J2A-MC-GZGI Statistical Analysis Plan Version 3

A Phase 2 Study of Once Daily LY3502970 Compared with Placebo in Participants Who Have Obesity or Are Overweight with Weight-Related Comorbidities

NCT05051579

Approval Date: 3-Oct-2022

Statistical Analysis Plan (J2A-MC-GZGI): A Phase 2 Study of Once-Daily LY3502970 Compared with Placebo in Participants Who Have Obesity or Are Overweight with Weight-Related Comorbidities

Protocol Title: A Phase 2 Study of Once Daily LY3502970 Compared with Placebo in Participants Who Have Obesity or Are Overweight with Weight-Related Comorbidities

Protocol Number: J2A-MC-GZGI

Amendment Number: a

Compound: LY3502970

Brief Title: Effect of LY3502970 versus Placebo in Participants Who Have Obesity or Are Overweight

Study Phase: 2

Sponsor Name: Eli Lilly and Company

Legal Registered Address: Indianapolis, Indiana, USA 46285

Regulatory Agency Identifier Number(s)

IND: 156143

EudraCT: 2021-002805-88

Table of Contents

Version History

Statistical Analysis Plan (SAP) Version 1 for Study J2A-MC-GZGI (GZGI*)* was approved on 26 January 2022 based on the protocol amendment (a) dated 29 September 2021.

Statistical Analysis Plan Version 2 was approved prior to the interim analysis after all participants complete Visit 12 (Week 16) of the treatment period. The main changes are summarized below:

- Changed wording in the postbaseline observation for the logistic regression "efficacy estimand" to clarify that only Visit 15 and 18 will be modelled for primary and final lock respectively
- Further clarified the hybrid estimand description and method
- Updated the "Treatment Discontinuation Reasons" table to match that of current case report form (CRF)
- Clarified analyses for the pancreatic enzymes
- Updated shift table categories for hepatic biomarkers alanine aminotransferase (ALT)
- Updated analyses related to hypersensitivity events and removed by period analysis
- Removed 2 listings related to coronavirus disease 2019 (COVID-19) pandemic
- Updated variables used in demographics table
- Clarified that mechanistic biomarkers analyses will be conducted through mixed-effect model for repeated measures (MMRM) on hepatic biomarkers (Cytokeratin 18 and Pro-C3), and
- Added the analysis for abuse liability.

Statistical Analysis Plan Version 3 was approved prior to the primary lock analysis after all participants complete Visit 15 (Week 26) of the treatment period. The main changes are summarized below:

- Changed Safety Analyses to be conducted using Safety Analysis Set (SS) for Week 26 instead of Full Analysis Set (FAS).
- Lipid analyses will be part of laboratory analyses for safety analyses set and will also be analyzed using efficacy analysis set separately.
- Removed typo for baseline analyses which will be using ANOVA instead of ANCOVA.
- Chi-square test has been added for categorical analysis in case Fisher's exact test fails.
- ABPM analyses will be using Safety Analysis Set (SS) instead of Efficacy Analysis Set (EAS).
- Clarified the limitations in the planned Bayesian model for long-term prediction due to the study design. Added that additional modelling will be explored for better weight loss prediction.
- Updated dose modification according to the protocol.
- Fasting Blood Glucose analyses have been changed to Fasting Serum Glucose.
- Updated the efficacy analysis for lipids from total cholesterol to all fasting lipids.
- Kaplan Meier plot of time to initially achieve >=5%, >=10% weight loss have been added.
- MMRM analyses for Heart Rate and PR-interval have been added.
- Updated the analyses for Liver enzymes based on new guidelines.

- Updated analyses for abuse liability potential and drug abuse, dependence and withdrawal based on new guidance.
- Updated “Changes to Protocol” section to “None”.

1. Introduction

1.1. Objectives, Endpoints, and Estimands

Objectives	Endpoints
Primary	
To demonstrate that at least 1 dose level of QD oral LY3502970 is superior in percent body weight reduction relative to placebo	• Percent change in body weight from baseline at Week 26
Secondary	
To compare the effect of QD LY3502970 versus placebo on body weight	• Percent change in body weight from baseline at Week 36 • Change in body weight (kg) from baseline at Week 26 and Week 36 • Percentage of study participants who achieve o ≥5% body weight reduction o ≥10% body weight reduction at Week 26 and Week 36 • Change in BMI (kg/m^2) from baseline at Week 26 and Week 36
To compare the effect of QD LY3502970 versus placebo on waist circumference	• Change in waist circumference from baseline at Week 26 and Week 36
To assess safety and tolerability of study interventions	• AEs overall • AEs of special interest • Laboratory parameters • Electrocardiogram • Vital signs
To assess the PK of LY3502970 and potential participant factors that may influence its PK	• Population PK Parameters

Objectives	Endpoints
Exploratory	
To assess the relationship between LY3502970 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 QD LY3502970 versus placebo on body weight control	• Percentage of study participants who achieve ≥15% body weight reduction at Week 26 and Week 36
To compare the effect of QD LY3502970 versus placebo on BP	• Change from baseline in o Systolic BP (mmHg) measured by ABPM o Diastolic BP (mmHg) measured by ABPM at Week 26
To compare the effect of QD LY3502970 versus placebo on heart rate	• Change in heart rate from baseline measured by ABPM at Week 26
To compare the effect of QD LY3502970 versus placebo on lipid parameters	• Change from baseline in fasting o Total cholesterol o HDL cholesterol o LDL cholesterol o VLDL cholesterol o Triglycerides at Week 26 and Week 36
To compare the effect of QD LY3502970 versus placebo on glucose control	• Change in HbA1c from baseline at Week 26 and Week 36 • Change in FBG from baseline at Week 26 and Week 36 • Percentage of study participants who develop T2DM at Week 26 and Week 36
To compare the effect of QD LY3502970 versus placebo on mechanistic biomarkers	• Change in mechanistic biomarkers from baseline at Week 26 and Week 36
To evaluate the effects of QD LY3502970 versus placebo on patient-reported outcomes • Health-related quality of life	

Objectives	Endpoints
• Participant experience	• Change from baseline in SF-36v2 acute form domain and summary scores at Week 26 and Week 36 • Change from baseline in IWQOL-Lite-CT composite and total scores at Week 26 and Week 36 • Actual and change from baseline in PGIS for Physical Activity at Week 26 and Week 36 • Summary statistics of actual responses to participant survey at Week 26

Abbreviations: ABPM = ambulatory blood pressure monitoring; AE = adverse event; BMI = body mass index; BP = blood pressure; FBG = fasting blood glucose; HbA1c = hemoglobin A1c; HDL = high-density lipoprotein cholesterol; IWQOL-Lite-CT = Impact of Weight on Quality of Life-Lite Clinical Trials Version; LDL = low-density lipoprotein cholesterol; PGIS = Patient Global Impression of Status for Physical Activity; PK = pharmacokinetics; QD = once daily; SF-36v2 = Short Form 36 Version 2 health survey acute form; T2DM = Type 2 Diabetes Mellitus; VLDL = very low-density lipoprotein cholesterol.

Primary Estimand

The primary clinical question of interest is: What is the treatment difference in percent change in body weight after 26 weeks of treatment in participants who meet the inclusion criteria and would have completed the treatment period?

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

- Population: participants who meet the inclusion criteria. Further details can be found in Section 5 and Section 9 of the Study Protocol J2A-MC-GZGI (a).
- Endpoint: percent change from baseline to 26 weeks in body weight
- Treatment condition: the randomized treatment with allowance for down-titration based on gastrointestinal (GI) tolerability.

The intercurrent event (ICE) "permanent discontinuation of study drug" is 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. There are no other defined ICEs. Down-titration will not be considered as ICEs for the definition of this estimand in this study.

Population-level summary: difference in mean percent changes in body weight at Week 26 between once daily (QD) LY3502970 and placebo.

Rationale for "efficacy" estimand: This Phase 2 study aims to study the efficacy of LY3502970 under the ideal condition that all participants adhere to the randomized treatment.

Estimand(s) for Secondary Objectives

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

- percent change in body weight from baseline at Week 36
- change in body weight (kg) from baseline at Week 26 and Week 36
- percentage of study participants who achieve ≥5% body weight reduction at Week 26 and Week 36
- percentage of study participants who achieve ≥10% body weight reduction at Week 26 and Week 36
- change in body mass index (BMI) (kg/m^2) from baseline at Week 26 and Week 36
- change in waist circumference from baseline at Week 26 and Week 36

Unless specified otherwise, safety and tolerability assessments will be conducted by comparing the safety of LY3502970 doses with placebo for the entire study period (the treatment period plus safety follow-up period where applicable), irrespective of adherence to study intervention for all study population (including inadvertent enrollment).

Supplemental Estimand(s) for Primary Efficacy Endpoint

An alternative estimand, called "hybrid estimand," is the treatment difference in the mean percentage change in body weight from baseline to 26 weeks between LY3502970 and placebo in participants who meet the inclusion criteria, with ICEs handled differently 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 will be handled by the hypothetical strategy.
- Category 3: All other ICEs will be handled by the treatment policy strategy.

Population-level summary: the mean percentage changes in body weight from baseline to Week 26

Rationale for "hybrid estimand:" Following the International Conference on Harmonisation (ICH) E9 (R1) guidance on estimands, this Phase 2 study will collect informative treatment disposition reasons, and ICEs will be handled differently according to the reasons of ICEs for this supplemental estimand. The ICEs of Category 1 are related to situations which will likely not occur in the future (inability to participate in the clinical trials due to corona virus 2019 (COVID-19) pandemic quarantines, schedule conflicts, travel, relocation, etc., which will not prevent patients continuing taking the medication in real world). For ICEs of Category 2, we would like to understand the efficacy if patients continue taking the medication, a hypothetical situation that patients can achieve. For ICEs of Category 3, patients may not be able to adhere to the study medication potential due to insurmountable reasons (for example, adverse events), estimating the potential outcome under the hypothetical strategy does not make sense and therefore treatment policy strategy is used.

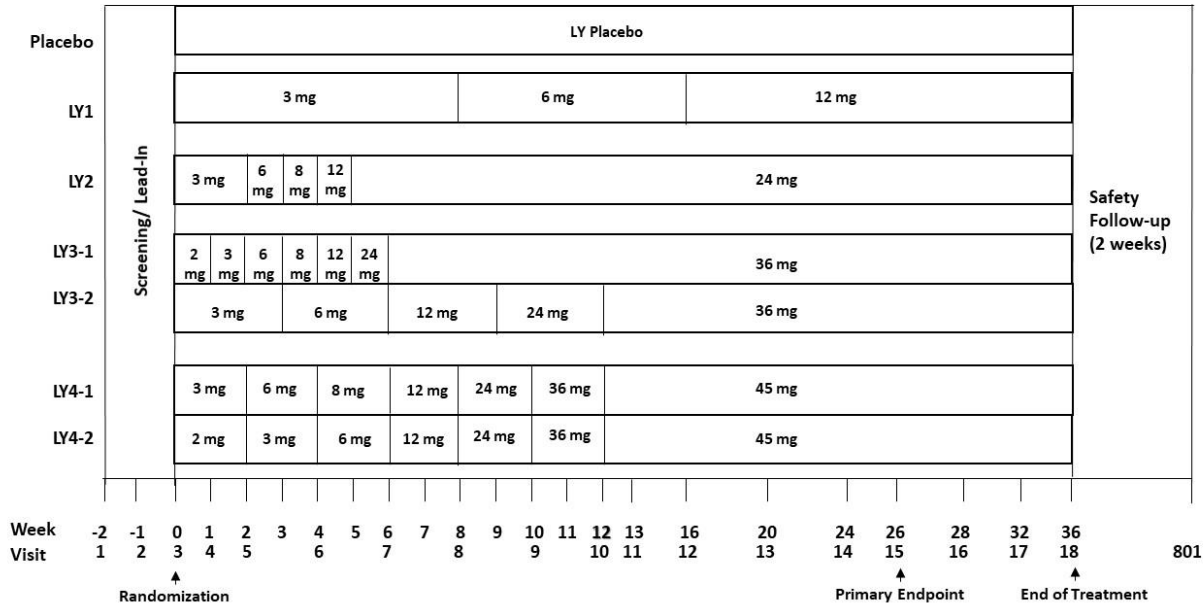
Further details can be found in Section 4.3.3.

1.2. Study Design

Study GZGI is a 36-week Phase 2, multicenter, randomized, double-blind, parallel, placebo-controlled study designed to examine the efficacy and safety of 4 dose levels of QD administered LY3502970 compared with QD administered placebo in participants who have obesity (BMI ≥30 kg/m^2) or are overweight (BMI ≥27 kg/m^2 and <30 kg/m^2) with weight-related comorbidities without diabetes. The duration of study participation for each participant will be approximately 40 weeks. The study will consist of an approximately 2-week screening/lead-in period followed by a 36-week treatment period. There will also be a 2-week off-drug safety follow-up period.

Study participants will be randomly assigned to study intervention in a randomization ratio of 5:5:5:3:3:3:3, including 50 participants per treatment group to placebo, LY3502970 12 mg (LY1) and LY3502970 24 mg (LY2) treatment groups, and 30 participants per treatment group to LY3502970 36 mg (LY3)-1, LY3-2, LY3502970 45 mg (LY4)-1, and LY4-2 treatment groups. This is a double-blind study. All participants will take daily oral blinded study drug. Participants will not know if they are receiving active or placebo oral study drug. Stratification will be by sex and BMI (≤35, >35 kg/m^2) at Visit 1.

Figure 1 illustrates the study design.



Abbreviations: LY = LY3502970; LY1 = LY 12 mg; LY2 = LY 24 mg; LY3 = LY 36 mg; LY4 = LY 45 mg.

Figure 1 Illustration of study design for clinical protocol GZGI.

Dose Modification If a participant does not tolerate a specific dose level for 1 week (for example, due to moderate-to-severe nausea, vomiting, diarrhea) and the Investigator does not believe that the participant will tolerate the dose with further exposure, then the investigator may reduce the dose to the next lower target dose per the instruction below.

To maintain blinding, the Investigator will use the Interactive Web Response System (IWRS) web site portal to receive the appropriate study drug dispensing information to preserve blinding of the study drug. The IWRS will manage the dose escalation.

LY1 (3/6/12 mg): A participant who does not tolerate the 3-mg dose will need to discontinue the study drug. If the participant does not tolerate either the 6 mg or 12 mg doses, then the participant will remain on the 3-mg dose (6 mg will not be a maintenance dose).

LY2, LY3-1, LY3-2, LY4-1, and LY4-2: A participant who does not tolerate the first dose level, either 3 mg or 2 mg, will need to discontinue the study drug. Participants who do not tolerate either 6 mg or 8 mg will also need to discontinue study drug: neither 6 mg or 8 mg will be a maintenance dose. If a participant does not tolerate a dose level above 12 mg, then the dose should be dropped to the next maintenance dose level (12 mg, 24 mg, or 36 mg). If this dose is tolerated for 2 weeks, the dose should be increased per original protocol dose escalation until the maintenance dose in achieved. However, if this dose escalation is not tolerated, the dose should be reduced to the next lower target dose that was tolerated (for example, 12 mg, 24 mg, or 36 mg). The participant will remain at that dose level for the duration of the study.

Temporary Discontinuation After randomization, the Investigator may interrupt study drug, for example, due to an adverse event (AE) (for example, nausea of moderate severity or vomiting), or a clinically significant laboratory value. If study drug interruption is due to an AE, the event is to be followed and documented. Every effort should be made by the Investigator to maintain participants in the study and to restart study drug promptly after any interruption, as soon as it is safe to do so (see the next paragraph for restarting study drug). The dates of study drug interruption and restart must be documented. The data related to interruption of study treatment will be documented in source documents and entered on the electronic case report form (eCRF).

Restarting Study Drug after Interruption If the number of consecutive missed capsule doses is ≤7, the treatment can be restarted at the same dose if the drug was well tolerated prior to discontinuation.

Participants who have missed >7 days of study drug will need to restart the study drug at the 8-mg dose (LY3502970 treatment groups 2, 3, 4-1, and 4-2) and dose escalate according to the protocol, including if the participants were in the early part of the dose escalation and were taking less than 8 mg. Participants who have missed >7 days of study drug in the LY1 treatment group will need to restart drug at 6 mg. To maintain blinding of the Investigator and participant, the Investigator should call the IWRS to explain that the participant needs the dose reduced or is restarting study drug and IWRS will provide dispensing information. Dose reductions may occur at unscheduled visits.

The Investigator will use the IWRS web site portal to receive the appropriate study drug dispensing information to preserve blinding of the study drug.

2. Statistical Hypotheses

The study hypothesis for the primary objective is that at least 1 dose level of QD oral doses of LY3502970 is superior in percent reduction from baseline in body weight relative to placebo at Week 26 in participants without diabetes who have obesity or are overweight with at least 1 weight-related comorbidity. Thus, the null hypothesis corresponding to the primary estimand is as follows:

- Null hypothesis: No dose level of QD oral doses of LY3502970 is different from placebo in percent reduction from baseline in body weight at Week 26 in participants without diabetes who have obesity or are overweight with at least 1 weight-related comorbidity.

2.1. Multiplicity Adjustment

Treatment comparisons will be performed for the primary objective at the full significance level of 0.05. No multiplicity adjustments will be made for the analysis of secondary and exploratory objectives.

3. Analysis Sets

For the purpose of analysis, the following analysis population and datasets are defined.

Table 1 Description of Analysis Datasets

Participant Analysis Population/Datasets	Description
Entered Participants	All participants who sign informed consent.
Randomized	All participants who are randomly assigned a study drug.
Efficacy Analysis Set (EAS)	Data obtained during the treatment period from all randomized participants who are exposed to at least 1 dose of study drug. Excludes data after permanent discontinuation of study drug. Participants will be included in the treatment group to which they were randomly assigned.
Full Analysis Set (FAS)	Data obtained during the treatment period from all randomized participants who are exposed to at least 1 dose of study drug, regardless of adherence to study drug. Participants will be included in the treatment group to which they were randomly assigned.
Safety Analysis Set (SS)	Data obtained during the treatment period plus safety follow-up from all randomized participants who are exposed to at least 1 dose of study drug, regardless of adherence to study drug. Participants will be included in the treatment group to which they were randomly assigned.

Note: Inadvertently enrolled participants who are discontinued from the study due to deviation from inclusion/exclusion criteria will not be included for efficacy analyses but will be included for all other analyses.

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 (eg, too few events to justify conducting an analysis). Additional analyses of the data may be conducted as deemed appropriate.

Unless otherwise noted, all tests of treatment effects will be conducted at a 2-sided alpha level of 0.05, and confidence intervals (CIs) 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, LY1, LY2, LY3-1, LY3-2, LY4-1, or LY4-2), regardless of the actual treatment(s) received by the participant due to any dose modification. The evaluation of the efficacy and safety endpoints for LY3 (36 mg) and LY4 (45 mg) compared with placebo will be made by pooling 2 dose escalation regimens, ie, combine LY3-1 and LY3-2 for LY3 and combine LY4-1 and LY4-2 for LY4. Therefore, the statistical comparisons between treatment groups will be between each of LY3502970 treatment groups (LY1, LY2, LY3, or LY4) and placebo. In addition, to better understand the tolerability of LY3502970, statistical comparisons will be conducted between the 2 dose escalation regimens with a maintenance dose of 36 mg or 45 mg on endpoints regarding GI reactions.

The primary estimand (a precise definition of the treatment effect to be estimated) of interest in comparing efficacy of LY3502970 doses with placebo is the "efficacy estimand" (Section 1.1). The primary efficacy assessment, guided by the "efficacy estimand" will be conducted using the Efficacy Analysis Set (EAS) (Section 3). A restricted maximum, likelihood-based 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. The model for the analysis of the primary efficacy endpoint of percent change from baseline in body weight will include the fixed class effects of treatment group (placebo, LY1, LY2, LY3-1, LY3-2, LY4-1, LY4-2), strata (sex, baseline BMI stratum [$\leq$35 kg/m^2, >35 kg/m^2]), visit, and treatment-by-visit interaction, as well as the continuous, fixed covariate of baseline value. An unstructured covariance structure will be used to model the within-participant errors. Significance tests will be based on least squares (LS) means and Type III tests. If this analysis fails to converge, the following covariance structures will be tested in order:

- Toeplitz with heterogeneity
- Autoregressive with heterogeneity
- Compound symmetry with heterogeneous variances
- Toeplitz
- Autoregressive
- Compound symmetry without heterogeneous variances

The first covariance structure that converges will be used. The Kenward-Roger approximation will be used to estimate denominator degrees of freedom.

Unless specified otherwise, safety assessments will be conducted by comparing safety of LY3502970 doses with placebo irrespective of adherence to study drug. Thus, safety analyses will be conducted using the Safety Analysis Set (SS) for the primary database lock at Week 26 and for the final database lock.

For continuous measures, summary statistics will include sample size, mean, standard deviation (SD), median, minimum, and maximum for both the actual and the change from baseline measurements. Least squares 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 (ANOVA) model that has treatment group as the model term.

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 (Section 4.3.3). Fisher's exact test (or Chi-square test where Fisher's exact test fails) will be used for treatment comparisons in other categorical outcomes.

For laboratory values, both conventional (CN) and Systeme International (SI) units will be presented. For body weight, kilogram (kg) will be presented.

Details about the analyses regarding demographic and baseline characteristics, historical illnesses and pre-existing conditions, treatment compliance, concomitant medications, and important protocol deviations can be found in Appendices 1-5 (Section 7.1 through Section 7.6), respectively.

Baseline is defined as the last nonmissing measurement (including both scheduled and unscheduled visits) at or before randomization visit (prior to first dosing of study drug) unless otherwise specified. For variables that are not collected at each postbaseline visit, data may exist at visits where the variable was not scheduled to be collected. In these situations, data from the early termination (ET) visit that do not correspond to the planned collection schedule will be excluded from the MMRM, ANCOVA, or logistic regression analysis.

Table 2 summarizes the rules for determining the estimand, analysis set, and baseline and postbaseline observations by study period and type of analysis.

Table 2 Baseline and Postbaseline Definitions, Analysis Set, and Estimand by Study Period and Type of Analysis

Study Period/Analysis Type (Estimand)	Analysis Set	Baseline Observations	Postbaseline Observations
26-Week Treatment Period			
Body Weight, BMI, MMRM ("efficacy estimand")	EAS	Last of Visits 1 to 3	Visits 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, and 15 Prior to any ICEs

Approved on 03 Oct 2022 GMT

Study Period/Analysis Type (Estimand)	Analysis Set	Baseline Observations	Postbaseline Observations
Body Weight, BMI ANCOVA ("hybrid estimand")	FAS	Last of Visits 1 to 3	Visit 15, with missing values or observed values at the visit that are collected after ICEs are handled as described in Section 4.3.3
Study Period/Analysis Type (Estimand)	**Analysis Set**	**Baseline Observations**	**Postbaseline Observations**
Body Weight categorical analyses logistic regression ("efficacy estimand)	EAS	Last of Visits 1 to 3	Visit 15, with missing values imputed using non-missing data from Visits 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, and 15 prior to any ICEs
Body Weight categorical analyses logistic regression ("hybrid estimand)	FAS	Last of Visits 1 to 3	Visit 15, with missing values or observed values at the visit that are collected after ICEs are handled as described in Section 4.3.3 and Section 4.3.2
Waist Circumference MMRM ("efficacy estimand")	EAS	Last of Visits 1 to 3	Visits 6, 8, 10, 12, 13, and 15 prior to any ICEs
Waist Circumference ANCOVA ("hybrid estimand")	FAS	Last of Visits 1 to 3	Visit 15, with missing values or observed values at the visit that are collected after ICEs are handled as described in Section 4.3.3
ABPM ANCOVA and categorical analysis	SS	Visit 3	Visit 15 prior to any ICEs
HbA1c and FSG MMRM ("efficacy estimand")	EAS	Last of Visits 1 to 3	Visits 10 and 15 prior to any ICEs
Develop T2DM categorical analysis logistic regression ("efficacy estimand")	EAS	Last of Visits 1 to 3	Visits 10 and 15 prior to any ICEs
Population PK parameters (N/A)	FAS	N/A	Predose at Visits 8, 10, and 13 Postdose at Visits 6, 8, 12, 15, and ET
Patient-reported outcomes: SF-36v2, IWQOL-Lite CT, and PGIS for Physical Activity ("efficacy estimand")	EAS	N/A	Visits 12 and 15 prior to any ICEs

Study Period/Analysis Type (Estimand)	Analysis Set	Baseline Observations	Postbaseline Observations
Patient-reported outcomes: Participant Survey Statistical Summary ("efficacy estimand")	EAS	N/A	Visit 15 prior to any ICEs
26-Week Treatment Period			
1.1) Treatment-Emergent Adverse Events	SS	The baseline period is defined as the start of screening and ends prior to the first dose of study drug (or prior to Visit 3 date if first dose date is missing).	Starts after the first dose of study drug and ends prior to the study drug disposition visit unless otherwise noted.
1.2) Treatment-Emergent Abnormal Labs, Vital Signs, Lipids, and ECGs	SS	Baseline will be all scheduled and unscheduled measurements recorded during the baseline period as defined above (1.1).	Postbaseline will be defined as above (1.1). All scheduled and unscheduled measurements will be included.
1.3) Change from Last Baseline to Week *xx* and to Last Postbaseline for Labs, Vital Signs, and ECGs	SS	The last scheduled non-missing assessment recorded prior to the date of first dose of study treatment during the baseline period defined above (1.1).	Postbaseline will be defined as above (1.1). Only scheduled visits will be included. The ET visits are considered scheduled visits.
36-Week Treatment Period (not including Safety Follow-up Visit)			
Body Weight, BMI, MMRM ("efficacy estimand")	EAS	Last of Visits 1 to 3	Visits 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, and 18 prior to any ICEs
Body Weight, BMI ANCOVA ("hybrid estimand")	FAS	Last of Visits 1 to 3	Visit 18, with missing values or observed values at the visit that are collected after ICEs handled as described in Section 4.3.3
Body Weight categorical analyses logistic regression ("efficacy estimand")	EAS	Last of Visits 1 to 3	Visit 18 with missing values imputed using non-missing data from Visits 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, and 18 prior to any ICEs
Body Weight categorical analyses logistic regression ("hybrid estimand)	FAS	Last of Visits 1 to 3	Visit 18, with missing values or observed values at the visit that are collected after ICEs are handled as described in Section 4.3.3 and Section 4.3.2

Study Period/Analysis Type (Estimand)	Analysis Set	Baseline Observations	Postbaseline Observations
Waist Circumference MMRM ("efficacy estimand")	EAS	Last of Visits 1 to 3	Visits 6, 8, 10, 12, 13, 15, 17, and 18 prior to any ICEs
Waist Circumference ANCOVA ("hybrid estimand")	FAS	Last of Visits 1 to 3	Visit 18, with missing values or observed values at the visit that are collected after ICEs are handled as described in Section 4.3.3
HbA1c and FBG MMRM ("efficacy estimand")	EAS	Last of Visits 1 to 3	Visits 10, 15, and 18 prior to any ICEs
Develop T2DM categorical analysis logistic regression ("efficacy estimand")	EAS	Last of Visits 1 to 3	Visits 10, 15, and 18 prior to any ICEs
Population PK parameters (N/A)	FAS	N/A	Predose at Visits 8, 10, 13, and 18 Postdose at Visits 6, 8, 12, 15, and ET
Patient-reported outcomes: SF-36v2, IWQOL-Lite CT, and PGIS for Physical Activity ("efficacy estimand")	EAS	N/A	Visits 12, 15, and 18 prior to any ICEs
Patient-reported outcomes: Participant Survey Statistical Summary ("efficacy estimand")	EAS	N/A	Visit 15 and ET prior to permanent discontinuation of study drug
36-Week Treatment Period (including Safety Follow-Up Visit where applicable)			
1.1) Treatment-Emergent Adverse Events	SS	The baseline period is defined as the start of screening and ends prior to the first dose of study drug (or prior to Visit 3 date if first dose date is missing).	Starts after the first dose of study drug and ends at the end of the study period (including off-drug follow up visit)..
1.2) Treatment-Emergent Abnormal Labs, Vital Signs, and ECGs	SS	Baseline will be all scheduled and unscheduled measurements recorded during the baseline period as defined above (1.1).	Postbaseline will be defined as above (1.1). All scheduled and unscheduled measurements will be included.

Study Period/Analysis Type (Estimand)	Analysis Set	Baseline Observations	Postbaseline Observations
1.3) Change from Last Baseline to Week *xx* and to Last Postbaseline for Labs, Vital Signs, and ECGs	SS	The last scheduled nonmissing assessment recorded prior to the date of first dose of study treatment during the baseline period defined above (1.1).	Postbaseline will be defined as above (1.1). Only scheduled visits will be included. The ET visits are considered scheduled visits.

Abbreviations: ABPM = ambulatory blood pressure monitoring; ANCOVA = analysis of covariance; BMI = body mass index; EAS = Efficacy Analysis Set; ECG = electrocardiogram; ET = early termination; FAS = Full Analysis Set; FBG = fasting blood glucose; HbA1c = hemoglobin A1c; ICE = intercurrent event; IWQOL-Lite-CT = Impact of Weight on Quality of Life-Lite Clinical Trials Version; MMRM = mixed-effect model repeated measures; N/A = not applicable; PGIS = Patient Global Impression of Status for Physical Activity; PK = pharmacokinetics; SF-36v2 = Short Form 36 Version 2 health survey acute form; SS = Safety Analysis Set; T2DM = type 2 diabetes mellitus.

Note: If a participant ED from the study drug, no further PK samples will be collected following ED PK sample.

Therefore, FAS for PK parameters only includes data that are collected while participant is on study drug.

4.2. Participant Dispositions

A listing of study disposition for all randomized participants will be provided. Summaries of study disposition and study drug disposition for all randomized participants will be provided by treatment groups.

Kaplan-Meier plots of time from randomization to premature study drug discontinuation, and premature study drug discontinuation due to AE, may be provided based on all randomized population. Time-to-event analyses of premature study treatment discontinuation, study drug discontinuation due to AE, and study discontinuation may be conducted.

4.3. Primary Estimand Analysis

The primary efficacy assessment, guided by the “efficacy estimand”, will be conducted using the EAS for the primary endpoint (percent change from baseline in body weight at Week 26).

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

To confirm efficacy of LY3502970 with adequate statistical power, the evaluation of the primary efficacy endpoint for LY3 (36 mg) and LY4 (45 mg) compared to placebo will be made by pooling 2 dose escalation regimens, i.e., combine LY3-1 and LY3-2 for LY3 and combine LY4-1 and LY4-2 for LY4.

The primary efficacy comparison will be based on the contrast between each treatment group of LY3502970 and placebo at Week 26 (Visit 15) from the MMRM analysis of percentage change from baseline in body weight using the EAS (Section 3). 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(s)

The primary efficacy measure will be the percent change in body weight from baseline at Week 26.

The percent change in body weight at each nominal visit is defined as:

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

4.3.2. Main Analytical Approach

The percent change in body weight from baseline at Week 26 will be analyzed using an MMRM model for the "efficacy estimand" as described in Section 4.1.

4.3.3. Analyses for a Supplemental Estimand

ICH E9 (R1) Addendum provides a general framework for defining estimands for clinical trials. The Addendem recommends handling intercurrent events according to the study objectives and the causes of the intercurrent events. Naturally, it requires mixed strategies to handle various types of intercurrent events due to different reasons in the same study. This approach has been advocated in the CHMP draft guidance for the development of treatment for diabetes mellitus (CPMP/EWP/1080/00 Rev.2) and several recent publications (Darken et al. 2020, Qu et al 2021, Qu and Lipkovich 2021). In addition, the imputation of missing values should be consistent with the potential outcome of interest and the assumption for the missingness should be based on the causes of the intercurrent events and missingness (ICH E9 (R1), Qu and Lipkovich 2021).

Therefore, an analysis related to a supplemental estimand, "hybrid estimand" (Section 1.1), will be conducted following the potential outcomes framework introduced by Lipkovich et al (2020).

The hybrid estimand is defined as the treatment difference in the mean percent change in body weight from baseline to 26 weeks between LY3502970 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. Generally, these ICEs will also cause participants to discontinue from the study. The potential outcome of interest is the response in the efficacy measurement if participants had continued the study drug. This is of best clinical interest as in real life—either these ICEs do not prevent participants from taking the medication or these ICEs (pandemic or geographic conflicts) do not represent the situation at normal time.
- Category 2: The ICEs of permanent discontinuation of study drug due to lack of efficacy will be handled by the hypothetical strategy. The potential outcome of interest is the response in the efficacy measurement if participants had continued the study drug.

- Category 3: All other ICEs will be handled by the treatment policy strategy. The potential outcome of interest is the response without taking the study drug from the time of treatment discontinuation events.

In this study, following ICH E9 (R1) guidance, a plan was made to collect informative treatment disposition reasons through the eCRF for why data intended for collection are missing and classify them into Categories 1-3 as shown in Table 3 (EMA 2020).

Table 3 Treatment Disposition Reasons

Disposition Reason	Associated subcategories	Category
Adverse Event		3
Death		3
Protocol Deviation	Due to Epidemic/Pandemic	1
	Other (option to include a specify field)	3
Pregnancy		3
Lack of Efficacy		2
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
Lost to Follow Up	Due to Epidemic/Pandemic	3
	Other	3
Other	Due to Epidemic/Pandemic	1
	Other	3

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

To estimate the “hybrid estimand”, multiple imputation will be used to impute the corresponding missing potential outcome according to the missingness patterns (Table 4, with ICEs handled differently according to the reasons of the events.

When participants have missing values without ICEs, the missing values will be imputed using data from participants from the same treatment group who do not have ICE or missing values under the missing at random (MAR) assumption.

Percent change from baseline in body weight will then be analyzed using an ANCOVA with terms for treatment (placebo, LY1, LY2, LY3-1, LY3-2, LY4-1, LY4-2) and strata, and the baseline value as a covariate.

Table 4 Strategy to Handle ICE and Missingness for Hybrid Estimand

ICE	Strategy to Handle ICE	Potential Outcome	Assumptions for Missingness	Methods to Handle Missing Values at Endpoint
Category 1: Treatment discontinuation due to reasons unlikely related to efficacy/safety outcome	Hypothetical	The response if participants had continued the study drug	MAR (since missing data unlikely depend on the efficacy outcomes)	Data collected after the ICE will be discarded. Missing values will be imputed using all non-missing data (excluding data collected after ICEs) from the same treatment group under the MAR assumption.
Category 2: Treatment discontinuation due to lack of efficacy	Hypothetical	The response if participants had continued the study drug	MAR (since such an intercurrent event can likely be modeled by the observed efficacy data)	Data collected after the ICE will be discarded. Missing values will be imputed using all non-missing data (excluding data collected after ICEs) from the same treatment group under the MAR assumption.
Category 3: All other Treatment discontinuations	Treatment policy	The response without taking study drug from the time of treatment discontinuation events	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 group with similar ICEs but non-missing values (retrieved dropout imputation), including the data collected after the ICE. In cases where there are not enough retrieved dropouts to provide a reliable imputation model, missing values will be imputed using the jump-to-reference (placebo) imputation approach.

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

4.4. Secondary Endpoint(s)/Estimands Analysis

4.4.1. Secondary Endpoints

The following secondary study objectives will be analyzed with the "efficacy estimand" using the data in the EAS:

- Percent change in body weight from baseline at Week 36
- Change in body weight (kg) from baseline at Week 26 and Week 36
- Change in BMI (kg/m^2) from baseline at Week 26 and Week 36
- Change in waist circumference from baseline at Week 26 and Week 36

- Percentage of study participants who achieve ≥5% body weight reduction at Week 26 and Week 36
- Percentage of study participants who achieve ≥10% body weight reduction at Week 26 and Week 36

4.4.1.1. Main Analytical Approach

Actual and change from baseline in body weight, BMI, and waist circumference and percent change in body weight will be analyzed using the MMRM model for the "efficacy estimand" as described in Section 4.1.

Missing values for binary endpoints will be imputed by categorizing (Yes, No) the imputed values for the corresponding continuous endpoints, as described in Section 4.3.3, and then analyzed using a logistic regression model with treatment and strata as fixed effects and the continuous baseline value as a covariate.

4.4.2. Analyses for a Supplemental Estimand

The secondary endpoints listed in Section 4.4.1 will also be analyzed with the "hybrid estimand" using the data in the FAS. Continuous endpoints with missing values imputed and then will be analyzed by an ANCOVA model as described in Section 4.3.3. Binary endpoints will be analyzed using the same way as described in Section 4.4.1.1. Such analyses for exploratory endpoints (e.g., percentage of study participants who achieve ≥15% body weight reduction at Week 26 and Week 36) may also be performed.

4.5. Exploratory Endpoint

4.5.1. Exploratory Efficacy Analysis

The following exploratory study objectives will be analyzed using the MMRM model for the "efficacy estimand" as described in Section 4.1 or using the logistic regression model for the "efficacy estimand" as described in Section 4.3.3.

- Change from baseline in fasting lipids at Week 26 and Week 36 using MMRM
- Change from baseline in hemoglobin A1c (HbA1c) at Week 26 and Week 36 using MMRM
- Change from baseline in FSG at Week 26 and Week 36 using MMRM
- Change from baseline in mechanistic biomarkers (Hepatic Biomarkers - Cytokeratin 18 and Pro-C3) at Week 26 and Week 36 using MMRM
- Percentage of study participants who achieve ≥15% body weight reduction at Week 26 and Week 36 using logistic regression.
- Percentage of study participants who develop T2DM at Week 26 and Week 36 using logistic regression.
- Kaplan-Meier plot of time to initially achieve body weight loss of ≥5% and ≥10%

4.5.2. Pharmacokinetic and Pharmacokinetic/Pharmacodynamic Methods

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

A summary of LY3502970 concentration-time data will be reported in the clinical study report. Exposure-response analysis between LY3502970 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 (NONMEM) 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.

4.5.3. Patient-Reported Outcomes

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

In addition to the following analyses, actual responses to participant survey at Week 26 and ET visits will be summarized by treatment group.

Item-level missingness is dealt with 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.

4.5.3.1. Patient Global Impression of Status for Physical Activity

The counts and percent of participants for Patient Global Impression of Status for Physical Activity (PGIS) response categories at each time point will be summarized by nominal visit and by treatment. A shift table from baseline to postbaseline of 5 PGIS response categories will be created at each postbaseline visit.

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

Per copyright owner, the PRO CoRE 2.0 Smart Measurement® System will be used to derive the following domain and component scores:

- Mental Component Score
- Physical Component Score
- Physical Functioning domain
- Role-Physical domain
- Bodily Pain domain
- General Health domain
- Vitality domain
- Social Functioning domain
- Role-Emotional domain, and
- Mental Health domain.

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

- Descriptive summaries by treatment group, and
- MMRM analysis as described in Section 4.1.

4.5.3.3. Impact of Weight on Quality of Life-Lite Clinical Trial

The following parameters will be included from the Impact of Weight on Quality of Life-Lite Clinical Trial (IWQOL-Lite-CT):

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

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

The IWQOL-Lite-CT scores are computed according to the IWQOL-Lite scoring rules (Kolotkin and Crosby 2002), as following:

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

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

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

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

- descriptive summaries by treatment group and
- MMRM analysis described in Section 4.1.

4.5.4. Bayesian Analyses for Dose-Response

The longitudinal dose-response model proposed in Qu et al. (2019) will be applied. This approach models the impact of dose titration on response longitudinally over time and can flexibly account for deviations from the pre-planned dose titration schedule that is likely to occur within the titration period of the clinical trial.

Let $\theta = (\theta_1, \theta_2, \ldots, \theta_m)$ be the m doses a participant has planned to take and $t_c = (t_{c1}, t_{c2}, \ldots, t_{cm})$ be the corresponding time points when the dose changes where t_{ci} indicates the time for dose to change from θ_i to θ_{i+1}. For characterizing the body weight change, we use a formulation of $f_{\theta,t_c}(t)$ that showed a good fitting in the analyses of historical trials and simulation studies. Therefore, the mean function of the parameter of interest at time t is modelled by:

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

where I(X) is the indicator function that takes value 1 when the condition X holds. The function, $f(t;\theta)$, 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= 26 for primary analysis and d= 36 for final analysis), $\lambda(\theta)$ is the dose-response function for the maximum response at dose θ and $k(\theta)$ is dose θ's rate parameter. In this equation, $h(t - t_{ci}) = \frac{1-e^{-\tau(t-t_{ci})}}{1-e^{-\tau d}}$. The parameter τ controls the rate at which the response (body weight) changes when titrating from one dose to the next dose.

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. 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, $Y_{\theta,t_c,jt} \sim N(f_{\theta,t_c}(t) + \frac{s_j(1-e^{-k(\theta_1)t})}{(1-e^{-k(\theta_1)d})}, \sigma^2)$.

The dose-maximum response function $\lambda(\theta)$ for body weight is provided below:

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 (NDLM). Additionally, the different formulation of $f_{\theta,t_c}(t)$ may also be considered if the current specification does not provide an adequate fit to the data.

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

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

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

Additional modelling may be explored for better long-term weight loss prediction if the model does not fit the data well, potentially due to the follow study features: (1) fast titration (weekly), (2) lack of long-term follow-up for lower doses, and (3) weight is not always measured at the week of titration.

4.6. Safety Analyses

Unless specified otherwise, safety will be assessed by comparing safety of LY3502970 doses with placebo irrespective of adherence to study drug. Thus, safety analyses will be conducted using the SS (Table 2) for 26 weeks and 36 weeks except ambulatory blood pressure monitoring (ABPM) measurements, which will be conducted using the SS at only 26 weeks. For selected lab values that are only scheduled to be measured for the treatment period, the MMRM model or ANCOVA (if MMRM model is not applicable) using the SS, will only show nominal visits during the treatment period.

4.6.1. Extent of Exposure

Summary of duration of follow-up (defined as time in days from date of randomization to the date of the last study visit) and/or duration of exposure to study drug (defined as time in days from date of first dose of study drug to date of last dose of study drug plus 1 day) will be provided by treatment group using data from the SS for the primary and final lock, in the following periods:

- 26 weeks plus safety follow-up (Visit 801) for all randomized participants
- 36 weeks plus safety follow-up (Visit 801) for all randomized participants

The following descriptive statistics will be provided: n, mean, SD, median, minimum, maximum, and sum (that is, total participant-years of exposure).

4.6.2. Treatment Emergent Adverse Events

A treatment-emergent adverse event (TEAE) is defined as an event that first occurred or worsened in severity after baseline (Table 2). The Medical Dictionary for Regulatory Activities (MedDRA) Lowest Level Term (LLT) will be used in the treatment-emergent derivation. The maximum severity for each LLT during the baseline period including ongoing medical history will be used as baseline severity. For events with a missing severity during the baseline period, it will be treated as "mild" in severity for determining treatment-emergence. Events with a missing severity during the postbaseline period will be treated as "severe" and treatment-emergence will be determined by comparing to baseline severity.

For events occurring on the day of first taking study medication, the case report form (CRF)-collected information (for example, treatment-emergent flag, start time of study treatment and event) will be used to determine whether the event was pre- versus 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 Preferred Terms (PTs) nested within System Organ Classes (SOCs). Statistical comparisons will be applied at both the SOC and PT levels. Events will be ordered by decreasing frequency within SOC. The SOC will be in alphabetical order. For events that are sex-specific, the denominator and computation of the percentage will include only participants from the given sex.

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

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

4.6.2.1. Adverse Events

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

4.6.2.2. Deaths

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

4.6.2.3. Other Serious Adverse Events

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

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

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

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 where 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.

Counts and percentages of participants with treatment-emergent abnormal sitting systolic blood pressure (SBP), sitting diastolic blood pressure (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 participants with treatment-emergent vital sign abnormalities are stated in Table 5.

Table 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

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

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

In addition, the following analyses will be conducted by treatment:

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

4.6.5. Electrocardiograms

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

Change from baseline to postbaseline values for ECG parameters will be summarized for participants 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 are based on Table 6.

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

- treatment-emergent ECG abnormalities as listed in Table 6
- QT greater than 500 msec
- QTcF greater than 500 msec, and
- treatment-emergent increase from the maximum baseline in QTcF interval of greater than 30 msec, 60 msec, or 75 msec. Maximum baseline will be the maximum non-missing observation in the baseline period. The maximum value during the treatment period will be analyzed. Scheduled and unscheduled measurements will be included.

Treatment-emergent quantitative ECG abnormalities are defined as quantitative abnormalities that first occurred after baseline. A listing of abnormal quantitative ECGs will be created.

An MMRM analysis will be conducted to analyze changes from baseline in heart rate and PR interval, separately.

Table 6 Selected Categorical Limits for ECG Data

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

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

4.6.6. Laboratory Data

4.6.6.1. Central Laboratory Measures

All laboratory data will be reported in International System of Units. Selected laboratory measures will also be reported using conventional 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 non-missing observation prior to taking first study drug. Unplanned measurements will be excluded from plots.

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

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

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

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

4.6.7. Special Safety Topics

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, in a model with strata (sex, baseline BMI stratum [≤ 35, >35] as fixed effects.

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

An MMRM analysis will be used to analyze each pancreatic enzyme with a log-transformed (postbaseline measure/baseline measure) response variable and treatment, nominal visit, strata (sex, baseline BMI stratum [≤35, >35]), and treatment-by-nominal visit interaction as fixed effects.

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 Appendix 6 (Section 7.6).

4.6.7.2. Major Adverse Cardiovascular Events

Major adverse cardiovascular events (MACE) reported by investigators are adjudicated by an independent clinical endpoint committee (CEC) in a blinded fashion.

The cardiovascular (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 [CABG] or percutaneous coronary intervention [PCI]), and cerebrovascular events, including cerebrovascular accident (stroke) and transient ischemic attack (TIA).

Only adjudicated MACE will be considered as adverse events of special interest (AESIs). 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 an 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).

4.6.7.3. Supraventricular Arrhythmias and Cardiac Conduction Disorders

Treatment-emergent supraventricular arrhythmias and cardiac conduction disorders will be considered as AESIs. 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 Appendix 6 (Section 7.6).

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

4.6.7.4. Hepatic Safety

4.6.7.4.1. Hepatobiliary Disorders

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

4.6.7.4.2. Liver Enzymes

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

Hepatic labs include alanine aminotransferase (ALT), aspartate aminotransferase (AST), total bilirubin (TBL), direct bilirubin (DBL), serum alkaline phosphatase (ALP), and gamma-glutamyltransferase (GGT).

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

- ALT: The number and percentage of participants with a measurement greater than or equal to 1 time (1X), 3 times (3X), 5 times (5X), 10 times (10X), and 20 times (20X) the performing lab upper limit of normal (ULN) during the treatment period will be summarized for all participants with a postbaseline value.
- AST: The number and percentage of participants with a measurement greater than or equal to 1 time (1X), 3 times (3X), 5 times (5X), 10 times (10X), and 20 times (20X) the performing lab upper limit of normal (ULN) during the treatment period will be summarized for all participants with a postbaseline value.
- ALP: The number and percentage of participants with a measurement greater than or equal to 2 times (2X) and 3 times (3X) the performing lab ULN during the treatment period will be summarized for all participants with a postbaseline value.
- TBL: The number and percentage of participants with a measurement greater than or equal to 2 times (2X), 5 times (5X), and 8 times (8X) the performing lab ULN during the treatment period will be summarized for all participants with a postbaseline value.
- DBL: The number and percentage of participants with a measurement greater than or equal to 2 times (2X) and 5 times (5X) the performing lab ULN during the treatment period will be summarized for all participants with a postbaseline value.

- GGT: The number and percentage of participants with a measurement greater than or equal to 2 times (2X) the performing lab ULN during the treatment period will be summarized for all participants with a postbaseline value.

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

Two plots will be provided as follows:

- Hepatocellular Drug-Induced Liver Injury (DILI) Screening Plot (TBL versus ALT or AST): Each patient is plotted (that is, 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 ≥2 × ULN within 30 days after maximum postbaseline ALT or AST ≥3 × ULN, without findings of cholestasis (defined as ALP <2 × ULN). Include all scheduled and unscheduled laboratory test values.
- Cholestatic DILI Screening Plot (TBL versus ALP): Each patient is plotted (that is, 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 ≥2 × ULN within 30 days after maximum postbaseline ALP ≥3 × 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 treatment group of LY3502970 and placebo, if data warrant. For the potential hepatocellular DILI plot, the quadrants will be: Potential Hy's Law (right upper), Cholestasis (left upper), Temple's corollary (right lower). For the potential cholestatic DILI plot, the quadrants will be: 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 ≥3 × ULN (right lower).

4.6.7.5. Hypoglycemia

All hypoglycemia will be considered as AESIs. The summaries of all hypoglycemic events will be provided by treatment group. A listing of all events of hypoglycemia along with the treatment allocation may be provided, if deemed necessary.

4.6.7.6. Thyroid Safety Monitoring

4.6.7.6.1. Calcitonin

Observed calcitonin data (a thyroid-specific laboratory assessment) will be summarized by treatment group 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 maximum baseline calcitonin value

(≤20 ng/L, >20 ng/L to ≤35 ng/L, >35 ng/L). Postbaseline: ≤20 ng/L, >20 ng/L to ≤35 ng/L, >35 ng/L to ≤50 ng/L, >50 ng/L to ≤100 ng/L, and >100 ng/L.

4.6.7.6.2. C-Cell Hyperplasia and Thyroid Malignancies

Thyroid malignancies and C-cell hyperplasia will be considered as AESIs. Treatment-emergent thyroid malignancies and C-cell hyperplasia will be identified using the MedDRA High Level Term (HLT) for Thyroid neoplasms and the 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.

4.6.7.7. Gastrointestinal Safety

4.6.7.7.1. Nausea, Vomiting, Constipation and Diarrhea

Summaries and analyses for incidence and severity of nausea, vomiting, diarrhea, constipation and the 4 events combined, will be provided by each treatment group by week and overall. Incidence of these events will be plotted by treatment group representing the actual dose that participants received.

Summary of the prevalence over time for nausea, vomiting, diarrhea, constipation and the 4 events combined will also be presented through plots.

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

4.6.7.7.2. 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 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 group.

4.6.7.8. 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 estimated glomerular filtration rate (eGFR) estimated by the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation with unit mL/min/1.73m^2, using categories (<30, ≥30 to <45, ≥45 to <60, ≥60 to <90, and ≥90 mL/min/1.73m^2) and a max-to-max shift table of urine albumin-to-creatinine ratio (UACR), using the categories UACR <30 mg/g, 30 mg/g ≤ UACR ≤ 300 mg/g, and UACR >300 mg/g (respectively, these represent normal, microalbuminuria, and macroalbuminuria).

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

4.6.7.8.1. 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 AESIs.

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

- Acute renal failure: Narrow terms in Acute renal failure SMQ and
- Chronic kidney disease: Narrow terms in Chronic kidney disease SMQ

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

4.6.7.8.2. Dehydration

Dehydration events will be captured in the Narrow terms in Dehydration SMQ.

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

4.6.7.9. Ambulatory Blood Pressure Monitoring

Ambulatory monitoring of HR and BP will be performed prior to Visit 3 (baseline) and again at Visit 15 (Week 26).

The ABPM measurements are downloaded to the ABPM vendor by the study site at the time of the participant visit. If the ABPM recording does not have ≥70% valid readings (assessed by the vendor), the entire 24-hour ABPM recordings are considered not valid and the participant is asked to wear the ABPM device for another 24-hour period. Only the valid recording session will be used in the summary and analyses of the ABPM measurements.

Only the valid readings will be used in the summary and analyses of the ABPM measurements.

For an ABPM recording session that is valid, all of the individual ABPM measurements that are flagged as valid and were recorded during the first 24 hours of recordings will be used in the analyses. The start of the ABPM session is defined as the first valid measurement. No further editing of the values will be performed beyond the values flagged as invalid by the ABPM machine.

The following ABPM derived measurements will be evaluated:

- Mean 24-hour SBP, DBP, and HR
- Mean daytime SBP, DBP, and HR. For analysis of daytime recordings, only sessions with >14 valid measurements during daytime (0700 to 2200 hours) will be included.
- Mean nighttime SBP, DBP, and HR. For analysis of nighttime recordings, only sessions with >7 valid measurements during nighttime (2200 to 0700 hours, inclusive) will be included.
- Dipper and nondipper response. Blood pressure normally has a circadian pattern in which BP drops at night and is higher during the awake hours. This is referred to as dipping. The following equation (Anwar and White 2001) will be used to assess dipping status:

(mean daytime SBP – mean nighttime SBP) / (mean daytime SBP) × 100%

Participants will be classified into the following categories based on their dipping status:

  - dipper: ≥10% nondipper: 0% to 10%, and
  - riser (reverse dipper): <0%

- Mean 24-hour, daytime, and nighttime BP load. Blood pressure load will be defined as follows (Ernst and Bergus 2002):
  - daytime: percent of SBP readings >140 mmHg and percent of DBP readings >90 mmHg (reported separately)
  - nighttime: percent of SBP readings >120 mmHg and percent of DBP readings >80 mmHg (reported separately)
  - mean 24-hour: percent above systolic limits (daytime and nighttime combined) and percent above diastolic limits (daytime and nighttime combined)
- Mean 24-hour, daytime, and nighttime pulse pressure (PP)
  - PP = SBP – DBP
- Mean 24-hour, daytime, and nighttime mean arterial pressure (MAP)
  - MAP = (2 × diastolic + systolic) / 3
- Treatment-emergent abnormally high HR (based on mean 24-hour values) as defined in Table GZGI.4.4.
- Treatment-emergent abnormally high SBP and DBP defined as following (adapted from Pickering et al. 2005):
  - Mean Daytime SBP/DBP: >140/90 mmHg
  - Mean Nighttime SBP/DBP: >125/75 mmHg
  - Mean 24-hour SBP/DBP: >135/85 mmHg

The daytime (0700 to 2200 hours) and nighttime (2200 to 0700 hours, inclusive) definitions will be used to provide accurate estimates of the BPs and HRs during the awake and sleeping periods.

Mean plots for each derived continuous measurement above by treatment group by visit may be created.

Summary statistics for actual and change from baseline for each derived measurements above will be performed by treatment group and by visit.

Blood pressure load is an indicator of hypertensive burden and has been shown to correlate with markers of CV morbidity in participants with hypertension and has been used as a method of assessing antihypertensive drug efficacy. Summary statistics of 24-hour BP load for all participants and those with or without hypertension at baseline will be calculated at baseline and at 36 weeks.

In addition to these derived measurements, the following summary measurements that provide information on the individual ABPM recordings sessions will be assessed:

- total duration of recording session
- total duration of recording session in the first 24 hours
- total number of measurements in the first 24 hours
- number of valid measurements in the first 24 hours, and
- percent of valid measurements in the first 24 hours.

For continuous outcomes collected from the ABPM, including BP and HR, the analysis will be conducted using an ANCOVA model, with terms of treatment group (LY1, LY2, LY3, LY4, and placebo), baseline measurement, BMI stratum (<35, >=35) and Sex using Safety Analysis Set (SS).

4.6.7.10. Hypersensitivity Events

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

Summaries of all potential hypersensitivity reactions will be generated by PT with decreasing frequency by treatment group. The AE database will be searched using pre-defined SMQs to identify events consistent with hypersensitivity events. Detailed searching criteria for hypersensitivity events can be found in Appendix 6 (Section 7.6). 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 will be summarized and compared by treatment group using Fisher's exact test.

4.6.7.11. Major Depressive Disorder/Suicidal Ideation

The major depressive disorder/suicidal ideation or behavior will be captured as AESIs. Adverse events will be searched using MedDRA PT terms. Detailed searching criteria can be found in Appendix 6 (Section 7.6).

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

Additionally, suicidal ideation and behavior, and depression will be assessed by the investigator via spontaneously reported AEs.

4.6.7.12. Abuse Liability

The counts and percentages of participants with treatment emergent potential abuse liability events and treatment-emergent drug abuse, dependence and withdrawal will be summarized by treatment group with decreasing frequency. Detailed searching criteria can be found in Section 7.6.

4.7. Other Analyses

4.7.1. Subgroup Analyses

Subgroup analyses of the primary endpoint (percentage change in body weight) will be made to assess consistency of the intervention effect across the following subgroups using the "efficacy estimand":

- Age group: <65 vs ≥65 years

- Sex: female vs male
- Baseline BMI (≤35 kg/m^2, >35 kg/m^2)
- Prediabetes Status (Baseline HbA1c <5.7% vs ≥5.7%)
- Race
- Ethnicity
- Country/Region

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

For body weight and percentage change from baseline in body weight, for each subgroup analyses aforementioned, the following 2 models will be conducted:

- Conduct MMRM model on the subgroup only with terms of treatment group, visit, treatment group-by-visit-interaction, strata, and baseline as a covariate. Variance-covariance structure for within-participant errors will be the same as Section 4.1.
- Full MMRM model: treatment group, visit, subgroup, treatment group-by-visit-interaction, treatment-by-subgroup-interaction, subgroup-by-visit-interaction, treatment-visit-subgroup-interaction, strata as fixed effects, and baseline as a covariate. Variance-covariance structure for within-participant errors will be the same as Section 4.1.

Sex will not be considered as a stratification factor in the analysis for the subgroup of sex and BMI stratum will not be considered as a stratification factor for the subgroup of BMI stratum.

Additional subgroup analyses may also be performed.

4.8. Interim Analyses

An interim efficacy and safety assessment after all participants complete Visit 12 (Week 16) of the treatment period may be conducted to provide information for dose escalation schemes and clinical trial material packaging for future studies. If conducted, an internal Assessment Committee (AC) will be formed to review the interim analyses for the safety and efficacy reports in an unblinded manner. Additional interim analyses may be conducted. Details on the timing of the interim analyses, operational support, and unblinding will be specified in the AC charter and in the study unblinding plan. Neither unblinded information nor results will be shared with study sites or blinded study team members before the study has been unblinded for the final data base lock. Study sites will receive information about interim results only if deemed necessary for the safety of the participants. The trial will not be stopped based on the superiority of LY3502970 versus placebo. Therefore, there will be no inflation of the Type 1 error rate and no need to employ an alpha spending function or multiplicity adjustment.

The primary database lock and primary data analysis for Study GZGI will occur when all participants have completed the Week 26 and the final database lock and final data analysis will occur when all participants have completed the study. Participants and investigators will remain blinded until the completion of the study.

The cancellation or addition of an interim analysis can be determined at any time during the study and will not require a protocol amendment.

Unblinding details are specified in a separate unblinding plan document.

The AC charter will describe the planned interim analyses for AC and the analyses for corresponding endpoints will follow the SAP described analysis method.

4.9. Changes to Protocol-Planned Analyses

- No changes have been made from protocol planned analyses.

5. Sample Size Determination

Approximately 270 participants will be randomly assigned to study drug in a randomization ratio of 5:5:5:3:3:3:3 to Study GZGI, including 50 participants per treatment group to the placebo, LY1, and LY2 groups and 30 participants per group to the LY3-1, LY3-2, LY4-1, and LY4-2 groups.

The sample size calculation is based on the primary efficacy estimand and its end point (percent change from baseline to Week 26 in body weight). Assuming a SD of 6%, a 2-sided alpha level of 0.05, 40 completers for 1 LY3502970 treatment group and 40 completers for the placebo group can provide approximately >90% power to detect a treatment difference of –5% between the LY3502970 treatment group and placebo (LY3502970 – Placebo) in percent change from baseline in body weight at Week 26. Assuming a 20% dropout rate for placebo, LY1, LY2, and 25% dropout rate for LY3 and LY4, 50 participants per group should be randomized for each of the placebo, LY1, and LY2 groups and 60 participants per group should be randomized for each of the LY3 and LY4 groups.

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 the COVID-19 pandemic
- Listing of AEs or deaths related to the COVID-19 pandemic
- Listing of Participants using Extended Visit Window due to COVID-19

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

All demographic and baseline clinical characteristics will be summarized by treatment groups and dose escalation subgroups (placebo, LY1, LY2, LY3-1, LY3-2, LY4-1, LY4-2) for all randomized participants.

The following variables will be included but not limited to: age (years), age groups (<65 and ≥65 years), sex, country, ethnicity, race, height (cm), body weight, HbA1c at baseline (% and mmol/mol), BMI at baseline, BMI group at baseline (≤35 and >35), BMI groups at baseline (≤30 and >30 kg/m^2; <25, ≥25 to <30 kg/m^2, ≥30 to <35 kg/m^2, ≥35 to <40 kg/m^2, and ≥40 kg/m^2), waist circumference (cm) at baseline, fasting serum glucose (mg/dl and mmol/L), eGFR at baseline, eGFR groups (eGFR based on the modified Modification of Diet in Renal Disease [MDRD] equation: ≥90 mL/min/1.73 m^2, <90 and ≥60 mL/min/1.73 m^2, <60 and ≥45 mL/min/1.73 m^2, <45 and ≥30 mL/min/1.73 m^2, and <30 mL/min/1.73 m^2), tobacco use, baseline SBP, and baseline DBP, and baseline pulse rate. A listing of participant demographic and baseline characteristics at study entry will be provided for all randomized participants.

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

The number of randomized participants and the number of randomized participants discontinued per investigator within each country for each treatment group will be summarized. In addition, the number of enrolled participants per investigator within each country will also be summarized.

7.2. Appendix 2: Historical Illnesses and Pre-existing Conditions

The count and percentages of participants with historical illnesses and pre-existing conditions will be summarized by treatment group using the MedDRA PTs nested within SOCs. The SOCs will be in alphabetical order. Conditions (i.e., PTs) will be ordered by decreasing frequency within SOCs. This will be summarized for all randomized participants. Historical illnesses are illnesses that end prior to informed consent and pre-existing 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: Treatment Compliance

If data warrants, the counts and percentages of participants who follow the planned dose escalation scheme (IWRS data), have missed dose of study drug (eCRF data) ≥7 days, or have dose de-escalation (IWRS data) will be summarized for each treatment group. Listings of such participants will also be provided.

Non-compliance is defined as having ≥75% of days of missed doses before permanent study drug discontinuation, will also be summarized by treatment group and dose escalation subgroups.

7.4. Appendix 4: 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 nested within Anatomical Therapeutic Chemical (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
- baseline obesity medication
- changes to baseline medication in post randomization (in term of type/class and dose):
    - antihypertensive therapy, and
    - lipid lowering therapy
    - obesity medication.
- utilization after randomization of:
    - antidiarrheal medication,
    - constipation medication, and
    - antiemetic medication

7.5. Appendix 5: Important Protocol Deviations

Important protocol deviations are identified in the Trial Issues Management Plan (TIMP). A listing of important protocol deviations by treatment groups will be provided at the end of 26-week treatment for primary lock and 36-week treatment for the final lock (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.

Hepatic Treatment-Emergent Adverse Events

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).

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:

1) Supraventricular Arrhythmias:

- For symptoms: Arrhythmia related investigations, signs, and symptoms SMQ (20000051), narrow and broad terms
- For supraventricular arrhythmias: In Cardiac arrhythmia SMQ, under tachyarrhythmia sub SMQ
    - Supraventricular tachyarrhythmia SMQ (20000057), broad and narrow terms
    - Tachyarrhythmia terms, nonspecific SMQ (20000164), narrow terms only; and
    - Ventricular tachyarrhythmia SMQ (20000058), narrow terms only.

2) Cardiac Conduction Disorders

- Conduction defects SMQ (20000056), narrow terms only; and
- Cardiac conduction disorders HLT (10000032), all PTs.

Major Depressive Disorder/Suicidal Ideation

The major depressive disorder/suicidal ideation or behavior will be captured as AESIs. Adverse events will be searched using MedDRA PTs. The PTs from the Depression and suicide/self-injury SMQ as defined in MedDRA (SMQs: 20000037 [Suicide/self-injury] and 20000167 [Depression (excluding suicide and self-injury)]) will be summarized.

Hypersensitivity Events

The hypersensitivity TEAE are characterized as follows:

- Anaphylactic reaction SMQ (20000021; narrow, algorithm per SMQ guide, and broad)
- Hypersensitivity SMQ (20000024; narrow and broad)
- Angioedema SMQ (20000024; narrow and broad)

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

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

Abuse Liability

To identify AE terms suggestive of potential abuse liability, a list of MedDRA preferred terms referred to as Abuse Term list will be used. In addition, narrow and broad terms from SMQ of Drug abuse, dependence and withdrawal (20000101 and 20000102) will be used.

8. References

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

Darken P, Nyberg J, Ballal S, Wright D. The attributable estimand: a new approach to account for intercurrent events. *Pharm Stat*. 2020;19(5):626-635. https://doi.org/10.1002/pst.2019

[EMA] European Medicines Agency. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. Step 5. EMA/CHMP/ICH/436221/2017. Published February 17, 2020. Accessed June 30, 2022. https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-sensitivity-analysis-clinical-trials-guideline-statistical-principles_en.pdf

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

Fu H, Manner D. Bayesian adaptive dose-finding studies with delayed responses. *J Biopharm Stat*. 2010;20(5):1055-70. https://doi.org/10.1080/10543400903315740

Kolotkin RL, Crosby RD. Psychometric evaluation of the impact of weight on quality of life-lite questionnaire (IWQOL-lite) in a community sample. *Qual Life Res*. 2002;11(2):157-71. https://doi.org/10.1023/a:1015081805439

Lipkovich, I, Ratitch, B, Mallinckrodt, CH. Causal Inference and Estimands in Clinical Trials. *Stat. in Biopharm Research*. 2020;12:54-67. https://doi.org/10.1080/19466315.2019.1697739

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

Qu Y, Lipkovich I. Implementation of ICH E9 (R1): a few points learned during the COVID-19 pandemic. *Ther Innov Regul Sci*. 2021;55(5):984-988. https://doi.org/10.1007/s43441-021-00297-6

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

Qu Y, Shurzinske L, Sethuraman S. Defining estimands using a mix of strategies to handle intercurrent events in clinical trials. *Pharm Stat*. 2021;20(2):314-323. https://doi.org/10.1002/pst.2078.

Signature Page for VV-CLIN-074789 v1.0

Approval	PPD PPD 03-Oct-2022 19:19:23 GMT+0000

Signature Page for VV-CLIN-074789 v1.0