for the primary outcome of change in HbA1c from baseline at Week 24 and adherence status under all the treatments are available. Then, the mean response for each treatment group can be estimated by simply taking the average of the potential primary outcomes for the adherers. The estimation of the treatment difference for the "adherer estimand" can be subsequently derived by taking the difference between the calculated mean responses. The inference is provided based on variance estimated by bootstrap.

4.4. Secondary Endpoints Analysis

Unless otherwise specified, secondary efficacy analyses will be conducted for EAS. The decision will be guided by the 2-sided p-values in each objective (see Table GZBD.4.3).

Table GZBD.4.3. Secondary Measures Not Controlled for Type 1 Error

Objectives	Relative to the efficacy measure	Analysis conducted in a manner similar to	Additional information
To compare the effect of QW LY3437943 versus placebo and versus dulaglutide on glucose control	Change in HbA1c (%) from baseline to 24 (for dulaglutide comparison only) and 36 weeks	MMRM model in Section 4.3.2	Same independent variables as in the MMRM model in Section 4.3.2. LSM estimates will be plotted by treatment through 36 weeks.
	Percentage of participants reaching HbA1c of <7.0% at 24 and 36 weeks	Logistic regression model with multiple imputations assuming ignorable missingness (Ma et al. 2022)	Use treatment group, stratification factors, and continuous baseline value as covariates. Missing postbaseline continuous-valued HbA1c data are imputed first within each treatment arm before deriving the binary outcome.
	Change in fasting blood glucose from baseline to 24 and 36 weeks (mg/dL, mmol/L)	MMRM model in Section 4.3.2 with change in fasting blood glucose (mg/dL, mmol/L) as the response variable	LSM estimates through 36 weeks will be plotted by study treatment.
To compare the effect of QW LY3437943 versus placebo and versus dulaglutide on body weight control	Change in body weight (kg) from baseline to 24 and 36 weeks	MMRM model in Section 4.3.2 with change in body weight (kg) as the response variable	LSM estimates through 36 weeks will be plotted by study treatment.

Abbreviations: HbA1c = hemoglobin A1c; LSM = least squares mean; MMRM = mixed model for repeated measures; QW = once-weekly.

4.4.1. Analyses for Supplemental Estimand(s)

Mean changes in HbA1c (%), body weight (kg), and fasting blood glucose (mg/dL, mmol/L) at Weeks 24 and 36 will also be analyzed by ANHECOVA guided by the hybrid estimand (similar analysis in Section 4.3.3). Missing HbA1c, body weight, and fasting blood glucose values at Weeks 24 and 36 will be imputed by the multiple imputation method described in Section 4.3.3.

Analysis guided by "adherer estimand" will also be conducted for these endpoints (similar analysis in Section 4.3.3) at Weeks 24 and 36.

4.5. Tertiary/Exploratory Endpoints Analysis

Unless otherwise specified, exploratory analyses will be conducted for EAS. Decision will be guided by the 2-sided p-values in each objective (see Table GZBD.4.4).

Table GZBD.4.4. Secondary Measures Not Controlled for Type 1 Error

Objectives	Relative to the efficacy measure	Analysis conducted
To compare the effect of QW LY3437943 versus placebo and versus dulaglutide on glucose control	Change in SMBG profile from baseline to 24 and 36 weeks	MMRM model in Section 4.3.2. LSM estimates through 36 weeks will be plotted by study treatment.
	Percentage of participants reaching HbA1c of <5.7%, ≤6.5%, and <7.0% at 24 and 36 weeks	Same logistic regression model in Table GZBD.4.3. Plot of the proportion of patients achieving the HbA1c target (<5.7%, ≤6.5%, and <7%) will also be provided.
To compare the effect of QW LY3437943 versus placebo and versus dulaglutide on body weight control	Percentage of participants with ≥5%, ≥10%, and ≥15% body weight loss from baseline to 24 and 36 weeks	Same logistic regression model in Table GZBD.4.3. Kaplan-Meier plots of time to initially achieve a body weight loss of ≥5%, ≥10%, and ≥15% will be provided.
	Change in BMI (kg/m^2) from baseline to 24 and 36 weeks	MMRM model in Section 4.3.2 with change in BMI (kg/m^2) as the response variable. LSM estimates through 36 weeks will be plotted by study treatment.
	Change in waist circumference (cm) from baseline to 24 and 36 weeks	MMRM model in Section 4.3.2 with change in waist circumference (cm) as the response variable. LSM estimates through 36 weeks will be plotted by study treatment.
To compare the effect of QW LY3437943 versus placebo and versus dulaglutide on Appetite VAS	Change in overall Appetite VAS from baseline to 4, 8, 12, 16, 24, and 36 weeks	MMRM model in Section 4.3.2.
To compare the effect of QW LY3437943 versus placebo and versus dulaglutide on lipids	Change from baseline to Weeks 4, 8, 12, 16, 24, and 36 in the following fasting lipid parameters • total cholesterol • HDL cholesterol • LDL cholesterol • VLDL cholesterol • Triglycerides • Non-HDL cholesterol	MMRM model in Section 4.3.2 including the log transformation for the response variables. LSM estimates through 36 weeks will be plotted by study treatment.
To compare the effect of QW LY3437943 versus placebo and versus dulaglutide on biomarkers	Change from baseline to Weeks 16, 24, and 36 in mechanistic biomarkers: • fasting insulin • fasting C-peptide • total adiponectin	MMRM model in Section 4.3.2 including the log transformation for the response variables. LSM estimates through 36 weeks may be plotted by study treatment.

	• IGFBP-2 • CCI • leptin • intact proinsulin • HOMA2-IR • HOMA2-B • intact proinsulin/C-peptide ratio • intact proinsulin/insulin ratio [redacted] • beta-hydroxybutyrate • free fatty acids • glycerol • uric acid • CTX-1 • P1NP • Apo B • Apo C3	
To compare the effect of QW LY3437943 versus placebo and versus dulaglutide	Change from baseline to Weeks 16, 24, and 36 in • SF-36v2 acute form domain scores • Eating Inventory domain scores	See Sections 4.7.1.1 and 4.7.1.2.

Abbreviations: Apo B = apolipoprotein B; Apo C3 = apolipoprotein C-III; BMI = body mass index; CTX-1 = C-terminal telopeptide of Type I collagen; CCI [redacted] HbA1c = hemoglobin A1c; HDL = high-density lipoprotein; HOMA2-B = homeostasis model assessment of beta-cell function; HOMA2-IR = homeostasis model assessment of insulin resistance; IGFBP-2 = insulin-like growth factor binding protein 2; LDL = low-density lipoprotein; LSM = least square means; MMRM = mixed-effect model repeated measures; P1NP = procollagen Type I N-terminal propeptide; QW = once weekly; SF-36v2 = Short Form-36 version 2 Health Survey; SMBG = self-monitoring of blood glucose; VAS = Visual Analog Scale; VLDL = very low-density lipoprotein.

4.5.1. Pharmacokinetic and Pharmacokinetic/Pharmacodynamic Methods

Pharmacokinetic, PD, and PK/PD analysis are the responsibility of Lilly's PK/PD group.

A summary of LY3437943 concentration-time data may be reported in the CSR. Data may be integrated with data from other studies for a combined PK and PK/PD analyses. Exposure-response analysis between LY3437943 concentration and safety, pharmacology, and efficacy may be performed using population PK and population PK/PD nonlinear mixed-effects modeling techniques implemented on Nonlinear Mixed Effects Modeling software. Additionally, the impact of intrinsic and extrinsic factors (such as age, weight, sex, renal, and hepatic functions) on PK and/or PD parameters may be evaluated.

4.5.2. Bayesian Analyses for Dose-Response

The longitudinal dose-response model as proposed by Qu et al. (2019) will be adapted here. This model also considers the preplanned dose changes that occur during the titration period. Let $\theta = (\theta_1, \theta_2, \dots, \theta_m)$ be the doses a participant has planned to take and $t_c = (t_{c1}, t_{c2}, \dots, t_{cm})$ be the corresponding times when the dose changes where t_{ci} indicates the time for the dose to change from θ_i to θ_{i+1}. Therefore, the mean function of the parameter of interest at time t is modelled by:

$$f_{\theta,t_c}(t) = f(t;\theta_1) + \sum_{i=1}^{m-1} [f(t - t_{ci};\theta_{i+1}) - f(t - t_{ci};\theta_i)]I(t > t_{ci}),$$

where $I(X)$ is the indicator function that takes value 1 when the condition X holds. The $f(t;\theta)$ is defined such that

$$f(t;\theta) = \frac{\lambda(\theta)(1 - e^{-k(\theta)t})}{1 - e^{-k(\theta)d}},$$

where d is the maximum duration of the treatment period in weeks ($d = 24$ for primary analyses while $d = 36$ for final analysis), $\lambda(\theta)$ is the dose-response function for the maximum response at dose θ and $k(\theta)$ is dose $\theta's$ rate parameter. This formulation of the mean function $f(t;\theta)$ was introduced by Fu and Manner (2010) to characterize the change from baseline over time in a continuous outcome that could be approximated with a pattern of exponential decay. It assumes a monotone time profile with the maximum effect reached at time d. Historical data showed that it has a good fitting to the change from baseline in body weight and HbA1c. The longitudinal data $Y_{\theta,t_c,jt}$ for participant j at time t with titration scheme (θ, t_c) will be fitted by adding the error terms to the mean function $f_{\theta,t_c}(t)$ where

$$Y_{\theta,t_c,jt} = f_{\theta,t_c}(t) + \frac{s_j(1 - e^{-k(\theta_1)t})}{1 - e^{-k(\theta_1)d}} + \epsilon_{jt},$$

$s_j \sim N(0, \sigma_s^2)$ and $\epsilon_{jt} \sim N(0, \sigma^2)$ are independent, denoting between-subject variation and within-subject variation respectively. Given s_j, $Y_{\theta,t_c,jt} \sim N(f_{\theta,t_c}(t) + \frac{s_j(1-e^{-k(\theta_1)t})}{(1-e^{-k(\theta_1)d})}, \sigma^2)$.

Such longitudinal dose-response analysis applies to the change in HbA1c from baseline. For characterizing the body weight change, we use a different formulation of $f_{\theta,t_c}(t)$ that showed better fitting in the analyses of historical trials and simulation studies while keeping all other elements unchanged. In the new formulation,

$$f_{\theta,t_c}(t) = f(t;\theta_1) + \sum_{i=1}^{m-1} h(t-t_{ci})[f(t;\theta_{i+1}) - f(t;\theta_i)]I(t > t_{ci}),$$

Where $h(t-t_{ci}) = \frac{(1-e^{-\tau(t-t_{ci})})}{1-e^{-\tau d}}$. The parameter τ controls the rate of change in the body weight when the dose changes. The dose-maximum response function $\lambda(\theta)$ is provided below:

- HbA1c

A maximum effect model is assumed where

$$\lambda(\theta) = \alpha_0 + \frac{\alpha_1\theta}{\alpha_2 + \theta},$$

and parameters $\alpha_0,\ \alpha_1,$ and α_2 represent, respectively, the basal effect when the dose level is zero (placebo), the maximum effect that can be achieved by any dose level on top of placebo, and the dose level that produces half of the maximum improvement. For dulaglutide 1.5 mg, $\lambda(\theta)$ will be modelled separately as a distinct dose.

- Body weight

A power model is assumed where

$$\lambda(\theta) = a + b * \theta^{\gamma},$$

γ is a sigmoidicity parameter indicating shape or steepness of dose response.

Other dose-response models for $\lambda(\theta)$ may be explored if the aforementioned dose-response models do not fit the data well, for example, Simple Normal Dynamic Linear Modeling.

The estimation of those parameters will be carried out in a Bayesian framework assuming noninformative priors for the hyperparameters in the model as follows:

$$\begin{cases} k(\theta), \tau \sim Uniform(0,1), \\ \alpha_0, \alpha_1, \alpha_2, a, b \sim N(0, 100^2), \\ \frac{1}{\sigma^2}, \frac{1}{\sigma_s^2} \sim Gamma(0.01, 0.01), \\ \gamma \sim N(1,5). \end{cases}$$

Posterior inference will be drawn for the dose-response at time t of clinical interest and the 95% credible intervals will also be plotted.

4.6. Safety Analyses

Unless specified otherwise, safety assessments will be guided by an estimand comparing the safety of LY3437943 doses with placebo irrespective of adherence to study drug. Thus, safety analyses will be conducted using the SS.

4.6.1. Extent of Exposure

Listing of exposure to LY3437943, placebo, and dulaglutide will be provided by the treatment group using data from the SS. Summary of the duration of follow-up (defined as time in days from date of randomization to the date of the last study visit) and/or duration on study treatment

(defined as time in days from date of the first dose of study treatment to date of the last dose of study treatment plus 7 days) will be provided by treatment group using data from SS for the following period:

- 36 weeks plus safety follow-up (Visit 801).

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

- greater than 0
- 1 weeks or longer
- 2 weeks or longer
- 4 weeks or longer
- 8 weeks or longer
- 12 weeks or longer
- 16 weeks or longer
- 20 weeks or longer
- 24 weeks or longer
- 30 weeks or longer, and
- 36 weeks or longer.

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

- 0 weeks
- longer than 0 weeks to less than 1 week
- 1 week or longer to less than 2 weeks
- 2 weeks or longer to less than 4 weeks
- 4 weeks or longer to less than 8 weeks
- 8 weeks or longer to less than 12 weeks
- 12 weeks or longer to less than 16 weeks
- 16 weeks or longer to less than 20 weeks
- 20 weeks or longer to less than 24 weeks
- 24 weeks or longer to less than 30 weeks, and
- 30 weeks or longer to less than 36 weeks.

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

4.6.2. Adverse Events

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

For events occurring on the day of first taking study medication, the CRF-collected information (e.g., treatment emergent flag, start time of study treatment, and event) 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 posttreatment.

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

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

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

4.6.3. Patient Narratives

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

- death
- SAE, or
- permanent discontinuation of study treatment due to AEs.

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

4.6.4. Vital Signs

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

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

Treatment differences in mean change will be analyzed using the MMRM model as described in Section 4.1 for the SS.

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

value less than the low limit at any time that meets the specified change criteria during the postbaseline period. To assess decreases, change from the minimum value during the baseline period to the minimum value during the postbaseline period will be used. To assess increases, changes from the maximum value during the baseline period to the maximum value during the postbaseline period will be used. Both planned and unplanned measurements will be included in the analysis. The criteria for identifying patients with treatment-emergent vital sign abnormalities are stated in Table GZBD.4.5.

Table GZBD.4.5. Categorical Criteria for Abnormal Treatment-Emergent Blood Pressure and Pulse Measurements

Parameter	Low	High
Systolic BP (mm Hg) (Supine or sitting – forearm at heart level)	≤90 and decrease from baseline ≥20	≥129 and increase from baseline ≥20
Diastolic BP (mm Hg) (Supine or sitting – forearm at heart level)	≤50 and decrease from baseline ≥10	≥90 and increase from baseline ≥10
Pulse (bpm) (Supine or sitting)	<50 and decrease from baseline ≥15	>100 and increase from baseline ≥15

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

The treatment-emergent high systolic blood pressure will also include the criterion (140 or higher and increase from baseline 20 or higher). The data summaries and analyses specified in this section will also apply to the measurements from ABPM where the same criteria applied to pulse rate will be applied to the heart rate.

In addition, following analyses will be conducted by treatment

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

4.6.5. Electrocardiograms

Summary statistics by treatment and by nominal visit will be provided for ECG parameters (heart rate, PR, QRS, RR, and QTcF). When the QRS is prolonged (e.g., a complete bundle branch block), QTcF should be used to assess ventricular repolarization. Thus, for a particular ECG, the following will be set to missing (for analysis purposes) when QRS is 120 msec or longer: QT and QTcF.

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

The criteria for identifying participants with treatment-emergent quantitative ECG abnormalities is based on Table GZBD.4.6.

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

- treatment-emergent ECG abnormalities as listed in Table GZBD.4.6
- PR 220 msec or greater with 0% through 25% and higher than 25% increase from baseline
- QTcF greater than 500 msec, and
- treatment-emergent increase from baseline in QTcF interval of greater than 30 msec, 60 msec, or 75 msec.

Treatment-emergent qualitative ECG abnormalities are defined as qualitative abnormalities that first occurred after baseline. A listing of abnormal qualitative ECGs will be created (see Table GZBD.4.6).

Table GZBD.4.6. Selected Categorical Limits for ECG Data

	Low		High	
Parameter	**Males**	**Females**	**Males**	**Females**
Heart Rate (bpm)	<50 and decrease ≥15	<50 and decrease ≥15	>100 and increase ≥15	>100 and increase ≥15
PR Interval (msec)	<120	<120	≥220	≥220
QRS Interval (msec)	<60	<60	≥120	≥120
RR Interval (msec)	<300	<300	>1714	>1714
QTcF (msec)	<330	<340	>450	>470

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

4.6.6. Clinical Laboratory Evaluation

All laboratory data will be reported in SI units. Selected laboratory measures will also be reported using CN units. Limits from the performing lab will be used to define low (L) and high (H). Descriptive summaries by treatment and by nominal visit will be provided for the baseline and postbaseline values as well as the change from baseline values.

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

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

For qualitative laboratory analytes, the number and percentage of participants with normal and abnormal values will be summarized by treatment.

A listing of abnormal findings will be created for laboratory analyte measurements, including qualitative measures. The listing will include participant identification, treatment group, laboratory collection date, study day, analyte name, and analyte finding.

The MMRM model as described in Section 4.1 or ANHECOVA (if MMRM model is not applicable) as described in Section 4.3 will be used for the analysis during the treatment period for the continuous measurements for selected lab tests. The ANHECOVA, if used in this case, will use EAS with multiple imputation performed by applying only hypothetical strategy to ICE.

4.6.7. Additional Safety Assessments

4.6.7.1. Exocrine Pancreas Safety

4.6.7.1.1. Pancreatic Enzyme

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

The counts and percentages of participants with maximum postbaseline pancreatic enzyme value exceeding the following thresholds will be provided by baseline pancreatic enzyme value (less than or equal to 1 × ULN, greater than 1 × ULN), and treatment: less than or equal to 1 × ULN, (greater than 1 to less than or equal to 3) × ULN, (greater than 3 to less than or equal to 5) × ULN, (greater than 5 to less than or equal to 10) × ULN, greater than 10× ULN. Missing will be considered as a separate group when calculating the counts and percentages.

An MMRM analysis as described in Section 4.1 will be used to analyze each pancreatic enzyme with log transformed (postbaseline measure/baseline measure) response variables using SS.

4.6.7.1.2. Pancreatic Events

Summaries of adjudicated and investigator-reported pancreatic events will be provided by treatment. Detailed searching criteria can be found in Section 7.6 (Appendix 6).

4.6.7.2. Gastrointestinal Safety

4.6.7.2.1. Nausea, Vomiting, and Diarrhea

Summaries and analyses for incidence and severity of nausea, vomiting, diarrhea, and 3 events combined, will be provided by each treatment group.

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

Time to the onset of nausea, vomiting, and diarrhea will be plotted. Prevalence and incidence of treatment-emergent nausea, vomiting, and diarrhea will also be plotted.

4.6.7.2.2. Severe Gastrointestinal Events

Severe GI AEs (GI SOC) will be captured with the AE-CRF form and serious cases will be captured with the SAE form. The PTs in the GI SOC MedDRA V23.1 will be used to identify GI AEs, and only the PTs with serious/severe cases will be considered as AESIs.

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

4.6.7.3. Hepatic Safety

4.6.7.3.1. Hepatobiliary Disorders

Hepatobiliary disorders will be considered as AESI. The counts and percentages of participants with treatment-emergent potentially drug-related hepatic disorders will be summarized by treatment using the MedDRA PTs. Detailed searching criteria can be found in Section 7.6 (Appendix 6).

4.6.7.3.2. Liver Enzymes

Analyses for laboratory analyte measurements are described in Section 4.6.6. This section describes additional analyses of liver enzymes.

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

- Alanine aminotransferase and AST
  The counts and percentages of participants whose maximum post-baseline results are less than or equal to 1× ULN, (greater than 1 to less than or equal to 3) × ULN, (greater than 3 to less than or equal to 5) × ULN, (greater than 5 to less than or equal to 10) × ULN, greater than 10× ULN or missing will be summarized by treatment group by maximum baseline result in less than or equal to 1× ULN, greater than 1× ULN or missing.
- Alkaline phosphatase, TBL, and direct bilirubin
  The counts and percentages of participants whose maximum post-baseline results are less than or equal to 1× ULN, (greater than 1 to less than or equal to 2) × ULN, greater than 2× ULN or missing will be summarized by treatment group by maximum baseline result in less than or equal to 1× ULN, (greater than 1 to less than or equal to 2) × ULN, greater than 2× ULN or missing.

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

Two plots will be provided as follows:

- Hepatocellular DILI Screening Plot (TBL versus ALT or AST): Each patient is plotted (i.e., ALT or AST): Each patient is plotted (i.e., scatterplot) based on their maximum postbaseline TBL (y-axis) versus transaminase (ALT or AST, whichever is higher), regardless of the time between the 2 maximum values. Dashed lines represent TBL and transaminase cut-offs of 2× ULN and 3× ULN (default) respectively. A potential Hy's Law case is circled and defined as having a maximum postbaseline TBL greater than or equal to 2× ULN within 30 days after maximum post-baseline ALT or AST greater than or equal to 3× ULN, without findings of cholestasis (defined as ALP less than 2× ULN). Include all scheduled and unscheduled laboratory test values.
- Cholestatic DILI Screening Plot (TBL versus ALP): each patient is plotted (i.e., scatterplot) based on their maximum postbaseline TBL (y-axis) versus ALP (x-axis), regardless of the time between the 2 maximum values. Dashed lines represent TBL and

ALP cut-offs of 2× ULN and 3× ULN (default) respectively. A potential cholestatic liver injury case is circled and defined as having a maximum postbaseline TBL greater than or equal to ≥2× ULN within 30 days after maximum postbaseline ALP greater than or equal to 2× ULN. Include all scheduled and unscheduled laboratory test values.

The counts and percentages of participants in each quadrant of the respective plots will be provided by the treatment group of LY3437943, dulaglutide, and placebo, as shown in Table GZBD.4.7 and Table GZBD.4.8, if data warranted.

Table GZBD.4.7. Summary of Participants with Potential Hepatocellular DILI in LY3437943 Group Versus Dulaglutide Versus Placebo

Quadrant	**LY3437943 (N=XXX)** **n (%)**	**Dulaglutide (N=XXX)** **n (%)**	**Placebo (N=XXX)** **n (%)**
Potential Hy's Law (Right upper)			
Cholestasis (Left upper)			
Temple's corollary (Right lower)			
Total			

Abbreviations: DILI = Drug-Induced Liver Injury; N = number of patients in the analysis population; n = number of patients in the specified category.

Table GZBD.4.8. Summary of Participants with Potential Cholestatic DILI in LY3437943 Group Versus Dulaglutide Versus Placebo

Quadrant	**LY3437943 (N=XXX)** **n (%)**	**Dulaglutide (N=XXX)** **n (%)**	**Placebo (N=XXX)** **n (%)**
TBL ≥2× ULN and ALP ≥2× ULN (Right upper)			
TBL ≥2× ULN and ALP <2× ULN (Left upper)			
TBL <2× ULN and ALP <2× ULN (Left lower)			
TBL <2× ULN and ALP ≥2× ULN (Right lower)			
Total			

Abbreviations: ALP = alkaline phosphatase; N = number of patients in the analysis population; n = number of patients in the specified category; TBL = total bilirubin; ULN = upper limit of normal

4.6.7.4. Severe Persistent Hyperglycemia

In this study, investigators will be trained on how to apply decision criteria for the timing and method of intervention in participants who do not reach glycemic targets during the 36-week treatment period (protocol Section 8.3.2.2). If any of the criteria is met, then participants may begin treatment with another antihyperglycemic agent as determined by their physician. This will be defined as severe persistent hyperglycemia for analysis. Severe persistent hyperglycemia will be considered as AESI.

Summaries of participants who had hyperglycemia rescue medicine (i.e., severe persistent hyperglycemia) will be provided by the treatment group.

4.6.7.5. Hypoglycemia

Per the study protocol, investigators should use the following definitions and criteria when diagnosing and categorizing an episode considered to be related to hypoglycemia (the blood glucose values in this section refer to values determined by a laboratory or International Federation of Clinical Chemistry and Laboratory Medicine blood-equivalent glucose meters and strips) in accordance with the 2020 American Diabetes Association position statement on glycemic targets (ADA 2020) as below. Level 2 and Level 3 hypoglycemia events are considered safety topics of special interest.

- **Level 1 Hypoglycemia (Level 1):**
  Glucose less than 70 mg/dL (3.9 mmol/L) and glucose greater than or equal to 54 mg/dL (3.0 mmol/L)
- **Level 2 Hypoglycemia (Level 2):**
  Glucose less than 54 mg/dL (3.0 mmol/L)
- **Severe Hypoglycemia (Level 3):**
  Severe hypoglycemia (in adults): A severe event characterized by altered mental and/or physical status requiring assistance for treatment of hypoglycemia. For example, participants had altered mental status and could not assist in their own care, or were semiconscious or unconscious, or experienced coma with or without seizures, and the assistance of another person was needed to actively administer carbohydrate, glucagon, or other resuscitative actions. Glucose measurements may not be available during such an event, but neurological recovery attributable to the restoration of glucose concentration to normal is considered sufficient evidence that the event was induced by a low glucose concentration.
    - The determination of a hypoglycemic event as an episode of severe hypoglycemia, as defined above, is made by the investigator based on the medical need of the participant to have required assistance and is not predicated on the report of a participant simply having received assistance.
    - If a hypoglycemic event meets the criteria of severe hypoglycemia, the investigator must record the event as serious on the AE CRF and report it to Lilly as an SAE.
- **Other hypoglycemia categories:**
  Nocturnal hypoglycemia is a hypoglycemia event (including severe hypoglycemia) that occurs at night and presumably during sleep.

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

Both the incidence (percent of patients experiencing 1 episode or more) and the rate (episodes/patient/year) of level 2 or level 3 hypoglycemia, and level 1 hypoglycemia will be reported by treatment group. For these analyses, only hypoglycemia events prior to permanent discontinuation of study drug or initiation of rescue medication will be included. A listing of all events of severe hypoglycemia may be provided, if deemed necessary. This listing will provide treatment allocation, clinical characteristics of the hypoglycemic event, and concomitant medications.

Summary of level 1 hypoglycemia will be provided for both FAS and EAS.

4.6.7.6. Immunogenicity

Treatment-emergent anti-drug antibodies are defined as those with a titer 2-fold (1 dilution) greater than the minimum required dilution if no ADAs were detected at baseline (treatment-induced ADA) or those with a 4-fold (2 dilutions) increase in titer compared to baseline if ADAs were detected at baseline (treatment-boosted ADA). A patient is evaluable for TE ADA if the patient has a nonmissing baseline ADA result, and at least 1 nonmissing postbaseline ADA result.

Listings of patients who are not TE ADA evaluable, patients with at least 1 test having detected LY3437943 ADAs, and patients having LY3437943 ADAs present or TEAE: hypersensitivity reactions or injection site reactions will be provided.

The frequency and percentage of patients with preexisting ADA, with TE ADA, with cross-reactive antibodies, and with neutralizing antibodies will be tabulated by dose (if data warrant), where proportions are relative to the number of patients who are TE ADA evaluable. The frequency and percentage of patients with hypersensitivity and injection site reaction TEAEs by TE ADA status will be tabulated if data warrant.

4.6.7.7. Hypersensitivity Events

Hypersensitivity reactions and related information reported in the eCRF will be listed and summarized by treatment.

Two main analyses are performed:

- **Potential Immediate Hypersensitivity:** Analysis of TEAEs occurring from the start of study drug administration up to 24 hours after the end of study drug administration. For events without the hypersensitivity eCRF, only date (no time) information is collected, the events that occurred on the same date as the study drug injection date will be included.
- **Potential Nonimmediate Hypersensitivity:** Analysis of TEAEs occurring more than 24 hours after the end of study drug administration, but prior to subsequent study drug administration.

Analyses of both time periods use the current standard MedDRA SMQs, published by Maintenance and Support Services Organization, to search for relevant events.

Summaries of all potential hypersensitivity reactions will be generated by PT with decreasing frequency by treatment. The AE database will be searched using predefined SMQs to identify events consistent with hypersensitivity events.

Detailed searching criteria for hypersensitivity events can be found in Section 7.6 (Appendix 6). Within query, individual PTs that satisfied the queries will be summarized. Also, a single event may satisfy multiple SMQs, in which case the event contributes to every applicable SMQ.

The number and proportion of participants experiencing treatment-emergent potential systemic hypersensitivity reactions may be summarized and compared by treatment group using Fisher's exact test.

4.6.7.8. Injection Site Reactions

Injection site reactions, incidence, and related information reported in eCRF will be summarized by treatment. Information to be summarized includes the timing of the reaction relative to study drug administration, and characteristics of the injection site reaction: erythema, induration, pain, pruritus, and edema.

Additionally, potential injection site reactions will be searched by predefined MedDRA HLTs of injection site reactions, administration site reactions, and infusion related reactions. Detailed searching criteria for injection site reaction events can be found in Section 7.6 (Appendix 6). The PT will be used for summary by treatment within each HLT category.

4.6.7.9. Renal Safety

Laboratory measures related to renal safety will be analyzed as specified for laboratory measurements in Section 4.6.6.

Two shift tables examining renal function will be created. A min-to-min shift table of eGFR estimated by the Chronic Kidney Disease Epidemiology Collaboration equation with unit mL/min/1.73m^2, using categories (less than 30, greater than or equal to 30 to less than 45, greater than or equal to 45 to less than 60, greater than or equal to 60 to less than 90, and greater than or equal to 90 mL/min/1.73m^2). A max-to-max shift table of UACR, using the categories UACR less than 30 mg/g, UACR greater than or equal to 30 and less than or equal to 300 mg/g, UACR greater than 300 mg/g (respectively, these represent normal, microalbuminuria, and macroalbuminuria).

Mixed model repeated measure analyses for eGFR and UACR will be provided as described in Section 4.1. Log transformation will be performed for UACR. SS will be used.

4.6.7.9.1. Acute Renal Events

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

Acute renal events will be considered as AESI.

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

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

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

4.6.7.9.2. Dehydration

Dehydration events will be captured in the narrow terms in Dehydration SMQ (20000232).

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

4.6.7.10. Thyroid Safety Monitoring

4.6.7.10.1. Calcitonin

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

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

4.6.7.10.2. C-Cell Hyperplasia and Thyroid Malignancies

Thyroid malignancies and C-cell hyperplasia will be considered as AESI. Treatment-emergent thyroid malignancies and C-cell hyperplasia will be identified using Malignancies SMQ (20000090), MedDRA HLT for Thyroid neoplasms and PT for Thyroid C-cell hyperplasia.

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

4.6.7.11. Major Adverse Cardiovascular Events

Major adverse CV events reported by investigators are adjudicated by an independent CEC in a blinded fashion.

The CV AEs to be adjudicated include deaths due to CV cause, myocardial infarction, hospitalization for unstable angina, hospitalization for heart failure, coronary interventions (such as coronary artery bypass graft or percutaneous coronary intervention); and cerebrovascular events, including cerebrovascular accident (stroke) and transient ischemic attack.

Only adjudicated MACE will be considered as AESI. The counts and percentages of participants with adjudicated MACE may be summarized by treatment.

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

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

4.6.7.12. Supraventricular Arrhythmias and Cardiac Conduction Disorders

Treatment-emergent supraventricular arrhythmias and cardiac conduction disorders will be considered as AESI. The CV events will include clinically relevant rhythm and conduction disorders.

The treatment-emergent supraventricular arrhythmias and cardiac conduction disorders events will be included using the MedDRA PTs. Detailed searching criteria can be found in Section 7.6 (Appendix 6).

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

4.6.7.13. Metabolic Acidosis

Treatment-emergent metabolic acidosis will be considered as AESI. The metabolic acidosis events will be included using the MedDRA PTs: Diabetic ketoacidosis, Ketoacidosis, Euglycaemic diabetic ketoacidosis, Ketonuria, Diabetic ketosis, Diabetic ketoacidotic hyperglycaemic coma, Ketosis, Urine ketone body present, Blood ketone body, Blood ketone body increased, Urine ketone body, Blood ketone body present, and lactic acidosis.

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

4.6.7.15. Drug Abuse, Dependence and Withdrawal

Subjects with drug abuse, dependence and withdrawal will be listed. Related events can be identified by searching the MedDRA PTs contained in any of the following SMQs:

- Broad and narrow terms in drug abuse and dependence SMQ (20000101) and
- Broad and narrow terms in drug withdrawal SMQ (20000102).

4.7. Other Analyses

4.7.1. Health Outcomes

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

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

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

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

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

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

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

- Descriptive summaries by treatment group, and
- Analysis from MMRM (as described in Section 4.1)/ANHECOVA model with treatment group, stratification factors as fixed effects, and the baseline value as a covariate (as described in Section 4.3), depending on whether there is more than 1 postbaseline response variables. The multiple imputation will use only hypothetical strategy to deal with ICE if ANHECOVA model is used.

4.7.1.2. Eating Inventory

The following domain scores related to eating behavior will be derived based on the 51-item Eating Inventory questionnaire:

- dietary restraint (21 items)
- disinhibition (16 items), and
- perceived hunger (14 items).

Responses to items in each domain are summed to obtain domain scores with the following ranges:

- dietary restraint (Range: 0 through 21)
- disinhibition (Range: 0 through 16), and
- perceived hunger (Range: 0 through 14).

The following analyses for the actual value and change from baseline value will be conducted on EAS:

- descriptive summaries by treatment group, and
- analysis from MMRM (as described in Section 4.1)/ANHECOVA model (as described in Section 4.3) depending on whether there is more than 1 postbaseline response variables. The multiple imputation will use only hypothetical strategy to deal with ICE if ANHECOVA model is used.

Subjects who do not speak English as their native language may have missing baseline data and will be excluded from this analysis.

4.7.2. Subgroup Analyses

Subgroup analyses of HbA1c (both % and mmol/mol) change from baseline (Week 0) at Week 24 and Week 36 will be made to assess consistency of the intervention effect across the following subgroups using the "efficacy estimand":

- age group: younger than 65 versus 65 years or older
- sex: female versus male
- baseline HbA1c (8.5% or less, greater than 8.5%)
- baseline BMI (less than 30 kg/m^2, 30 kg/m^2 or greater)
- baseline eGFR Group (less than 60, 60 $mL/min/1.73m^2$ or higher)
- race
- ethnicity
- duration of T2D (less than or equal to 5, greater than 5 and less than or equal to 10, greater than 10 years)
- duration of T2D (less than median duration, greater than or equal to median duration)
- baseline BMI (the minimum or greater to less than the first quartile, the first quartile or greater to less than the median, the median or greater to less than the third quartile, the third quartile or greater to the maximum)

If the number of participants is too small (less than [10%]) within a subgroup, then the subgroup categories may be redefined prior to unblinding the study.

For HbA1c and change from baseline in HbA1c, for each subgroup analysis aforementioned, the MMRM model as described in Section 4.1 on the subgroup will be conducted.

Subgroup analyses of body weight change from baseline (Week 0) at Week 24 and Week 36 will be conducted similarly using the following subgroups:

- sex: female versus male
- baseline HbA1c (8.5% or less, greater than 8.5%)
- baseline BMI (less than 30 kg/m^2, 30 kg/m^2 or greater)
- baseline BMI (the minimum or greater to less than the first quartile, the first quartile or greater to less than the median, the median or greater to less than the third quartile, the third quartile or greater to the maximum)

Additional subgroup analyses may also be performed.

4.8. Interim Analyses

An interim analysis for Study GZBD will be conducted when approximately 15% of the participants complete the Week 16 visit to evaluate the safety and tolerability profile of LY3437943. This safety interim analysis was planned to address potential tolerability concerns. It was planned after the approval of the protocol and thus was not documented in the protocol. A second interim analysis for Study GZBD will be conducted when approximately 60% of the participants complete the Week 16 visit. This interim analysis will evaluate both safety and efficacy profile of LY3437943 and may be used to determine the doses of future studies of LY3437943. A third interim analysis will be conducted when 100% of the participants complete the 36-week treatment period. This interim analysis will evaluate both safety and efficacy profile of LY3437943 and may be used to support end of phase 2 interactions with regulatory agencies. The second and third interim analyses were planned after the approval of the protocol and thus were not documented in the protocol.

An AC will be formed to review the interim analyses in an unblinded manner. The details regarding endpoints of analysis and number of participants will be provided in the AC charter. Since the endpoints for interim analysis is a subset of the endpoints in Section 1.1, the analysis methods of these endpoints should be consistent with the ones described in this document.

Study team members who have potential contact with the sites will remain blinded throughout the study. Information that may unblind the study during the analyses will not be reported to study sites or blinded study team members before the study has been unblinded. Study sites will receive information about interim results only if deemed necessary for the safety of the participants.

5. Sample Size Determination

Approximately 300 participants will be randomized in a 2:2:2:1:1:1:1:2 ratio to either placebo, dulaglutide 1.5 mg, LY3437943 0.5 mg, LY3437943 4 mg (with starting dose at 2 mg), LY3437943 4 mg (with starting dose at 4 mg), LY3437943 8 mg (with starting dose at 2 mg), LY3437943 8 mg (with starting dose at 4 mg), or LY3437943 12 mg (with starting dose at 2 mg). Assuming a 20% dropout rate, this results in approximately 40 completers per arm. Sample size selection is guided by the objective of establishing superiority of each LY3437943 maintenance dose to placebo relative to the change in HbA1c from baseline to 24 weeks. The evaluation of superiority to placebo will be conducted for each of the 4 LY3437943 maintenance doses at 2-sided significance level of 0.05 using 2-sample t-test. The LY3437943 group mean reduction in HbA1c at Week 24 from baseline compared to placebo is assumed to be -2.1% assuming a common SD of 1.1%. The chosen sample size provides at least 99% power to establish superiority of LY3437943 0.5 mg, LY3437943 4 mg, LY3437943 8 mg, or LY3437943 12 mg compared to placebo. No adjustment for multiplicity will be performed.

Approved on 13 Dec 2022 GMT

6. Novel Coronavirus Disease 2019 Impact

The following additional statistical analyses may be performed at the primary database lock and final database lock to assess the impact of COVID-19 pandemic for all randomized participants if the data warrants:

- listing of all randomized participants who discontinue study due to COVID-19 pandemic,
- listing of all study disruptions related to COVID-19 pandemic,
- listing of AEs or deaths related to COVID-19 pandemic, and
- listing of important protocol deviations due to COVID-19 pandemic.

In case there is a larger impact of COVID-19 on the study, due to a shut-down or any other reason, more details for additional analyses may be provided.

For the primary endpoints and key secondary endpoints, missing data due to COVID-19 will be handled as described in Section 4.3.3.

7. Supporting Documentation

7.1. Appendix 1: Demographic and Baseline Characteristics

A listing of participant demographics for all randomized participants will be provided. All demographic and baseline clinical characteristics will be summarized by treatment groups and dose escalation subgroups for all randomized participants.

The following variables will be included, but not limited to, age (years), age groups (younger than 65 and 65 years or older, and younger than 65, 65 or older to less than 75, 75 or older to younger than 85, 85 years or older), sex, ethnicity, race, height (cm), body weight, HbA1c at baseline, HbA1c stratum at baseline (8.5% or less, greater than 8.5%), fasting glucose at baseline, BMI at baseline, BMI groups at baseline (less than 30, 30 kg/m^2 or greater; less than 25, greater than or equal to 25 to less than 30, greater than or equal to 30 to less than 35, greater than or equal to 35 to less than 40, and greater than or equal to 40 kg/m^2), waist circumference at baseline (cm), metformin use (yes, no), renal function (normal, mild, moderate, severe and end stage), eGFR groups (eGFR rate based on the modified Modification of Diet in Renal Disease equation: greater than or equal to 90, less than 90 and greater than or equal to 60, less than 60 and greater than or equal to 30, and less than 30 mL/min/1.73 m^2), duration of diabetes (years), tobacco baseline use status, baseline SBP, and baseline DBP. A listing of participant demographic and baseline characteristics at baseline will be provided for all randomized participants.

A listing of participants whose stratification factor value(s) that are entered into the IWRS (for treatment group assignment) is different from the clinical database will also be provided.

7.2. Appendix 2: Historical Illnesses and Preexisting Conditions

The count and percentages of participants with historical illnesses and preexisting conditions will be summarized by treatment groups and dose escalation subgroups using the MedDRA PTs nestled within SOC. The SOC will be in alphabetical order. Conditions (i.e., 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 informed consent and preexisting conditions are conditions that are still ongoing at informed consent. Events will be ordered by decreasing frequency. No statistical comparisons between treatment groups will be performed.

7.3. Appendix 3: Concomitant Medications

Concomitant medications will be summarized by treatment group. The percentages of participants who took concomitant medication will be summarized by treatment using PTs nestled within ATC Level 3 codes. The concomitant medications will be ordered by decreasing frequency within each ATC level.

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

Additionally, medications of interest (as defined below) will be summarized by treatment groups and dose escalation subgroups for the SS.

Concomitant medications of interest include the following:

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

7.4. Appendix 4: Treatment Compliance

Listing and summary of prematurely discontinuing study treatment (including discontinuation reason) and discontinuing study will be provided by treatment groups and dose escalation subgroups. Kaplan-Meier plots of time to premature study treatment discontinuation and time to premature study treatment discontinuation due to AE will be provided at Weeks 24 and 36 based on all randomized population. Kaplan-Meier plots of time to premature study discontinuation will be provided at Weeks 24 and 36 based on all randomized population.

If data warrants, the counts and percentages of participants who follow the planned dose escalation scheme (IWRS data), have dose temporary discontinuation of study drug (eCRF data), or have dose de-escalation (IWRS data) will be summarized for each treatment group. Listings of participants who had missing dose or overdose during the study as reported through eCRF will be provided.

Per protocol, treatment compliance will be assessed every 4 weeks, or every 6 weeks after Visit 11, at the time of visits to the study site. Treatment compliance will be defined as taking at least 75% of the scheduled SC doses. Compliance over the entire study period will be calculated using the number of doses administered (regardless of the actual dose in unit or mL administered) divided by the total number of doses expected to be administered ×100 during the study period. Treatment compliance will be summarized for the entire study period by treatment groups and dose escalation subgroups using the SS.

7.5. Appendix 5: Important Protocol Deviations

Important protocol deviations are identified in the Trial Issues Management Plan. A listing and summary of important protocol deviations by treatment groups and dose escalation subgroups will be provided at the end of the study (for all randomized participants).

7.6. Appendix 6: Searching Criteria for Additional Safety Assessments

Pancreatitis Events

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

Treatment-Emergent Hepatobiliary Disorders

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

- broad and narrow terms in the Liver related investigations, signs, and symptoms SMQ (20000008)
- broad and narrow terms in the Cholestasis and jaundice of hepatic origin SMQ (20000009)
- broad and narrow terms in the Hepatitis non-infections SMQ (20000010)
- broad and narrow terms in the Hepatic failure, fibrosis and cirrhosis and other liver damage SMQ (20000013)
- narrow terms in the Liver-related coagulation and bleeding disturbances SMQ (20000015)
- narrow PTs in Gallbladder related disorders SMQ (20000124)
- narrow PTs in Biliary tract disorders SMQ (20000125); and
- narrow PTs in Gallstone related disorders SMQ (20000127).

Injection Site Reactions

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

- MedDRA HLT of injection site reaction
- HLT of administration site reactions NEC, and
- HLT of infusion Site Reactions

Supraventricular Arrhythmias and Cardiac Conduction Disorders

Treatment-emergent supraventricular arrhythmias, arrhythmias and cardiac conduction disorders will be considered as an AESI. The CV events will include clinically relevant rhythm and conduction disorders. The treatment-emergent supraventricular arrhythmias and cardiac conduction disorders events will be included using the MedDRA PT contained in any of the following SMQs:

- Supraventricular arrhythmias:
- For symptoms: Arrhythmia related investigations, signs, and symptoms SMQ (20000051), narrow and broad terms
    - For supraventricular arrhythmias:
        - Supraventricular tachyarrhythmia SMQ (20000057), broad and narrow terms
        - Tachyarrhythmia terms, nonspecific SMQ (20000164), narrow terms only; and
        - Ventricular tachyarrhythmia SMQ (20000058), narrow terms only.
- Cardiac conduction disorders
    - Conduction defects SMQ (20000056), narrow terms only; and
    - Cardiac conduction disorders HLT (10000032), all PTs.

Hypersensitivity Events

The hypersensitivity TEAE are characterized as follows:

- Anaphylactic reaction SMQ (20000021; narrow, algorithm per SMQ guide, and broad)
- Hypersensitivity SMQ (20000214; narrow and broad)
- Angioedema SMQ (20000024; narrow and broad), and
- Event maps to PT of Injection related reaction (10071152)

The number and percentage of patients who experienced a TEAE for the following will be analyzed for each of the 2 time periods:

- any narrow or algorithmic term from any 1 of the 4 SMQs indicated above (that is, combined search across narrow and algorithmic portions of all 4 SMQs)
- any narrow scope term within each SMQ, separately (that is, narrow SMQ search), and
- any term within each SMQ, separately (that is, broad SMQ search)

7.7. Appendix 7: Dual-Energy X-Ray Absorptiometry and Ambulatory Blood Pressure Monitoring

This section is applicable to the participants who are enrolled in the DXA and ABPM addendum.

Baseline clinical characteristics and demographic variables (including but not limited to total fat mass, total lean mass, visceral fat mass, SBP and DBP, heart rate, and other variables described in Section 7.1 [Appendix 1]) will be summarized by treatment for the participants who are enrolled in the addendum.

7.7.1. Primary Efficacy Analyses

The primary objective of DXA and ABPM addendum is to assess the effect of 0.5, 4, 8, and 12 mg doses of LY3437943 compared with placebo on percent change in total fat mass measured by DXA from baseline (Visit 2 at Week -2) to Week 36.

An ANHECOVA model as described in Section 4.3 will be used for EAS to analyze the primary endpoint using the baseline total fat mass as a covariate. The multiple imputation will use only hypothetical strategy to deal with ICEs. The LS mean of the treatment difference estimated from the model and 2-sided 95% CIs will be provided.

7.7.2. Secondary Efficacy Analyses

The secondary study objectives related to ABPM are to assess the effect of 0.5, 4, 8, and 12 mg doses of LY3437943 compared with placebo on change in SBP and DBP measured by ABPM from baseline (Visit 2 at Week -2) to Week 36.

These objectives will be analyzed using a mixed model for repeated measures (MMRM) as described in Section 4.1 using SS. Baseline is defined as the averaged measurement value of the data collected during the 24-hour monitoring period prior to first dosing of study drug. At the primary analysis at week 24 when there is only one post-baseline measurement of ABPM parameters, the ANHECOVA as described in Section 7.7.1 will be used, and multiple imputation will be done by treatment group with all available data in SS.

7.7.3. Exploratory Efficacy Analyses

The following exploratory objectives are considered in DXA and ABPM addendum:

To assess the effect of 0.5, 4, 8, and 12 mg doses of LY3437943 versus placebo and dulaglutide on the following DXA-derived measures:

- percent change in total fat mass from baseline (Visit 2 at Week -2) to Week 36 (dulaglutide comparison only)
- percent change in body weight from baseline (Visit 2 at Week -2) to Week 36
- percent change in total lean mass from baseline (Visit 2 at Week -2) to Week 36
- change in ratio between total fat mass and total lean mass from baseline (Visit 2 at Week -2) to Week 36, and
- percent change in visceral fat mass from baseline (Visit 2 at Week -2) to Week 36.

To assess the effect of 0.5, 4, 8, and 12 mg doses of LY3437943 versus placebo and dulaglutide on the following ABPM-derived measures:

- mean change in SBP from baseline (Visit 2 at Week -2) to Weeks 24 and 36 (dulaglutide comparison only)
- mean change in DBP from baseline (Visit 2 at Week -2) to Weeks 24 and 36 (dulaglutide comparison only), and
- mean change in heart rate (bpm) from baseline (Visit 2 at Week -2) to Weeks 24 and 36.

For DXA-derived endpoints, the same model as described in Section 7.7.1 will be used. For ABPM-derived endpoints, the same model as described in Section 7.7.2 will be used.

8. References

[ADA] American Diabetes Association. 6. Glycemic targets: standards of medical care in diabetes – 2020. *Diabetes Care*. 2020;43(suppl 1):S66-S76. https://doi.org/10.2337/dc20-s006

Fu H, Manner D. Bayesian adaptive dose-finding studies with delayed responses. *J Biopharm Stat*. 2010;20(5):1055-1070. https://doi.org/10.1080/10543400903315740

[ICH] International Council for Harmonisation. Addendum on estimands and sensitivity analysis in clinical trials. To the guideline on statistical principles for clinical trials. E9(R1). November 20, 2019. Accessed April 07, 20202. https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf.

Luo J, Ruberg SJ, Qu Y. Estimating the treatment effect for adherers using multiple imputation. *Pharm Stat*. 2021; In press. https://doi.org/10.1002/pst.2184

Ma C, Shen X, Qu Y, Du Y. Analysis of an incomplete binary outcome dichotomized from an underlying continuous variable in clinical trials. *Pharm Stat*. 2002; In press.

Qu Y, Cui X, Bray R, Sun N. Efficient estimation of the efficacy and safety endpoints for clinical trials with preplanned dose titrations. Working paper; 2021.

Qu Y, Liu Z, Fu H, et al. Modeling the impact of preplanned dose titration on delayed response. *J Biopharm Stat*. 2019;29(2):287-305. https://doi.org/10.1080/10543406.2018.1535499

Ye T, Shao J, Yi Y, Zhao Q. Toward better practice of covariate adjustment in analyzing randomized clinical trials. 2009. Accessed March 30, 2022. https://arxiv.org/pdf/2009.11828.pdf

Signature Page for VV-CLIN-074016 v1.0

Approval	PPD 13-Dec-2022 12:12:32 GMT+0000

Signature Page for VV-CLIN-074016 v1.0