This study W8M-MC-LAA1 (NCT06230523) is a sub-study of Master Protocol W8M-MC-CWMM (NCT06143956)

# Statistical Analysis Plan W8M-MC-LAA1 (Version 2.0)

A Phase 2, Parallel-Group, Double-Blind Study to Investigate Weight Management with LY3841136 Compared with Placebo in Adult Participants with Obesity or Overweight

NCT06230523

Approval Date: 26-Nov-2024

# Title Page

**Protocol Title:** A Phase 2 Study of Once-Weekly LY3841136 Compared with Placebo in Participants Who Have Obesity or Are Overweight with Weight-Related Comorbidities

**Protocol Number:** W8M-MC-LAA1

**Compound Number:** LY3841136

**Short Title:** Effect of LY3841136 versus Placebo in Participants Who Have Obesity or Are Overweight

**Sponsor Name:** Eli Lilly and Company

**Legal Registered Address:** Indianapolis, Indiana 46285 USA

**Regulatory Agency Identifier Number(s):**

| Registry | ID |
|---|---|
| IND for master protocol CWMM | 163848 |



**Document ID:** VV-CLIN-151350

**Table of Contents**

Approved on 26 Nov 2024 GMT

## Table of Contents

**Tables** **Page**

## Table of Contents

**Figure** **Page**

## Version history

This statistical analysis plan (SAP) for Study W8M-MC-LAA1 (LAA1) is based on the protocol LAA1(b) dated 19 August 2024.

### SAP Version History Summary

| SAP Version | Approval Date | Change | Rationale |
|---|---|---|---|
| 1.0 | 21 May 2024 | Not Applicable | Original version |
| 2.0 | See date on Page 1 | Major changes include:<br>1. Updating analysis with 9 mg as top dose.<br>2. Updated approach to custom AE searches<br>3. Adding actigraphy analyses<br>4. DXA analyses removed requirement to have at least one dose to be included in analysis; removed initiation of rescue med as ICE.<br>5. Removing dose pooling approach as main analysis method.<br>6. Indicating fasting visits to be used for efficacy analyses<br>7. Definition of 1 dose as 1 week of study treatment<br>8. Removing strata factor variable and including separate strata factors (not including DXA participation) in MMRM and ANCOVA models<br>9. Indicating the populations used for certain AE analysis<br>10. Specifying: shift tables for Health Outcome measures; orthostatic analysis of vitals | 1. To be in line with protocol amendment (b)<br>2. Introduced medical review process<br>3. Actigraphy included here instead of separate SAP<br>4. Clarify DXA analysis<br>5. Titration scheme better suited to MMRM instead of cMMRM<br>6. Clarification on model approach<br>7. Not previously specified.<br>8. Model preference<br>9. Highlighting changes from the CWMM-SAP<br>10. Were not previously specified |

# 1. Introduction

Study LAA1 is an intervention specific appendix (ISA) of the chronic weight management master protocol (CWMM). The CWMM-specific SAP (CWMM-SAP) describes and specifies the analyses relevant to all ISAs. The purpose of this SAP is to describe and specify the analyses specific to the Study LAA1 ISA. As such, analyses for Study LAA1 which have not already been covered by the CWMM-SAP will be described in this document.

## 1.1. Objectives, Endpoints, and Estimands

The CWMM master protocol contains objectives and endpoints that will be evaluated for all study interventions evaluated under the CWMM master protocol.

Study LAA1 includes the following LY3841136-specific features:

- primary analysis endpoint occurs at Week 48
- primary intervention is LY3841136, and
- primary control is placebo.

For clarity in the table below, the primary and secondary objectives and endpoints covered in CWMM-SAP Section 1.1 are repeated here, and secondary objectives and endpoints specific to Study LAA1 have been added. Tertiary objectives and endpoints at the master protocol level are not repeated and can be found in CWMM-SAP Section 1.1; and those specific to Study LAA1 are outlined below.

| Objectives | Endpoints |
|---|---|
| **Primary** | |
| To compare the effect of LY3841136 versus placebo for weight reduction at Week 48 | Percent change from baseline in body weight |
| **Secondary** | |
| To compare the effect of LY3841136 versus placebo for weight reduction at Week 48 | • Change from baseline body weight (kg)<br>• Incidence of participants who achieve<br>  ○ ≥5% body weight reduction<br>  ○ ≥10% body weight reduction<br>• Change from baseline BMI |
| To assess safety and tolerability of study interventions | • TEAEs overall<br>• SAEs<br>• AEs leading to discontinuation<br>• AEs for special safety topics<br>• laboratory parameters<br>• electrocardiogram<br>• vital signs |
| To characterize the PK of LY3841136 | PK parameters AUC and $C_{max}$ |

| Objectives | Endpoints |
|---|---|
| **Tertiary/Exploratory** | |
| To evaluate the development of treatment-emergent antidrug antibodies to LY3841136 | Treatment-emergent antidrug antibodies |
| To compare the effect of LY3841136 versus placebo on change in physical activity from baseline to Weeks 24 and 48 | • Change from baseline in physical activity measured by:<br>o Daily step count<br>o M10 (activity intensity of most active 10 hours during 24-hour period)<br>o MVPA (moderate to vigorous physical activity in minutes during 24-hour period) |
| To compare the effects of LY3841136 versus placebo on health-related quality of life, function, and appetite | • Change from baseline in:<br>o SF-36 v2 acute form domain and summary scores<br>o CoEQ scores<br>o FACT-GP5 score<br>o PGIS – Physical Function due to Weight<br>o PGIC – Physical Function due to Weight<br>o PGIS – Food Craving<br>o PGIC – Food Craving<br>o Patient assessment of GI symptoms |
| To compare the effect of LY3841136 versus placebo on blood glucose regulation from baseline to Week 48 | • Change from baseline in:<br>o Fasting insulin<br>o Fasting glucose<br>o HbA1c<br>o HOMA2-IR<br>o HOMA2-B |
| To compare the effect of LY3841136 versus placebo on mechanistic biomarkers at longitudinal time points from baseline to Week 48 | Change from baseline in longitudinal and endpoint biomarkers |
| To assess the relationships between LY3841136 dose and/or exposure and key safety and efficacy measures, as applicable | Dose-response or exposure-response analyses for key safety and efficacy endpoints, as applicable |

Abbreviations: AE = adverse event; AUC = area under the concentration versus time curve; BMI = body mass index; $C_{max}$ = maximum observed drug concentration; COEQ = The Control of Eating Questionnaire; FACT-GP5 = Functional Assessment of Chronic Illness Therapy - Item GP5; GI = gastrointestinal; HbA1c = glycated hemoglobin A1c; HOMA2-B = homeostasis model of beta-cell function; HOMA2-IR = homeostasis model assessment of insulin resistance; PGIC = Patient's Global Impression of Change; PGIS = Patient Global Impression of Status; PK = pharmacokinetic(s); SAE = serious adverse event; SF-36 v2 acute form = Short Form-36 Version 2 Health Survey acute form; TEAE = treatment-emergent adverse event.

## Primary estimand

Efficacy estimand is used as the primary estimand for the primary objective.

The primary clinical question of interest is:

> *What is the difference in percent change in body weight from baseline at 48 weeks of treatment between LY3841136 and placebo in participants with obesity or overweight as an adjunct to healthy diet and physical activity, in individuals who meet the eligibility criteria if they would remain on their randomly assigned treatment for 48 weeks?*

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

- *Population* – Participants with obesity or overweight who meet the eligibility criteria.
- *Endpoint* – Percent change from baseline to 48 weeks in body weight.
- *Treatment condition* – The randomized treatment with allowance for potential LY3841136 dose interruptions and modification.
- *Population-level summary and treatment effect of interest* – Difference in mean percent changes in body weight from baseline to Week 48 between LY3841136 and Placebo. Intercurrent events (ICEs) include permanent discontinuation of any study drug or initiation of prohibited weight management treatments, which is handled by the hypothetical strategy. Dose modification and interruption for LY3841136 will not be considered as ICEs since dose modification and interruption are part of treatment condition.

***Rationale for efficacy estimand***

This Phase 2 study aims to study the efficacy of LY3841136 under the ideal condition that all participants adhere to their randomly assigned treatment without being confounded by the initiation of other weight management treatments.

**Estimand(s) for secondary objectives**

The efficacy estimand is used as the estimand for the secondary objectives. For endpoints related to weight reduction (including primary, secondary, and tertiary), a second estimand of "treatment regimen" may be evaluated.

**Supplemental estimand: Treatment regimen**

The clinical question of interest is:

> *What is the treatment difference in the percent change in body weight from baseline at 48 weeks between LY3841136 and placebo, as an adjunct to healthy diet and physical activity, in individuals who meet the eligibility criteria regardless of adherence to study intervention?*

***Treatment regimen estimand attributes***

- *Population* – Individuals who meet the eligibility criteria.
- *Endpoint* – Percent change in body weight from baseline to 48 weeks.
- *Treatment condition* - The randomized treatment with allowance for potential dose interruptions and modifications regardless of adherence to study intervention or initiation of prohibited weight management treatments.

- *Intercurrent event*: no ICEs since treatment adherence and the initiation of prohibited weight management treatments are part of the treatment condition.
- *Population-level summary and treatment effect of interest* – Difference in mean percent changes in body weight from baseline to Week 48 between LY3841136 group and placebo group.

***Rationale for treatment regimen estimand***

This estimand aims to evaluate the efficacy of LY3841136 that reflects the real-life behavior of the target population.

## 1.2. Study Design

Study LAA1 is a 48-week, Phase 2, multicenter, randomized, double-blind, placebo-controlled study designed to examine the safety and efficacy of across a range of dose levels of subcutaneous once weekly administered LY3841136 compared with subcutaneous once weekly administered placebo in participants with obesity or overweight who have not been diagnosed with any form of diabetes mellitus except for a prior diagnosis of gestational diabetes mellitus. The primary objective will be the effect of LY3841136 on percent change in body weight at 48 weeks. Four maintenance (top) doses of LY3841136 (1 mg, 3 mg, 6 mg, and 9 mg) will be evaluated in the trial across six LY3841136 treatment groups: 1 mg, 3 mg, 6 mg, 9 mg, 6/9 mg, and 3/6/9 mg.

Dose escalation to improve gastrointestinal tolerability will occur in 2 treatment groups where 9 mg is the maintenance dose. The first escalation treatment group will begin with 6 mg and escalate to 9 mg at week 20. The second dose escalation group will begin with 3 mg and escalate every 4 weeks to reach 9 mg LY3841136 (or placebo) at Week 8.

The study sample size approximately of 250 participants will be randomly assigned to placebo (approximately 50 participants), 1 mg LY3841136 (approximately 25 participants), 3 mg LY3841136 (approximately 25 participants), 6 mg LY3841136 (approximately 25 participants), 9 mg LY3841136 (approximately 50 participants), 6/9 mg LY3841136 (approximately 25 participants), or 3/6/9 mg LY3841136 (approximately 50 participants). To achieve this approximate allocation per arm of 2:1:1:1:2:1:2 as described previously for the overall study, participants enrolling under protocol amendment (a) or later will be randomized in a ratio of 2:1:1:1:2:3:2 to placebo, 1 mg, 3mg, 6 mg, 9 mg, 6/9 mg, or 3/6/9 mg.

Randomization will be stratified by body mass index (BMI) (<35, ≥35 $kg/m^2$ at screening), sex (female, male), and Dual-Energy X-Ray Absorptiometry (DXA) substudy participation (Y/N).

Approximately 70% enrollment of females will be used to ensure a sufficiently large sample of males.

This study includes a 6-week screening/lead-in period, a 48-week, double-blind treatment period, and a 10-week posttreatment follow-up period.

Participants will self-administer study intervention subcutaneously once weekly.

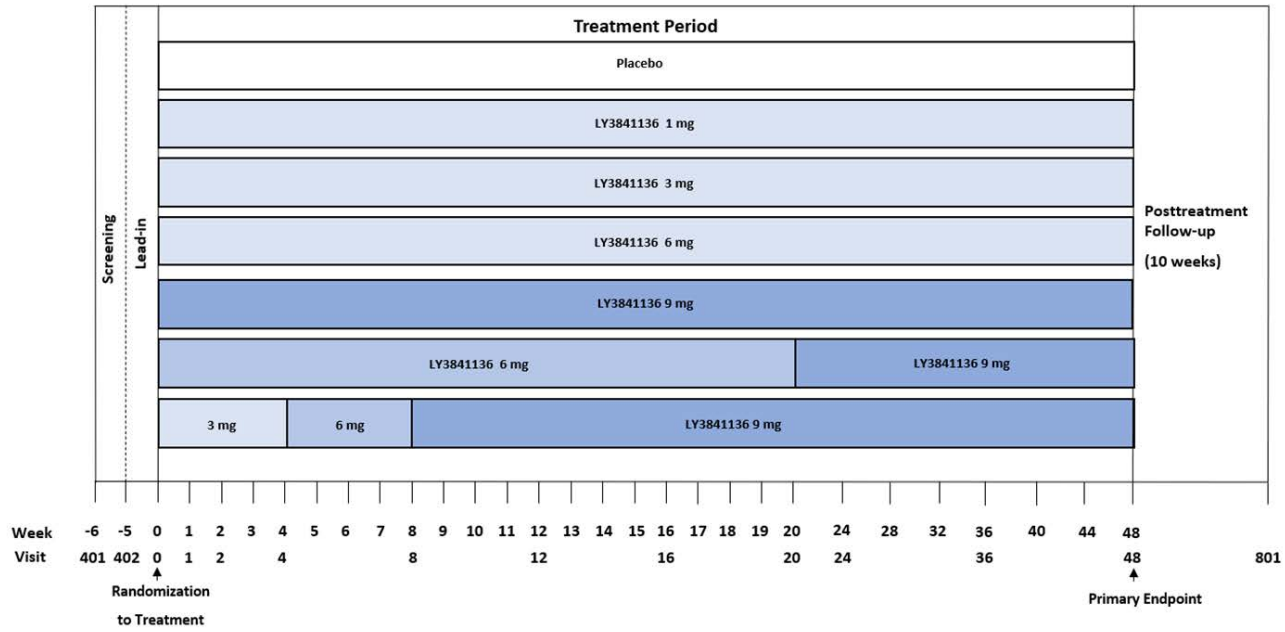


**Figure LAA1.1.1. Illustration of study design for clinical protocol W8M-MC-LAA1.**

# 2. Statistical Hypothesis

The primary null hypothesis is that there is no difference between LY3841136 and placebo on mean percent change from baseline in body weight at Week 48.

The null hypotheses for the secondary endpoints is that there is no difference between LY3841136 and placebo at Week 48 with respect to:

- mean change from baseline body weight (kg)
- proportion of participants who achieve
  - ≥5% body weight reduction
  - ≥10% body weight reduction
- mean change from baseline BMI

The null hypotheses will be tested on the following treatment comparisons:

- 1 mg LY3841136 vs placebo
- 3 mg LY3841136 vs placebo
- 6 mg LY3841136 vs placebo
- 9 mg LY3841136 vs placebo
- 6/9 mg LY3841136 vs placebo
- 3/6/9 mg LY3841136 vs placebo
- Pooled 9 mg (6/9 mg, 3/6/9 mg, or 9 mg) LY3841136 vs placebo

## 2.1. Multiplicity Adjustment

No adjustment for multiplicity will be performed.



Approved on 26 Nov 2024 GMT

# 3. Analysis Sets

The CWMM-SAP Section 3 includes the following definitions of the Participant Analysis Sets and Data Points Sets. It is copied here for ease of reference within this document.

| Participant Analysis Set | Description |
|---|---|
| Entered | All participants who signed informed consent. |
| Full analysis set (FAS) | All randomized participants. |
| Safety analysis set (SAS) | All participants who are exposed to study intervention. |

The following data points sets are defined:

| Data Points Sets | Description |
|---|---|
| DPS1 | Data points obtained during treatment period at or after randomization up to the date of discontinuation of the study drug. |
| DPS2 | Data points obtained during treatment period at or after randomization regardless of discontinuation of study intervention. |
| DPS3 | Data points obtained during treatment period plus safety follow-up period regardless of adherence to study drug. |

Abbreviation: DPS = data point set.

Unless otherwise specified:

- Full analysis set (FAS) and data point set 1 (DPS1) are used to estimate the primary estimand (efficacy estimand) for primary and secondary objectives.
- FAS and data point set 2 (DPS2) are used to estimate the additional treatment regimen estimand for the primary and secondary objectives.
- Safety analysis set and data point set 3 (DPS3) are used to present safety data.

For both efficacy analyses and safety analyses, participants will be included in the analyses according to the planned investigational intervention.

# 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.

Please see the CWMM-SAP Section 4.1 for general considerations, including:

- Change from baseline and percent change from baseline calculations
- summarization and reporting of continuous data and categorical data
- modeling of continuous longitudinal data and cross-sectional data, and binary data
- alpha level of statistical tests and confidence intervals
- summary statistics
- end of study participation
- Definition of Baseline
- Analysis Methods for Efficacy Estimand
- Analysis Methods for Safety
- Analyses for Supplemental Estimands
- Borrowing of Placebo Data (Note: The borrowing of placebo data is not planned for this ISA.)

Unless stated otherwise, statistical summaries and analyses will be conducted based on planned randomized treatment group:

- placebo
- 1 mg LY3841136
- 3 mg LY3841136
- 6 mg LY3841136
- 9 mg LY3841136
- 6/9 mg LY3841136, or
- 3/6/9 mg LY3841136

regardless of the actual treatment(s) received by the participant due to any dose modification.

The primary estimand (a precise definition of the treatment effect to be estimated) of interest in comparing efficacy of LY3841136 doses with placebo is the “efficacy estimand”.

The post-baseline values that will be included in the model for the efficacy analyses will be from the treatment period visits designated as fasting by the protocol SoA, that is, Visits 0, 4, 8, 12, 16, 20, 24, 36, 48. Values will be included from fasting visits regardless of whether the subject was truly fasting.

For modeling purposes where stratification variables are indicated as independent variables, participation in the DXA substudy will not be included as a stratification variable. The stratification variables (as factors) in the model will be the body mass index (BMI) (<35, ≥35 $kg/m^2$ at screening), and sex (female, male).

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 according to the estimand strategy.

The summary statistics for continuous measures will include sample size, mean, standard deviation (SD), median, minimum, and maximum. The analysis model to make comparisons among treatment groups relative to continuous measurements assessed over time (in addition to the baseline and end of treatment measurements) will be a mixed model for repeated measures (MMRM). Analysis of covariance (ANCOVA) or analysis of heterogeneous covariance (ANHECOVA) (Ye et al. 2022) may be used to make comparisons among treatment groups for continuous measurements with only 1 postbaseline assessment. Least squares means (LSM) 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 LSMs and the 95% confidence intervals for the treatment differences, along with the p-values for the treatment comparisons. All baseline measures will be analyzed using an analysis of variance model that has treatment group as the model terms.

Kaplan-Meier method will be used for estimation of cumulative event-free survival rates over time, and Cox proportional hazards regression analysis will be used to compare hazards rates among treatments.

Statistical treatment comparisons will only be performed between LY3841136 and placebo. Comparisons among LY3841136 doses will not be performed unless otherwise specified.

Not all analyses described in this SAP will necessarily be included in the Clinical Study Report (CSR). Any analysis described in this SAP and not provided in the CSR would be available upon request.

### 4.1.1. Definition of Baseline

Baseline for efficacy assessments is defined as the last available non-missing measurement prior to the first dosing of study intervention, which in most cases will be the measurement recorded at Week 0. If there are no doses of study intervention, baseline will be defined as the last available non-missing measure on or prior to randomization. In cases where the measurement is taken on the same day (where time is not collected or not reliable) as the first injection, this measurement will be used as the baseline values for data analysis.

Baseline period for safety assessment starts at screening and ends prior to the first dose (typically at Week 0).

## 4.2. Participant Dispositions

Please refer to CWMM-SAP Section 4.2.

## 4.3. Primary Endpoint Analysis

### 4.3.1. Definition of Endpoint

The primary efficacy measure is percent change in body weight at primary timepoint. The percent change in body weight is defined as:

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

### 4.3.2. Main Analytical Approach

The primary endpoint analysis using the FAS and DPS1 will be done for Week 48.

The primary efficacy assessment, guided by the "efficacy estimand," will be conducted where the hypothetical strategy is used to handle the ICE (permanent discontinuation of study intervention), so only data collected before the occurrence of any such ICEs will be used in the MMRM estimation (Section 4.1). Through the MMRM, the potential efficacy measures (after the ICEs) had participants not experienced ICEs will be implicitly imputed.

The strata variables sex (male, female) and baseline BMI category (<35, ≥35 kg/$m^2$ at screening) will be included in the model as factor variables. DXA substudy participation (Yes/No) will not be included in the model.

The MMRM will be used to analyze continuous measurements at postbaseline visits guided by efficacy estimand.

The MMRM model will include the following terms:

1. Baseline value (of the dependent variable)
2. Treatment group (factor variable)
3. Visit (factor variable)
4. Sex (factor variable)
5. Baseline BMI category (factor variable)
6. Two-way Interactions:
    a. Visit by Treatment
    b. Visit by Baseline BMI category
    c. Visit by Sex
    d. Visit by Baseline

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

1. heterogeneous Toeplitz
2. heterogeneous first-order autoregressive
3. heterogeneous compound symmetry
4. Toeplitz
5. first-order autoregressive, and
6. compound symmetry.

The first covariance structure that converges will be used.

If the full model fails to converge for all covariance structures tested, the model will be refit by removing one or more two-way interaction terms and then cycling through the covariance structures.

The sandwich estimator (Diggle et al., 1994) will be applied to provide robust standard errors accounting for potential misspecification of the covariance structure. The associated two-sided 95% confidence interval (CI) and corresponding p-values will also be reported. LSMeans will be reported which reflect the proportion of the categorical variable levels observed in the data (i.e. observed margins) rather than assuming equal proportions among the levels within the categorical variables.

To confirm efficacy of LY3841136 with adequate statistical power, the evaluation of the efficacy endpoints for the arms 9 mg, 6/9 mg and 3/6/9 mg combined (e.g. as 9 mg pooled) may be compared with placebo.

Additionally, a dose-pooling, constrained variant of MMRM (cMMRM) may be explored for the primary endpoint. This analysis replaces the fixed effect of treatment group with that of the preplanned time-varying treatment doses at each time point in an effort to derive a more efficient estimator when the same treatment regimen over time is shared across the treatment groups due to preplanned titration (Qu et al. 2022).

For exploratory efficacy endpoints with multiple post-baseline timepoints, the MMRM model used in the main analysis will be fit; however, due to the exploratory nature of these analyses the model may be simplified as necessary for the analysis. Additionally, for endpoints with very few post-baseline timepoints an ANCOVA/ANHECOVA may be used instead of MMRM.

### 4.3.3. Analysis Methods for Safety

See CWMM-SAP Section 4.1.3. for more complete details.

The MMRM analysis will be used for selected continuous safety data with multiple postbaseline measurements. The MMRM model will include the following terms:

1. Baseline value (of the dependent variable)
2. Treatment group (factor variable)
3. Visit (factor variable)
4. Two-way Interactions:
    a. Visit by Treatment
    b. Visit by Baseline

Alternatively, the model may be fit by removing the Visit by Baseline term.

### 4.3.4. Analysis of Supplemental Estimand

The treatment regimen estimand will also be evaluated. See CWMM-SAP Section 4.1.4. The ANCOVA model may be used to analyze continuous measurements guided by treatment regimen estimand. The model will include:

1. Baseline value (of the dependent variable)
2. Treatment group (factor variable)
3. Sex (factor variable)
4. Baseline BMI category (factor variable)
5. Two-way Interactions:
    a. treatment by baseline BMI category
    b. treatment by Sex
    c. treatment by Baseline

## 4.4. Secondary Endpoints Analysis

See CWMM-SAP Section 4.4 for the analysis of secondary endpoints.

### 4.4.1. Logistic Regression for Binary Endpoints

Binary outcomes will be analyzed with the following procedure:

1) Assuming missing at random (MAR), multiple imputation will be used to impute the missing continuous-valued measurements at scheduled visits, utilizing the randomized participants with efficacy estimand data points. The imputation is based on non-missing data from the same treatment group at the same visit.
2) At the visit of interest, convert the observed and imputed continuous values into binary values. For example, transform the continuous body weight value at a visit into a binary value based on whether the corresponding percent change from baseline meets a specified threshold.
3) Fit a logistic regression model to the transformed binary data considering the follow terms for inclusion in the model:
    - Baseline value (of the dependent variable, continuous).
    - Treatment group (factor variable)
    - Sex (factor variable)
    - Baseline BMI category (factor variable)
4) Derive the final inference using Rubin's Rule by combining estimates from the multiple imputed datasets.

### 4.4.2. Pharmacokinetic and Pharmacokinetic/Pharmacodynamic Methods

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

LY3841136 concentration-time data will be summarized in the clinical study report. Dose and exposure-response analyses between LY3841136 dose and concentration and key safety, tolerability, and efficacy endpoints may be explored graphically or performed using population PK and population PK/PD approaches implemented in the 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. Tertiary/Exploratory Endpoints Analysis

Unless otherwise specified, the following exploratory endpoints will be analyzed on the FAS and DPS1 dataset guided by the "efficacy estimand."

See CWMM-SAP Section 4.4 for the analysis of select exploratory endpoints not covered here. Table LAA1.4.1 outlines analyses for exploratory endpoints.

**Table LAA1.4.1. Exploratory Endpoints**

| Measure and Endpoint | Analysis | Additional Information |
|---|---|---|
| Development of treatment-emergent antidrug antibodies to LY3841136 | See Section 4.5.4 | |
| Change from baseline in physical activity measured by:<br>• daily step count<br>• M10 (activity intensity of most active 10 hours during 24-hour period)<br>• MVPA (moderate to vigorous physical activity in minutes during 24-hour period) | See section 4.5.1. See section 4.5.1. See section 4.5.1. See section 4.5.1. | |
| Mean change in<br>• fasting glucose<br>• fasting insulin<br>• fasting C-peptide<br>• HOMA2-IR<br>• HbA1c<br>• HOMA2-B<br>•<br>• Plasma renin activity<br>• Aldosterone<br>• Sodium<br>• PotassiumPotassium<br>• Fasting lipid panel:<br>  ○ total cholesterol<br>  ○ LDL cholesterol (LDL-c) – LDL direct measure if triglycerides >400 mg/dL<br>  ○ HDL cholesterol (HDL-c)<br>  ○ non-HDL cholesterol (non HDL-c)<br>  ○ triglycerides<br>  ○ VLDL cholesterol (VLDL-c)<br>• Calcium<br>• Phosphate<br>• Calcitonin | MMRM model in Section 4.3 | LSM estimates through 48 weeks will be plotted by study treatment.<br><br>Log transformation of the response variables may be conducted. |
| Mean change in<br>• total adiponectin<br>• total IGF-1<br>• leptin<br>• PTH<br>• 25-OH vitamin D<br>• CTX-1 | MMRM model in Section 4.3. MMRM model in Section 4.3. MMRM model in Section 4.3. May revert to ANCOVA/ANHECOVA as necessary. | LSM estimates through 48 weeks will be plotted by study treatment.<br><br>Log transformation of the response variables may be conducted. |



Approved on 26 Nov 2024 GMT

| Measure and Endpoint | Analysis | Additional Information |
|---|---|---|
| • P1NP | | |
| Proportion of participants with change in<br>• PGIC - Physical Function due to Weight<br>• PGIS - Food Craving<br>• PGIC - Food Craving<br>• PGIS - Physical Function due to Weight<br>• Patient assessment of GI symptoms | See Section 4.5.2. | |
| Mean change in<br>• SF-36 v2 acute form domain scores | See Section 4.5.2.1. | |
| Mean change in<br>• CoEQ scores | See Section 4.5.2.2. | |
| Mean change in<br>• FACT-GP5 score | See Section 4.5.2.3. | |
| Dose-response or exposure-response analyses for key safety and efficacy endpoints, as applicable | See Section 4.5.3. | |

Abbreviations: ANHECOVA = analysis of heterogeneous covariance; BMI = body mass index; BP = blood pressure; CoEQ = The Control of Eating Questionnaire; CTX-1 = collagen cross-linked C-telopeptide; DPS = data point set; FACT-GP5 = Functional Assessment of Chronic Illness Therapy - Item GP5; GI = gastrointestinal; HbA1c = glycated hemoglobin A1c; HDL = high-density lipoprotein; HOMA2-B = homeostasis model of beta-cell function; HOMA2-IR = homeostasis model assessment of insulin resistance; IGF-1 = insulin-like growth factor 1; LDL = low-density lipoprotein; LSM = least squares mean; MMRM = mixed model for repeated measures; P1NP = procollagen 1 Intact N-terminal propeptide; PGIC = Patient's Global Impression of Change; PGIS = Patient Global Impression of Status; PTH = parathyroid hormone; SAS = safety analysis set; SF-36 v2 = Short Form 36 version 2 Health Survey; VLDL = very low-density lipoprotein.

### 4.5.1. Actigraphy

The Axivity AX6 (actigraphy data capture tool) will be utilized to objectively assess various physical activity parameters including but not limited to change in daytime physical activity from baseline.

A wrist worn AX6 device will be dispensed to participants and placed on the non-dominant wrist at Visit 402 to collect activity data for at least 2 weeks before being returned to the site at Visit 0. The AX6 is to be worn continuously for 2 weeks during study Visits 402 to 0, Weeks 22 to 24, and Weeks 46 to 48.

The exploratory endpoints related to actigraphy are to compare the effect of LY3841136 versus placebo in change from baseline to Weeks 24 and 48 in physical activity as measured by:

- Daily step count
- M10 (activity intensity of most active 10 hours during 24-hour period)
- MVPA (moderate to vigorous physical activity in minutes during 24-hour period)

For each subject and endpoint, the data will be preprocessed by averaging across the measurements within the interval to provide a single value per timepoint for Baseline (Visits 402 to 0); Week 24 (Weeks 22-24) and Week 48 (Weeks 46-48). At least 3 valid days within the two-

week interval are needed to derive the corresponding timepoint value. Participants without baseline data will be excluded from actigraphy analyses.

Change from baseline to Weeks 24 and 48 in physical activity will be analyzed according to the efficacy estimand using an MMRM model and as detailed in Section 4.3.2. If a subject discontinues treatment during the interval of days comprising a timepoint, then the timepoint will not be included in the analysis.

Descriptive statistics will be prepared on physical activity measures (for example, mean, median, and standard deviation) for baseline and each postbaseline measurement per treatment group.

### 4.5.2. Health Outcomes

The patient-reported outcome questionnaires will be analyzed using the FAS and DPS1, unless specified otherwise.

Statistical significance will not be assessed for shift tables.

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

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

#### 4.5.2.1. Short-Form-36 Health Survey Version 2, Acute Form

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

- Mental Component Summary Score
- Physical Component Summary 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 above parameter, the raw scores will be transformed into the domain scores (t-scores) and the following analyses for the actual value and change from baseline value will be conducted on FAS and DPS1:

- descriptive summaries by treatment group, and
- analysis from MMRM (as described in Section 4.3).

#### 4.5.2.2. Control of Eating Questionnaire

The Control of Eating Questionnaire (Dalton et al. 2015) is a 19-item participant-completed questionnaire that assesses the intensity of food cravings, food types craved, appetite, and mood

over the past 7 days. The 11-point numeric rating scale ranges from 0 (not at all) to 10 (extremely) for each of the 19 items.

Seventeen of the 19 items are grouped into the 4 domains of:

- craving control
- positive mood
- craving for savory foods, and
- craving for sweet foods.

The remaining 2 items, which assess the degree of hunger and fullness, are scored individually.

Higher scores represent higher levels of the concept measured in each domain or individual item (for the 2 standalone items).

The 4 subscales should be calculated as:

- Craving Control
  - calculate average of items 9, 10, 11, 12 and 19
  - reverse the final subscale score so that a greater score represents greater levels of Craving Control
- Craving for Sweet
  - calculate average of items 3, 13, 14 and 15
- Craving for Savory
  - calculate average of items 4, 16, 17 and 18
- Positive Mood
  - reverse item 6, and
  - calculate average of items 5, 7, 8 and 6 (reversed).

The following analyses for the actual value and change from baseline value for each domain and for the hunger and fullness items will be conducted on FAS and DPS1:

- descriptive summaries by treatment group, and
- analysis from MMRM (as described in Section 4.3).

#### 4.5.2.3. Functional Assessment of Chronic Illness Therapy – Item GP5

This is a single, participant-completed item from the Functional Assessment of Chronic Illness Therapy general scale that assesses the degree to which side-effects of treatment are bothersome. Response options are on a 5-point Likert scale ranging from “not at all” to “very much.”

Sustained overall side-effect burden is defined as a Functional Assessment of Chronic Illness Therapy – Item GP5 score of 3 or 4 (“quite a bit” or “very much” bothered) at 3 or more subsequently measured postbaseline assessments or treatment discontinuation due to adverse events (AEs) or death on treatment due to AEs. Missing assessments are allowed within a sequence of the three assessments of score of 3 or 4 to consider side-effect burden as sustained, but no score of 0, 1, or 2 is allowed within the sequence.

Actual and percent change from baseline will be analyzed with MMRM (as described in Section 4.3).

Time to sustained side-effect burden may be analyzed. Time to sustained overall side-effect burden as measured by Functional Assessment of Chronic Illness Therapy – Item GP5 is defined as the time from the first dose of study treatment until the earliest date of the series of Functional Assessment of Chronic Illness Therapy – Item GP5 scores meeting the criteria for sustained overall side-effect burden defined above or the date of treatment discontinuation or the date of death due to AE (whichever is earlier).

#### 4.5.2.4. Patient Global Impression of Change – Physical Function due to Weight

The Patient Global Impression of Change – Physical Function Weight questionnaire assesses participants' overall perception of the efficacy of treatment. This is a single global item that asks participants to rate the overall change in their ability to perform physical activities due to their weight since starting the study medication. The responses are based on a 5-point scale ranging from "much better" to "much worse."

Shift tables will be provided to analyze the proportion of participants with a change in category from baseline. The number and percentage of patients in each category will be summarized by visit.

#### 4.5.2.5. Patient Global Impression of Severity – Physical Function due to Weight

The Patient Global Impression of Severity – Physical Function Weight questionnaire is designed to assess the participants' overall perception of their condition. This is a single global item that asks participants to rate how their weight limited their ability to perform physical activities in the past 7 days on a 5-point scale ranging from "not at all limited" to "extremely limited."

Shift tables will be provided to analyze the proportion of participants with a change in category from baseline. The number and percentage of patients in each category will be summarized by visit.

#### 4.5.2.6. Patient Global Impression of Severity – Food Craving

The Patient Global Impression of Severity – Food Craving item is a participant-completed, single-item measure that assesses overall perception of food craving in the past 7 days on a 5-point scale:

- much higher than usual
- somewhat higher than usual
- same as usual
- somewhat lower than usual, and
- much lower than usual.

Shift tables will be provided to analyze the proportion of participants with a change in category from baseline. The number and percentage of patients in each category will be summarized by visit.

#### 4.5.2.7. Patient Global Impression of Change – Food Craving

The Patient Global Impression of Change Food Craving item is a patient-reported, single-item measure that assesses participants' change in their food craving since they started taking the study medication. The item is rated on a 5-point scale:

- increased a lot
- increased a little
- no change
- decreased a little, and
- decreased a lot.

Shift tables will be provided to analyze the proportion of participants with a change in category from baseline. The number and percentage of patients in each category will be summarized by visit.

#### 4.5.2.8. Patient Assessment of Gastrointestinal Symptoms

Three questions will be used to assess the gastrointestinal symptoms of nausea, vomiting, and diarrhea. These single-item, participant-completed measures how often nausea, vomiting, and diarrhea occurred in the past 7 days. Response options are on a 5-point scale ranging from "0 days" to "every day."

Shift tables will be provided to analyze the proportion of participants with a change in category from baseline. The number and percentage of patients in each category will be summarized by visit.

### 4.5.3. Bayesian Analyses for Dose-Response

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

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

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

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

where $d$ is the maximum duration of the treatment period in weeks ($d = 48$ for primary analysis), $\lambda(\theta)$ is the dose-response function for the maximum response at dose $\theta$ and $k(\theta)$ is dose $\theta's$ rate parameter. This formulation of the mean function $f(t;\theta)$ was introduced by Fu and Manner (2010) to characterize the change from baseline over time in a continuous outcome that

could be approximated with a pattern of exponential decay. It assumes a monotone time profile with the maximum effect reached at time $d$. The longitudinal data $Y_{\theta,t_c,jt}$ for participant $j$ at time $t$ with titration scheme $(\theta, t_c)$ will be fitted by adding the error terms to the mean function $f_{\theta,t_c}(t)$ where

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

$$s_j \sim N(0, \sigma_s^2) \text{ and } \epsilon_{jt} \sim N(0, \sigma^2)$$

are independent, denoting between-subject variation and within-subject variation, respectively. Given

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

For characterizing the body weight change, we use a different formulation of $f_{\theta,t_c}(t)$ that showed better fitting in the analyses of historical trials and simulation studies while keeping all other elements unchanged. In the new formulation,

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

Where

$$h(t - t_{ci}) = \frac{\left(1-e^{-\tau(t-t_{ci})}\right)}{1-e^{-\tau d}}.$$

The parameter $\tau$ controls the rate of change in the body weight when dose changes.

The dose-maximum response function $\lambda(\theta)$ 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, eg, Simple Normal Dynamic Linear Modeling. 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 noninformative priors for the hyperparameters in the model as follows:

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

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

As this is an exploratory analysis, the method may be adjusted as necessary for analysis.

### 4.5.4. Immunogenicity

Treatment-emergent antidrug antibodies (TE-ADAs) are defined as:

- those with a titer 2-fold (1 dilution) greater than the minimum required dilution if no ADAs were detected at baseline (treatment-induced ADA), or
- those with a 4-fold (2 dilutions) increase in titer compared to baseline if ADAs were detected at baseline (treatment-boosted ADA).

A patient is evaluable for TE-ADA if the patient has a non-missing baseline ADA result, and at least 1 nonmissing postbaseline ADA result.

Listings of patients who are not TE-ADA evaluable, patients with at least 1 test having detected LY3841136 ADAs, and patients having LY3841136 ADAs present or treatment-emergent adverse event: hypersensitivity reactions or injection site reactions will be provided.

The frequency and percentage of patients with preexisting ADA, with TE-ADA, with cross-reactive antibodies and with neutralizing antibodies will be tabulated by dose (if data warrant), where proportions are relative to the number of patients who are TE-ADA evaluable. If cross-reactivity to native amylin or neutralizing antibodies against native amylin assays are performed, the frequency of each may be reported. The frequency and percentage of patients with hypersensitivity and injection site reaction treatment-emergent adverse events by TE-ADA status will be tabulated if data warrant.

For the TE-ADA-positive participants, the distribution of maximum titers may be described.

The relationship between the presence of antibodies and the PK parameters and PD response including safety and efficacy to LY3841136 may be assessed.

## 4.6. Safety Analyses

All safety analyses will be conducted on the SAS and DPS3 unless otherwise specified.

### 4.6.1. Extent of Exposure

Due to the weekly frequency of dosing administration, for analysis purposes one injection will be considered 7 days of treatment.

Listing of exposure to LY3841136 and placebo will be provided by treatment group. Summary of duration of follow-up (defined as time in days from date of randomization to the date of the last study visit) and/or duration on study treatment (defined as time in days from date of first dose of study treatment to date of last dose of study treatment plus 7 days) will be provided by treatment group, in the following period:

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

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

- greater than 0
- 4 weeks or longer
- 8 weeks or longer
- 12 weeks or longer
- 16 weeks or longer
- 20 weeks or longer
- 24 weeks or longer
- 28 weeks or longer
- 32 weeks or longer
- 36 weeks or longer
- 40 weeks or longer
- 44 weeks or longer, and
- 48 weeks or longer.

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

- 0 weeks
- longer than 0 weeks to less than 4 weeks
- 4 weeks or longer to less than 8 weeks
- 8 weeks or longer to less than 12 weeks
- 12 weeks or longer to less than 16 weeks
- 16 weeks or longer to less than 20 weeks
- 20 weeks or longer to less than 24 weeks
- 24 weeks or longer to less than 28 weeks
- 28 weeks or longer to less than 32 weeks
- 32 weeks or longer to less than 36 weeks
- 36 weeks or longer to less than 40 weeks
- 40 weeks or longer to less than 44 weeks, and
- 44 weeks or longer to less than 48 weeks.

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

### 4.6.2. Adverse Events

AEs will be coded from the actual term using the Medical Dictionary for Regulatory Activities (MedDRA) and reported with preferred terms and system organ class. A TEAE is defined as an event that first occurred or worsened in severity after baseline. The baseline period for safety assessment starts at screening and ends prior to the first dose (typically at Week 0). 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 CRF-collected information (e.g., treatment emergent flag, start time of study treatment and event) will be used to determine whether the event was pre versus posttreatment if available. If the relevant information is not available, then the events will be counted as posttreatment.

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.

The following analyses will be performed according to the Analysis/Data Point sets which supersede those outlined in the CWMM-SAP.

| Analysis | Method | Analysis/Data Point Sets |
|---|---|---|
| Overview of AEs, including<br>• TEAE<br>• SAE<br>• death, and<br>• discontinuation from study drug due to an AE. | Fisher's exact | SAS + DPS3 |
| TEAEs by PT within SOC | Fisher's exact | SAS + DPS3 |
| Maximum severity TEAEs by PT | Fisher's exact | SAS + DPS3 |
| Common TEAEs by PT | Fisher's exact | SAS + DPS3 |
| SAEs by PT within SOC | Fisher's exact | SAS + DPS3 |
| SAEs by PT | Fisher's exact | SAS + DPS3 |
| Primary AEs leading to permanent discontinuation of study drug by PT within SOC | Fisher's exact | SAS + DPS3 |
| AEs leading to study treatment interruption by PT within SOC | Fisher's exact | SAS + DPS3 |
| Listing of SAEs | | Entered |
| Listings of AEs | | Entered |
| Listing of primary AEs leading to permanent discontinuation of study drug | | SAS + DPS3 |
| Listing of deaths | | Entered |
| Narratives for participants with at least 1 notable event | | SAS + DPS3 |

Abbreviations: AE = adverse event; DPS3 = data point set 3; FAS = full analysis set; MedDRA = Medical Dictionary for Regulatory Activities; PT = MedDRA Preferred Term; SAE = serious adverse event; SOC = MedDRA System Organ Class; TEAE = treatment-emergent adverse event.

Note: For events that are sex-specific, the denominator and computation of the percentage will include only participants from the given sex.

### 4.6.3. Safety Laboratory Measures

Please see the analyses in CWMM-SAP Section 4.5.3.

### 4.6.4. Vital Signs and Physical Characteristics

Please see the CWMM-SAP Section 4.5.4 for a description of the analyses.

Additionally, vital signs will be assessed for orthostatic hypotension and orthostatic tachycardia. The proportion of subjects meeting the criteria in Table LAA1.4.2 will be summarized by treatment and visit. A summary of the change between supine and standing measurements will be provided by treatment and visit.

**Table LAA1.4.2. Criteria for Orthostatic Vital Sign Analyses**

| | |
|---|---|
| Orthostatic Hypotension (systolic) | Decrease of systolic blood pressure of >= 20 mm Hg between supine and standing |
| Orthostatic Hypotension (diastolic) | Decrease of diastolic blood pressure of >= 10 mm Hg between supine and standing |
| Orthostatic pulse (tachycardia) | Increase in pulse rate of >= 30 bpm between supine and standing |

### 4.6.5. Electrocardiograms

Please see the CWMM-SAP Section 4.5.5 for a description of the analyses. Patient Narratives

### 4.6.6. Adverse Events of Special Interest (AESI) and Other Safety Topics

All analyses will be done as outlined in the CWMM-SAP Section 4.5.6 Adverse Events of Special Interest (AESI) and Other Safety Topics.

Per the LAA1 protocol, the following AEs of special interest are defined:

- hypotension and related neurological signs and symptoms
- acute renal failure
- suicidal ideation and behavior, and
- depression.

#### 4.6.6.1. Custom Searches for Special topics

The following additional custom searches for LAA1 will be performed and will be identified through a Medical Dictionary for Regulatory Activities (MedDRA) search. The AEs that are included in the MedDRA search will be identified through a review and categorization of each TEAE occurring in the study by the blinded medical team (CRP and/or GPS representative).

Records of the search criteria used will be kept in the CLUWE directory:

"\prd\ly900038\w8m_mc_laa1\common\documentation\LAA1_MedDRA_Search"

The search criteria document may be updated due to MedDRA version update or evolving criteria of event or topic, and it will be reflected in the document's version history.

##### *4.6.6.1.1. Mood-related Events*

Treatment-emergent events considered to be related to mood will be summarized and listed.

The mood-related events search will be done for each of two categorizations: a 'Focused Mood' search for AEs that are considered directly or very closely related to mood, and an 'Expanded Mood' search for AEs considered directly, very closely related or potentially related to mood based on the judgment of the medical review.

##### *4.6.6.1.2. Fatigue Search*

Treatment-emergent events considered to be related to fatigue will be summarized and listed.

##### *4.6.6.1.3. Bradycardia/Sinus Bradycardia*

Treatment-emergent events considered to be related to bradycardia or sinus bradycardia will be summarized and listed.

##### *4.6.6.1.4. Abuse potential*

Treatment-emergent events considered to be related to study agent abuse potential will be summarized and listed.

##### *4.6.6.1.5. Agitation, Sleeplessness, or Anxiety*

Treatment-emergent events considered which are considered related to agitation, sleeplessness, or anxiety will be summarized and listed.

#### *4.6.6.1.6. Dehydration and Electrolytes Imbalance*

Treatment-emergent events considered to be related to dehydration and/or an electrolyte imbalance will be summarized and listed.

## 4.7. Other Analyses

### 4.7.1. Subgroup Analyses

Subgroup analyses of the primary endpoint will be made to assess consistency of the intervention effect across the following subgroups:

- age group: younger than 65 versus 65 years or older
- race
- sex: female versus male
- ethnicity
- BMI ($kg/m^2$) group: less than 30 versus greater than or equal to 30 and less than 35 versus greater than or equal to 35 and less than 40 versus greater than or equal to 40, and
- baseline BMI (the minimum or greater to less than the first quartile, the first quartile or greater to less than the median, the median or greater to less than the third quartile, the third quartile or greater to the maximum).

For each subgroup analysis, the MMRM model as described in Section 4.3 will be used.

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

## 4.8. Interim Analyses

An interim analysis for Study LAA1 will be conducted approximately when 100% of the participants complete the Week 24 visit. This interim analysis will evaluate both safety and efficacy profile of LY3841136 and may be used to determine the doses of future studies of LY3841136. Another interim analysis may be conducted approximately when 100% of the participants complete the 36-week treatment period.

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

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

# 5. Sample Size Determination

The study sample size approximately of 250 participants will be randomly assigned to placebo (approximately 50 participants), 1 mg LY3841136 (approximately 25 participants), 3 mg LY3841136 (approximately 25 participants), 6 mg LY3841136 (approximately 25 participants), 9 mg LY3841136 (approximately 50 participants), 6/9 mg LY3841136 (approximately 25 participants), or 3/6/9 mg LY3841136 (approximately 50 participants). To achieve this approximate allocation per arm of 2:1:1:1:2:1:2 as described previously for the overall study, participants enrolling under protocol amendment (a) or later will be randomized in a ratio of 2:1:1:1:2:3:2 to placebo, 1 mg, 3mg, 6 mg, 9 mg, 6/9 mg, or 3/6/9 mg.

Approximately 70% enrollment of females will be used to ensure a sufficiently large sample of males.

Sample size selection is guided by the objective of establishing superiority of each LY3841136 dose to placebo relative to the percent change from baseline in body weight at Week 48. Assuming a difference of 7.8% between the LY3841136 and placebo group means of percent change from baseline in body weight, and a standard deviation of 10%, a sample size of 50 per arm enrolled in the 3/6/9 mg and 9 mg LY3841136 arms provides over 90% power, and 25 per arm enrolled in the 1 mg, 3 mg, 6 mg, and 6/9 mg LY3841136 arms will provide over 80% power using 2-sample t-test at a 2-sided significance level of 0.05.



Approved on 26 Nov 2024 GMT

# 6. Supporting Documentation

## 6.1. Appendix 1: Demographics, Baseline Characteristics, Medical History, Concomitant Medications, Treatment Compliance, Important Protocol Deviations

Details are in CWMM-SAP Appendices for the following analyses:

- Demographic and Baseline Characteristics
- Historical Illnesses and Preexisting Conditions
- Concomitant Medications
- Treatment Compliance
- Important Protocol Deviations

## 6.2. Appendix 2: Dual-Energy X-Ray Absorptiometry

A substudy using DXA to assess body composition will be performed in a subset of Study LAA1 participants. The exploratory objective is to compare the effect of LY3841136 versus placebo on change in body composition from baseline to Week 48. The exploratory endpoints include:

- percent change from baseline to Week 48 in total fat mass, as measured by DXA
- percent change from baseline to Week 48 in body weight
- percent change from baseline to Week 48 in total lean mass, as measured by DXA
- percent change from baseline to Week 48 in visceral fat mass, as measured by DXA
- absolute and percent change from baseline to Week 48 in bone mineral density, as measured by DXA
- ratio of Android Fat mass to Gynoid Fat mass by visit, and
- absolute and percent change from baseline to Week 48 in appendicular lean mass.

Additional endpoints may also be assessed, such as the Fat Loss Index (FLI), and the actual and change from baseline in Fat Mass Index (FMI), defined as:

- FLI = *(DXA Change from Baseline in Body Fat [kg]) / (Change from Baseline in Body Weight [kg])x100*
- FMI = *DXA Body Fat [kg] / Height [m]^2*

Baseline clinical characteristics and demographic variables (including but not limited to total fat mass, total lean mass, visceral fat mass, systolic blood pressure and diastolic blood pressure, heart rate) will be summarized by treatment for the participants who are enrolled in the addendum.

Unless otherwise specified, all analyses for the variables measured or derived from DXA will be conducted on participants who are enrolled in this substudy, randomly assigned to study treatment, and have a baseline measurement. There will be no multiplicity adjustments.

MMRM as in Section 4.3 will be used on FAS (randomized participants in the substudy) and DPS1 for the absolute and percent change from baseline analyses. Since DXA is performed for only two post baseline timepoints, the model may be changed to ANCOVA if necessary. The LSM of the treatment difference estimated from the model and 2-sided 95% confidence intervals will be provided.

The Ratio of Android Fat mass to Gynoid Fat mass will be evaluated by fitting an MMRM for Android Fat mass and for Gynoid Fat mass separately and comparing the ratio of the LSMean estimate for Android Fat mass to the LSMean estimate for Gynoid Fat mass at the corresponding visits. Confidence intervals may be provided by using the delta method to approximate the variance. Analysis of FLI would be similarly performed.

## 6.3. Appendix 3: Qualitative Interviews

A substudy using qualitative interviews to assess participants' treatment experiences will be performed in a subset of Study LAA1 participants.

Qualitative data from the interviews will be extracted from interview transcripts. These data will be analyzed using a qualitative data analysis program that allows for qualitative data (participant's quotes) to be systematically organized, coded, and compared. Analysis of coded interview data results will be detailed in a separate document.

# 7. References

Dalton M, Finlayson G, Hill A, et al. Preliminary validation and principal components analysis of the Control of Eating Questionnaire (CoEQ) for the experience of food craving. *Eur J Clin Nutr*. 2015;69(12):1313-1317. https://doi.org/10.1038/ejcn.2015.57

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

Ma C, Shen X, Qu Y, Du Y. Analysis of an incomplete binary outcome dichotomized from an underlying continuous variable in clinical trials. *Pharm Stat*. 2022;21(5):907-918. https://doi.org/10.1002/pst.2204

Qu Y, Cui X, Bray R, Sun N. Efficient estimation of the efficacy and safety endpoints for clinical trials with preplanned dose titrations. *Stat Biopharm Res*. 2022;1(15):214-224 https://doi.org/10.1080/19466315.2022.2025893

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

Ye T, Shao J, Yi Y, Zhao Q. Toward better practice of covariate adjustment in analyzing randomized clinical trials. *J Am Stat Assoc*. 2022;544(118):2370-2382. https://doi.org/10.1080/01621459.2022.2049278

Signature Page for VV-CLIN-151350 v2.0

| | |
|---|---|
| Approval | PPD<br>26-Nov-2024 19:38:19 GMT+0000 |

Signature Page for VV-CLIN-151350 v2.0