Statistical Analysis Plan: W8M-MC-OXA1 (Version 3.0)

A Phase 2, Parallel-Group, Double-Blind, 4-Arm Study to Investigate Weight Management with LY3305677 Compared with Placebo in Adult Participants with Obesity or Overweight

NCT06124807

Approval Date: 30-May-2025

Title Page

Master Protocol Title: A Master Protocol for a Randomized, Controlled, Clinical Trial of Multiple Interventions for Chronic Weight Management in Adult Participants with Obesity or Overweight

ISA Title: A Phase 2, Parallel-Group, Double-Blind, 4-Arm Study to Investigate Weight Management with LY3305677 Compared with Placebo in Adult Participants with Obesity or Overweight

ISA Number: W8M-MC-OXA1

Compound Number: LY3305677

Short Title: A Study to Investigate Weight Management with LY3305677 Compared with Placebo in Adult Participants with Obesity or Overweight

Acronym: OXA1

Sponsor Name: Eli Lilly and Company

Legal Registered Address: Indianapolis, Indiana USA 46285

Regulatory Agency Identifier Number(s)

Registry	ID
IND for CWMM master protocol	163848
EU trial number	2022-502816-35-00

Document ID: VV-CLIN-128213

Table of Contents

List of Tables

Tables **Page**

List of Figures

Figure **Page**

List of Abbreviations

Term	Definition
ADA	antidrug antibodies
AE	adverse event
AESI	adverse event of special interest
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
ANHECOVA	analysis of heterogeneous covariance
ANOVA	analysis of variance
AST	aspartate aminotransferase
BMI	body mass index
CI	confidence interval
cMMRM	constrained mixed model for repeated measures
CRF	case report form
CSR	clinical study report
CWMM	Chronic Weight Management Master
DBL	direct bilirubin
DPS	data point set
ECG	electrocardiogram
ED	early discontinuation
ELF	enhanced liver fibrosis
FAS	full analysis set
FLI	fatty liver index
GGT	gamma-glutamyl transferase
GI	gastrointestinal
HbA1c	glycated hemoglobin

Approved on 30 May 2025 GMT

HFF	hepatic fat fraction
HOMA2-B	updated (1998) homeostasis model assessment of beta-cell function
HOMA2-IR	updated (1998) homeostatic model assessment for insulin resistance
ICE	intercurrent event
ISA	intervention-specific appendix
KM	Kaplan-Meier
LFC	liver fat content
Lilly	Eli Lilly and Company
LLT	lowest-level term
LSM	least squares mean
MAR	missing at random
MedDRA	Medical Dictionary for Regulatory Activities
MI	multiple imputation
MMRM	mixed model repeated measures
MRI-PDFF	magnetic resonance imaging – proton density fat fraction
NASH	non-alcoholic steatohepatitis
NIS4	NASH Index Score 4
NONMEM	nonlinear mixed-effects modeling
PD	pharmacodynamic(s)
PGIC	Patient Global Impression of Change
PGIS	Patient Global Impression of Severity
PK	pharmacokinetic(s)
PROMIS	Patient-Reported Outcome Measurement Information System
PT	MedDRA preferred term
QW	weekly
REML	restricted maximum likelihood
SAE	serious adverse event
SAP	Statistical Analysis Plan

SAS	safety analysis set
SD	standard deviation
SF-36	36-item Short Form Survey
SOC	MedDRA system organ class
TBL	total bilirubin
TE-ADA	treatment-emergent antidrug antibodies
TEAE	treatment-emergment adverse event
UACR	urine albumin-to-creatinine ratio
VAS	visual analog scale

Version history

This Statistical Analysis Plan (SAP) for Study W8M-MC-OXA1 (OXA1) is based on the protocol dated 28 Oct 2024.

SAP Version History Summary

SAP Version	Approval Date	Change	Brief Rationale
3	See date on page 1	In Section 4.2.4, added baseline HbA1c (<5.7%, ≥5.7%) as a possible factor variable to models of post-baseline hypoglycemia incidence and rate.	FDA feedback
		Section 4.9.1 is amended to state that subgroup analyses by race will include at least the 5 race categories specified by FDA guidance. Because some race subgroups are expected to be less than 10% of the study population, only summary statistics will be provided for these analyses.	FDA feedback
2	26 Feb 2025	Removed initiation of prohibited weight management treatments in estimand bullet in Section 1.1.1.1 as this was not in protocol	Deletion
		In Section 1.2, added that we will refer to arm associated with 3/6 mg dosing strategy as LY3305677 3/6 mg	Clarification
		In Section 3, modified Safety Analysis Set (SAS) to include the following underlined text, “All participants who are	Correction

		exposed to at least one dose of study intervention"	
		In Section 3, added a new application of FAS+DPS2 to include the following underlined text, "FAS and DPS2 are used to estimate the supplemental (treatment regimen) estimand for the primary objective and select secondary objectives."	To incorporate FDA feedback
		Combined BMI and sex into a single stratification factor in Sections 4.1 and 4.2	For modeling purposes
		To the end of Section 4.2.4, the following was added for the negative binomial model, "Due to the limited number of expected hypoglycemia events at baseline in this population, baseline will not be included as a covariate in the model."	Clarifying
		Modified in-text references for HOMA to HOMA2	Correction
		Added "In situations where higher-order interactions are included in the model, main effects may be dropped to simplify the model statement without any impact to the point estimates or inference" to Section 4.2.1	For modeling purposes
		Removed specification of Satterthwaite approximation for degrees of freedom adjustment in Section 4.2.1	Deletion for modeling flexibility

		Added G-computation procedure to logistic regression in Section 4.2.2	To align with FDA guidance
		Expanded exploratory endpoints related to body composition measures in Table OXA1.4.2 to be more explicit about what adipose tissue volumes, muscle volumes, and other parameters would be summarized.	Clarification
		Deleted, "Any exploratory endpoints using MMRM and ANCOVA in the CWMM SAP that are not specified in the below table should be updated to cMMRM and ANHECOVA." from Section 4.7	Correction
		Deleted text in Table OXA1.4.2 noting that actigraphy measures would be described in a separate analysis plan. Replaced with Section 4.7.1.	Modification
		Added calculations for fatty liver index (FLI) and body composition measures to Table OXA1.4.3.	Define calculations for analyses
		In Table OXA1.4.3, modified average for systolic and diastolic blood pressure to involve two measurements (instead of three).	Correction
		Added the following to Section 4.7, "Additionally, for endpoints with very few post-baseline timepoints an ANCOVA/ANHECOVA	Addition for reliable model-based estimates

Approved on 30 May 2025 GMT

		may be used instead of MMRM."	
		Added Section 4.7.1 describing actigraphy analysis plan	Addition
		Modified extent of study treatment exposure in Section 4.8.1 by planned treatment group (from 8-16 weeks to 8-12 and 12-16 weeks) to coincide with dose escalations	Correction
		In Section 4.10, excluded antidrug antibodies from secondary efficacy objectives evaluated at database locks and interim analyses.	Antidrug antibodies are not available until final database lock
		Minor formatting changes and clarifying text added throughout	Editorial and clarifications
1	28 Oct 2024	Not Applicable	Original version

1. Introduction

Study OXA1 is an intervention specific appendix (ISA) of the chronic weight management master (CWMM) protocol. The CWMM-specific SAP (CWMM-SAP) describes and specifies the analyses relevant to all ISAs. The purpose of this Statistical Analysis Plan (SAP) is to describe and specify the analyses relevant in ISA W8M-MC-OXA1 under the CWMM protocol in addition to the CWMM-SAP. The OXA1-specific SAP (OXA1-SAP) includes the additional secondary and exploratory analysis specific to this ISA.

Changes to the protocol-planned analyses are described in Section 4.11.

1.1. Objectives, Endpoints, and Estimands

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

Study OXA1 includes the following LY3305677-specific features:

- Primary analysis endpoint occurs at Week 32,
- Primary intervention is LY3305677, and
- Primary control is placebo.

For clarity in Table OXA1.1.1, the primary and secondary objectives and endpoints covered in CWMM-SAP Section 1.1 are repeated here. In addition to the primary, secondary, and exploratory objectives and endpoints stated in the CWMM protocol, secondary and exploratory objectives (and their respective endpoints specific for OXA1) are also listed in this table. Exploratory objectives and endpoints at the master protocol level are not repeated and can be found in CWMM-SAP Section 1.1.

Table OXA1.1.1. Objectives for Clinical Protocol W8M-MC-OXA1

Objectives	Endpoints	Analysis Methods
Primary		
To compare the effect of LY3305677 versus placebo for weight reduction at the primary analysis time point (Week 32)	Percent change from baseline in body weight (%)	• cMMRM, efficacy estimand, Section 4.2.1 • ANHECOVA, treatment regimen estimand, Section 4.3.1
Secondary		
To compare the effect of LY3305677 versus placebo for weight reduction at Week 48	Percent change from baseline in body weight (%)	• cMMRM, efficacy estimand, Section 4.2.1
To compare the effect of LY3305677 versus placebo for weight reduction at the primary analysis time point (Week 32) and at Week 48	Change from baseline in body weight (kg)	• cMMRM, efficacy estimand, Section 4.2.1
	Incidence of: • ≥5% body weight reduction • ≥10% body weight reduction	• Logistic regression, efficacy estimand, Section 4.2.2
	Change from baseline in BMI	• cMMRM, efficacy estimand, Section 4.2.1
To assess safety and tolerability of study interventions	• TEAEs overall • SAEs • AEs leading to discontinuation • AEs for special safety topics • Laboratory parameters • Electrocardiogram • Vital signs	• Section 4.2.4 • Table OXA1.4.4 • CWMM-SAP Tables CWMM.4.5 - CWMM.4.10

Objectives	Endpoints	Analysis Methods
To evaluate the development of treatment-emergent antidrug antibodies to LY3305677	Treatment-emergent antidrug antibodies (TE-ADA)	• Section 4.6.2
To characterize the pharmacokinetics (PK) of LY3305677	PK parameters AUC and Cmax	• Details will be specified in a separate PK/PD analysis plan. See Section 4.6.1 for further details.
To compare the effect of LY3305677 versus placebo on liver fat content (LFC) as measured by MRI-PDFF at Week 32 and Week 48	For participants with baseline LFC ≥5% and for participants with baseline LFC ≥10%: • Change and percent change from baseline in LFC • Incidence of ≥30% relative reduction in LFC	• Change and percent change: ○ Primary data lock: ANHECOVA, efficacy estimand, Section 4.2.3 ○ Final data lock: MMRM, efficacy estimand, Section 4.2.1 • Proportion: logistic regression, efficacy estimand, Section 4.2.2
Exploratory (in addition to exploratory objectives listed in the CWMM-SAP)		
To compare the effect of LY3305677 versus placebo on BP and heart rate at Week 32 and Week 48	Change from baseline in: • 24-hour mean systolic BP (mmHg) measured by ABPM • 24-hour mean diastolic BP (mmHg) measured by ABPM • 24-hour mean heart rate (beats/min) measured by ABPM	• Section 4.7
To compare the effect of LY3305677 versus placebo on physical activity at Week 32 and Week 48	Change from baseline in physical activity measured by: • Daily steps count	• Section 4.7

Objectives	Endpoints	Analysis Methods
	• M10 (activity intensity of most active 10 hours during 24-hour period) • moderate-to-vigorous physical activity (MVPA) in minutes during 24-hour period	
To compare the effects of LY3305677 versus placebo on HRQoL, function, and appetite at Week 32 and Week 48	Change from baseline in: • PROMIS SF Pain Interference 4a v1.1 • PGIS - Physical Function due to Weight • PGIC - Physical Function due to Weight • Domains of SF-36 v2, Acute • Appetite VAS	• Section 4.7
To compare the effect of LY3305677 versus placebo on glucose regulation at Week 32 and Week 48	Change from baseline in: • Fasting insulin • Fasting glucose • HbA1c • HOMA2-IR • HOMA2-B	• Section 4.7
To compare the effect of LY3305677 versus placebo on biomarkers at longitudinal time points and at Week 32 and Week 48	Change from baseline in longitudinal and endpoint biomarkers	• Section 4.7
To assess the relationships between LY3305677 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	• Section 4.7

Objectives	Endpoints	Analysis Methods
To compare the effect of LY3305677 versus placebo on body composition as measured by MRI at Week 32 and Week 48	Change from baseline in: • Adipose tissue volumes • Muscle volumes • Additional MRI-derived measures may be explored	• Section 4.7
To compare the effect of LY3305677 versus placebo on liver fat content (LFC) as measured by MRI-PDFF at Week 32 and Week 48	For participants with baseline LFC ≥5% and for participants with baseline LFC ≥10%: • Incidence of LFC < 5% • Incidence of ≥50% relative reduction in LFC	• Section 4.7

Abbreviations: ABPM = ambulatory blood pressure monitoring; ADA = antidrug antibodies; AEs = adverse events; AUC = area under the concentration versus time curve; BMI = body mass index; BP = blood pressure; C_{max} = maximum observed drug concentration; CWMM = chronic weight management mater protocol; HbA1c = glycated hemoglobin A1c; HOMA2-B = updated (1998) homeostasis model assessment of beta-cell function; HOMA2-IR = updated (1998) homeostasis model assessment of insulin resistance; HRQoL = health-related quality of life; LFC = liver fat content; MRI = magnetic resonance imaging; MRI-PDFF = magnetic resonance imaging proton density fat fraction; PD = pharmacodynamic(s); PGIC = Patient Global Impression of Change; PGIS = Patient Global Impression of Severity; PK = pharmacokinetic(s); PROMIS SF = Patient-Reported Outcome Measurement Information System Short Form; SAEs = serious adverse events; SAP = statistical analysis plan; SF-36 v2 = Short Form-36 Version 2; TE = treatment-emergent; TEAEs = treatment-emergent adverse events; VAS = visual analog scale.

1.1.1. Estimands for Primary Objective

1.1.1.1. Primary Estimand: Efficacy

An efficacy estimand is used as the primary estimand for the primary objective.

The primary clinical question of interest is:

> *What is the treatment difference in percent change in body weight from baseline to 32 weeks between LY3305677 and placebo in participants with obesity or overweight who meet the eligibility criteria if they would remain on their randomly assigned treatment for 32 weeks?*

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

- *Population* – participants with obesity or overweight who meet the eligibility criteria.
- *Endpoint* – percent change in body weight from baseline to 32 weeks.
- *Treatment condition* – the randomized treatment, as an adjunct to healthy diet and physical activity, with allowance for dose modification based on gastrointestinal (GI) tolerability, unless otherwise specified in the OXA1 protocol.
- *Intercurrent events* – Intercurrent events (ICEs) include permanent discontinuation of any study drug, which is handled by the hypothetical strategy. Dose modification and interruption will not be considered as ICEs because dose modification and interruption are part of treatment condition. Down-titration will not be considered as intercurrent events for the estimand definition, unless otherwise specified in OXA1 protocol.
- *Population-level summary and treatment effect of interest* – Difference in mean percent change in body weight from baseline to Week 32 between LY3305677 and placebo.

1.1.1.2. Rationale for Efficacy Estimand

This Phase 2 trial aims to study the efficacy of LY3305677 under the ideal condition that all participants adhere to the randomized treatment without being confounded by the initiation of other weight management treatment.

1.1.2. Estimand(s) for Secondary Objectives

The efficacy estimand is used as the estimand for the secondary objectives.

1.1.2.1. Supplemental Estimand: Treatment Regimen

An alternative estimand, the treatment regimen estimand which is guided by the treatment policy strategy, will be used to conduct supplementary analysis on the primary and select secondary objectives.

In the treatment regimen estimand, the primary clinical question of interest is:

> *What is the treatment difference in percent change in body weight from baseline to 32 weeks between LY3305677 and placebo in participants with obesity or overweight who meet the eligibility criteria regardless of their adherence to study intervention.*

The treatment regimen estimand is described by the following attributes:

- *Population* – participants with obesity or overweight who meet the eligibility criteria.
- *Measure* – percent change in body weight from baseline to 32 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 events* – no ICEs because 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 change in body weight from baseline to Week 32 between LY3305677 and placebo.

1.1.2.2. Rationale for Treatment Regimen Estimand

This treatment regimen estimand aims to evaluate the efficacy of LY3305677 that reflects the real-life behavior of the target population.

1.2. Study Design

OXA1 is a Phase 2, parallel-group, double-blind, four-arm study to investigate weight management with LY3305677 subcutaneously administered once weekly compared with placebo. For convenience, the functional combination of the CWMM master protocol plus the OXA1 ISA will be simply described as "the study."

The study consists of a 6-week screening/lead-in period, a 48-week treatment period, and a posttreatment follow-up period of approximately 8 weeks after the last visit in the treatment period.

The primary endpoint will be at Week 32.

The study schema is presented in Figure OXA1.1.1. Throughout the OXA1 SAP, we will refer to the arm associated with the 3/6 mg dosing strategy as LY3305677 3/6 mg and the other arms by their maximum planned dose.

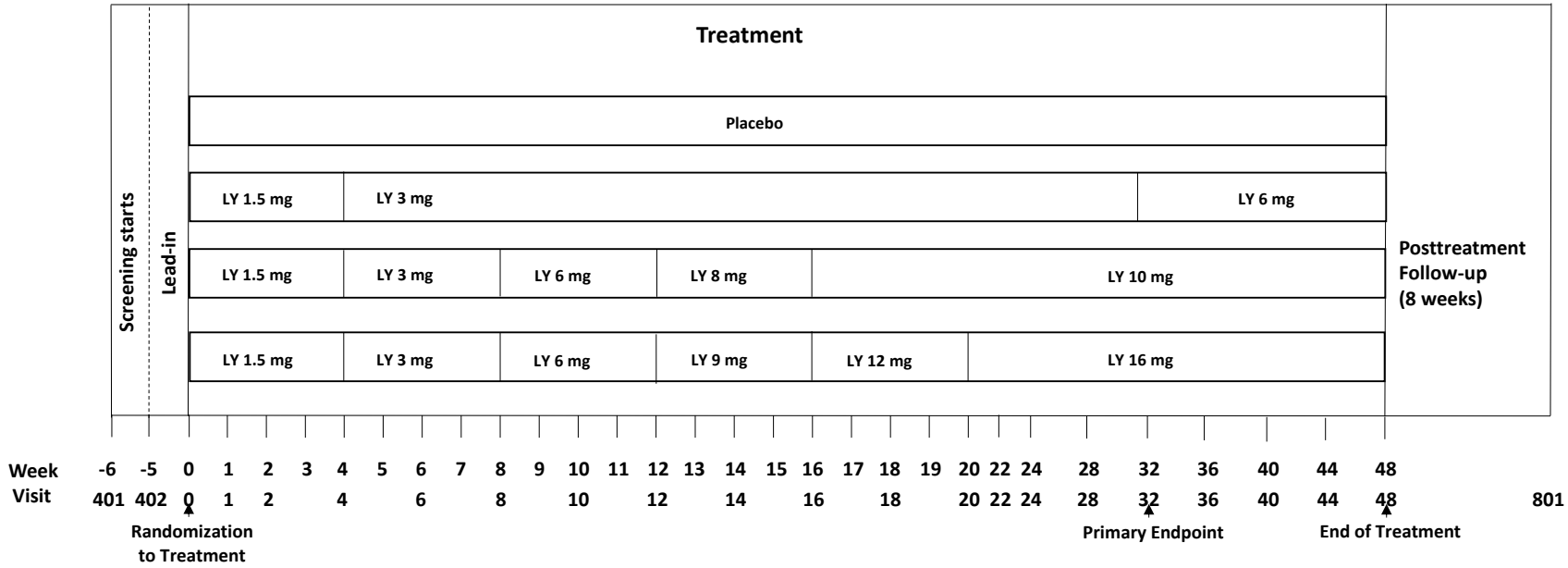

Figure OXA1.1.1. Illustration of study design for Clinical Protocol W8M-MC-OXA1.

2. Statistical Hypotheses

The primary null hypothesis is that there is no difference between LY3305677 and placebo in percent change in body weight between baseline and Week 32. Thus, the null hypothesis will be tested on the following treatment comparisons:

- QW LY3305677 3/6 mg versus placebo,
- QW LY3305677 10 mg versus placebo, and
- QW LY3305677 16 mg versus placebo.

The secondary null hypotheses are that there is no difference between LY3305677 and placebo with respect to:

- mean percent change from baseline in body weight (%) at Week 48,
- mean change from baseline in body weight (kg) at Week 32 and Week 48,
- proportion of participants who achieve:
  - ≥5% body weight reduction at Week 32 and Week 48,
  - ≥10% body weight reduction at Week 32 and Week 48,
- mean change from baseline in BMI at Week 32 and Week 48,
- for participants with baseline LFC ≥5% and for participants with baseline LFC ≥10%:
  - mean change and mean percent change from baseline in liver fat content by MRI-PDFF at Week 32 and Week 48, and
  - proportion of participants who achieve ≥30% reduction in liver fat content by MRI-PDFF at Week 32 and Week 48.

2.1. Multiplicity Adjustment

No adjustment for multiplicity will be performed.

Approved on 30 May 2025 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 at least one dose of study intervention.

The following data point sets are defined:

Data Point Sets	Description
DPS1	Data points obtained during treatment period at or after randomization up to the date of discontinuation of the study intervention.
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 intervention.

Abbreviations: DPS = data point set.

FAS and DPS1 are used to estimate the primary (efficacy) estimand for the primary and secondary objectives.

FAS and DPS2 are used to estimate the supplemental (treatment regimen) estimand for the primary objective and select secondary objectives.

SAS and DPS3 are used to present safety data.

For the efficacy analyses, participants will be included in the analyses according to the planned investigational intervention, whereas, for safety analyses, participants will be included in the analyses according to the investigational intervention they actually received.

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 not summarized here. Unless otherwise stated, statistical summaries and analyses will be conducted based on planned randomized treatment group:

- placebo,
- 3/6 mg LY3305677,
- 10 mg LY3305677, and
- 16 mg LY3305677.

The randomization of this study is stratified by sex (male, female) and BMI category at screening (≤ 30, >30 kg/m^2). A strata variable is defined for statistical modeling to consist of 4 joint levels of stratification factors:

- Male, BMI ≤ 30 kg/m^2
- Male, BMI >30 kg/m^2
- Female, BMI ≤ 30 kg/m^2, and
- Female, BMI >30 kg/m^2.

This strata variable will be referred to as "strata" through the remainder of the SAP.

Summary statistics for continuous measures will include sample size, mean, standard deviation (SD), median, minimum, and maximum values. To compare treatment groups for continuous measurements assessed over time, a mixed model for repeated measures (MMRM) or constrained mixed model for repeated measures (cMMRM) (Qu et al. 2023) will be utilized. For continuous measurements with only one postbaseline assessment, comparisons among treatment groups may be made using analysis of covariance (ANCOVA) or analysis of heterogeneous covariance (ANHECOVA) (Ye et al. 2023). Least squares means (LSMs) and standard errors derived from these analysis models will be presented for changes from baseline. Treatment comparisons will display the differences in LSMs, along with 95% confidence intervals (CIs) and p-values for these comparisons.

Data may exist at visits where the variable is not scheduled to be collected, because of, for example, early discontinuation (ED) visits. In these situations, data from the ED visit that does not correspond to the planned collection schedule will be excluded from the constrained mixed model for repeated measures (cMMRM) analysis (Andersen 2013).

In addition to the contents in CWMM SAP, any change to the data analysis methods described in OXA1 will require an amendment ONLY if it changes a principal feature of the OXA1 study. Any other change to the data analysis methods described in this study, and the justification for making the change, will be described in this SAP or the clinical study report (CSR). Additional exploratory data analyses may be conducted, as deemed appropriate.

Approved on 30 May 2025 GMT

Some analyses and summaries described in this analysis plan may not be conducted if not warranted by data (for example, too few events to justify conducting an analysis). Not all analyses described in this SAP will necessarily be included in the CSR. Any analysis described in this SAP and not provided in the CSR would be available upon request. Not all displays will necessarily be created as a "static" display. Some may be incorporated into interactive display tools instead of or in addition to a static display.

4.1.1. Definition of Baseline

Baseline for efficacy assessments is defined as the last non-missing measurement prior to the first dosing of study intervention, which in most cases will be the measurement recorded at Visit 0. If there are no doses of study intervention, baseline will be defined as the last 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 value for data analysis.

Baseline period for safety assessment starts at screening and ends prior to the first dose (typically at Visit 0). For triplicate ECG, the average of the 3 results will be used as the baseline.

4.2. Analysis Methods for Efficacy Estimand

4.2.1. Constrained Mixed Model for Repeated Measures (cMMRM) for Continuous Endpoints

For continuous outcomes with repeated postbaseline measurements, we will use a restricted maximum likelihood (REML)-based approach with dose-pooling constrained MMRM (cMMRM) analysis to assess longitudinal continuous variables (Qu et al. 2023). All scheduled postbaseline visits will contribute longitudinal observations to the analysis, aligned with the estimand strategy.

In this dose-pooling variant of MMRM, the fixed effect of treatment arm is replaced by the treatment arm with the lowest dose for periods that share the same dose. A general form for the cMMRM for participant i, at timepoint k, within assigned treatment arm j is

$$Y_{ijk} = \mu_{jk} + b_{ij} + \epsilon_{ijk},$$

where $b_{ij} \sim N(0, \sigma_s^2)$ is the between-subject error and $\epsilon_{ijk} \sim N(0, \sigma^2)$ is the residual error with the following constraints (see Figure OXA1.1.1) imposed on the postbaseline means

$$\mu_{3/6mg,k} = \mu_{10mg,k} = \mu_{16mg,k} \quad \text{for } 0 < k \leq 4$$

$$\mu_{3/6\text{mg},k} = \mu_{10mg,k} = \mu_{16mg,k} \quad \text{for } 4 < k \leq 8$$

$$\mu_{10mg,k} = \mu_{16mg,k} \quad \text{for } 8 < k \leq 12$$

for target doses of $j \in \{3/6, 10, 16\}$ mg. This approach derives a more efficient estimator when the same treatment regimen is shared across treatment groups due to a preplanned titration.

Logistically, this would involve creating a Pooled_Arm variable that assigns all patients in Weeks 0-8 to 3/6 mg (e.g., Pooled_Arm=3_6mg) and patients in the 16-mg arm to the 10-mg arm for Weeks 8-12 (e.g., Pooled_Arm=10mg). In addition to placebo, when graphing, this will generate one line segment for the 3/6-mg arm in Weeks 0-4, one line segment for the 3/6-mg arm in Weeks 4-8, two line segments for the 3/6-mg and 10-mg arms in Weeks 8-12, and three line segments for each treatment period in the remaining weeks.

The following fixed effects will be considered for inclusion in the cMMRM:

- baseline value (of the dependent variable)
- strata (factor variable)
- pooled treatment arm (factor variable)
- visit (factor variable)
- two-way interactions:
    - baseline and visit
    - strata and visit
    - pooled treatment arm and visit

In situations where higher-order interactions are included in the model, individual main effect terms may be dropped from the formula to simplify the model statement without any impact to the point estimates or inference. Restricted maximum likelihood (REML) will be used to obtain model parameter estimates. The robust 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% CI and corresponding p-values will also be reported.

Continuous outcomes lacking sufficient measurement frequency during the applicable pooling period will be analyzed using an MMRM instead of a cMMRM by replacing pooled treatment arm with treatment arm. This SAP offers the flexibility of reverting to an unconstrained MMRM, which involves fitting the model by treatment arm rather than pooled treatment arm, for any reason.

An unstructured covariance matrix will be used to model the within-participant errors for each treatment group, assuming heteroscedasticity and that measurements for different participants are independent. If the model fails to converge under the unstructured covariance structure, the following covariance structures will be tested in order:

- Heterogeneous Toeplitz,
- Heterogeneous first-order autoregressive,
- Heterogeneous compound symmetry,
- Toeplitz,
- First-order autoregressive, and
- Compound symmetry.

The first covariance structure that converges will be used. If the full model fails to converge for all covariance structures, the model will be refit by first removing an individual two-way interaction term and cycling through the covariance structures.

For some laboratory parameters, such as (but not limited to) lipid parameters, fasting insulin, glycated hemoglobin (HbA1c), homeostatic model assessment for insulin resistance (HOMA2-IR and HOMA2-B), and leptin, both the baseline and post-baseline values will be log-transformed before analyzing.

4.2.2. 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)
    - strata (factor variable)
    - treatment arm (factor variable)
4) Use G-computation to estimate the unconditional average treatment effect for each imputed dataset.
5) Derive the final inference using Rubin's Rule by combining estimates from the multiple imputed datasets.

4.2.3. ANHECOVA for Continuous Endpoints with a Single Post-Baseline Measurement

The analysis of heterogeneous covariance (ANHECOVA [Ye et al. 2023]) model will be used to analyze continuous outcomes guided by the treatment regimen estimand. The ANHECOVA estimator for the mean response was shown to be reliable, follow an asymptotically normal distribution, and outperform the traditional ANOVA or ANCOVA estimators in terms of asymptotic efficiency. The ANHECOVA model 1) includes all covariates used in randomization (e.g., strata) 2) specifies all treatment-by-covariate interactions, and 3) uses a robust estimate of standard error. The addition of interaction terms is not intended to estimate the heterogeneity effect but to provide robustness and efficiency for the estimate of treatment comparisons on the unconditional effect.

The following covariates will be considered for inclusion in the ANHECOVA:

- baseline value (of the dependent variable)
- strata (factor variable)
- treatment arm (factor variable)
- two-way interactions:
    - baseline and treatment arm

- strata and treatment arm

The associated two-sided 95% CI and corresponding p-values will also be reported. Missing values will follow the MI procedure outlined bullet 1) in Section 4.2.2.

For some parameters, such as (but not limited to) mean hepatic fat fraction, both the baseline and post-baseline values will be log-transformed before fitting the model.

4.2.4. Analysis Methods for Safety

Fisher's exact test will be used for treatment comparisons of proportions.

The cMMRM analysis described in Section 4.2.1 will be used for select continuous safety data with multiple postbaseline measurements. The following fixed effects will be considered for inclusion in the cMMRM:

- baseline value (of the dependent variable)
- pooled treatment arm (factor variable)
- visit (factor variable)
- two-way interactions:
  - visit by pooled treatment arm

The other model specifications will be similar to those outlined in Section 4.2.1.

Continuous safety outcomes lacking sufficient measurement frequency during the applicable pooling period will be analyzed using an MMRM instead of a cMMRM by replacing pooled treatment arm with treatment arm. Again, this SAP offers the flexibility of reverting to an unconstrained MMRM for any reason.

Analysis of variance (ANOVA) will be used to estimate baseline means under Type III sums of squares, with outcome as the actual value of the parameter and predictor of treatment arm as a factor variable.

The following baseline and postbaseline variables will be natural log transformed prior to analysis:

- lipids,
- calcitonin,
- p-amylase,
- lipase,
- hepatic enzymes including ALT, AST, ALP, total bilirubin (TBL), direct bilirubin (DBL), and gamma-glutamyl transferase (GGT), and
- urine albumin-to-creatinine ratio (UACR).

The Kaplan-Meier (KM) product-limit method will be used to estimate cumulative event-free survival rates over time in the time-to-event analyses. Cox proportional hazards regression will be used to compare hazard ratios between treatment groups. For Kaplan-Meier estimation, a dose-pooling strategy may be implemented, where subjects with the same pre-planned titration

history within the pre-specified analysis set will be grouped together to calculate the number of events and the total number of subjects at risk.

A logistic regression model will be used for baseline and postbaseline treatment comparisons of hypoglycemia incidence. At baseline, the outcome will be a binary variable (hypoglycemia, yes/no), and the predictors will include treatment arm as a factor variable. At postbaseline, the outcome will be a binary variable (hypoglycemia, yes/no), and the predictors may include baseline HbA1c (<5.7%, ≥5.7%), strata, and treatment arm as factor variables, if deemed appropriate.

A negative binomial regression model will be used for baseline and postbaseline treatment comparisons of hypoglycemia rates. At baseline, the outcome will be a count variable (number of hypoglycemia events), and the predictors will include treatment arm as a factor variable with log (exposure in days/365.25) as an offset variable. At postbaseline, the outcome will be a count variable (number of hypoglycemia events), and the predictors may include baseline HbA1c (<5.7%, ≥5.7%), strata, and treatment arm as factor variables, if deemed appropriate, with log(exposure in days/365.25) as an offset variable. Due to the limited number of expected hypoglycemia events at baseline in this population, baseline will not be included as a covariate in the model.

4.3. Analysis Methods for Treatment Regimen Estimand

4.3.1. Analysis of Heterogeneous Covariance (ANHECOVA) for Continuous Endpoints

The analysis of heterogeneous covariance ANHECOVA model (Ye et al. 2023) will be fit as previously outlined in Section 4.2.3. The handling of missing data will follow the procedure detailed in the subsequent Section 4.3.2.

4.3.2. Imputation Under Treatment-regimen Estimand

Objective: To estimate the effect of the treatment regimen in real-world conditions, including the potential impact of participants discontinuing the treatment.

Imputation Strategy: Efforts are made to minimize missing data by encouraging participants who discontinue the study intervention to continue in the study for the entire treatment and follow-up periods. If missing data occur despite these efforts, the missing values are imputed in a way that reflects what the values would likely have been had they been collected. This approach acknowledges that missing data might occur and aims to reflect the likely values in the presence of discontinuation or other real-world factors.

Missing data shall be minimized for estimating the treatment regimen estimand. Participants who discontinue study intervention (that is, discontinue study treatment) will be encouraged to continue in the study for treatment period/phase and follow-up period/phase. Note: the words period and phase are used interchangeably in this SAP, while the case report forms (CRFs) use "phase."

If there are occurrences of missing data despite the best precautions, missing data should be imputed in a fashion consistent with what the values would likely have been had they been collected.

4.3.2.1. Retrieved Dropout Imputation

Retrieved dropout imputation will be used as primary imputation method. If a participant discontinues the study intervention but has a non-missing endpoint from the same treatment arm, this data will be used to impute the missing value for another participant who discontinued.

4.3.2.2. Jump-to-reference Imputation

In the cases where there are not enough retrieved dropouts to provide a reliable imputation model, the imputation of missing data will be done with jump-to-reference imputation. In this case, missing data is imputed using data from the placebo group and provides a fallback imputation strategy that assumes the missing values behave like those in the reference group.

4.4. Participant Dispositions

A detailed description of participant disposition will be provided. The planned listings and summary tables for disposition are provided in Table CWMM.4.1 of the CWMM SAP.

4.5. Primary Endpoint Analysis

4.5.1. 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%.*

4.5.2. Main Analytical Approach

The main analysis guided by the efficacy estimand will be conducted on the primary endpoint using FAS and DPS1. See Section 4.2.1 for details on the modeling strategy (constrained MMRM).

4.5.3. Supplementary Analyses

A supplementary analysis guided by the treatment regimen estimand will be conducted on the primary endpoint using FAS and DPS2. See Section 4.3.1 for details on the modeling strategy (ANHECOVA).

4.6. Secondary Endpoints Analysis

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

The following are ISA-specific secondary endpoints that will also be analyzed using FAS and DPS1 guided by the efficacy estimand, with the exception of treatment-emergent antidrug antibodies (TE-ADAs), which will be analyzed using SAS and DPS3.

Table OXA1.4.1. ISA-Specific Secondary Endpoints for Clinical Protocol W8M-MC-OXA1

Measure	Endpoint	Analysis method	Timepoint
Treatment-emergent antidrug antibodies to LY3305677	Proportion with preexisting ADA, TE-ADA, cross-reactive antibodies, and neutralizing antibodies will be tabulated by dose	• Descriptive summaries by treatment group	• Week 32 • Week 48
Pharmacokinetics (PK) of the primary intervention	Described in a separate PK/PD analysis plan (see Section 6.2 [Appendix 2])	• Described in a separate PK/PD analysis plan (see Section 6.2 [Appendix 2])	• Week 32 • Week 48
Liver fat content (LFC) as measured by MRI-PDFF	For participants with baseline LFC ≥5% and for participants with baseline LFC ≥10%: • Change and percent change from baseline in LFC [%] • Proportion with ≥30% relative reduction in LFC [%]	• Change and percent change: ○ Primary data lock: ANHECOVA, efficacy estimand, Section 4.2.3 ○ Final data lock: MMRM, efficacy estimand, Section 4.2.1 • Proportion: logistic regression, efficacy estimand, Section 4.2.2	• Week 32 • Week 48

Abbreviations: ANHECOVA = analysis of heterogeneous covariance; ISA = intervention-specific appendix; MRI-PDFF = magnetic resonance imaging - proton density fat fraction.

4.6.1. Pharmacokinetic/Pharmacodynamic Methods

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

LY3305677 concentration-time data will be summarized in the clinical study report. Dose and exposure-response analyses between LY3305677 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.

4.6.2. 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 one non-missing postbaseline ADA result.

Listings of patients who are not TE-ADA evaluable, patients with at least one test having detected LY3305677 ADAs and patients having LY3305677 ADAs present or TEAEs (hypersensitivity reactions or injection site reactions) will be provided.

The frequency and percentage of patients with preexisting ADA, TE-ADA, cross-reactive antibodies, and neutralizing antibodies will be tabulated by treatment arm (if data warrant), where proportions are relative to the number of patients who are TE-ADA evaluable. The frequency and percentage of patients with hypersensitivity and injection site reaction 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 LY3305677 may be assessed.

4.7. Exploratory Endpoints Analysis

See CWMM SAP Section 4.4 for general overview of the exploratory endpoint analysis and select exploratory endpoints not covered here.

For exploratory efficacy endpoints with multiple post-baseline timepoints, the MMRM model used in the main analysis (Section 4.2.1) 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.

The following ISA-specific exploratory endpoints will also be analyzed on FAS and DPS1 guided by the efficacy estimand.

Table OXA1.4.2. ISA-Specific Exploratory Endpoints for Clinical Protocol W8M-MC-OXA1

Measure	Endpoint	Analysis method	Timepoint
BP and heart rate from ABPM	Change from baseline in: • 24-hour mean systolic BP (mmHg) measured by ABPM • 24-hour mean diastolic BP (mmHg) measured by ABPM • 24-hour mean heart rate measured by ABPM	• Primary data lock: ANHECOVA, efficacy estimand, Section 4.2.3 • Final data lock: MMRM, efficacy estimand, Section 4.2.1	• Week 32 • Week 48

Measure	Endpoint	Analysis method	Timepoint
Physical activity	Change from baseline in: • Daily steps count • M10 (activity intensity of most active 10 hours during 24-hour period) • moderate-to-vigorous physical activity (MVPA) in minutes during 24-hour period	Section 4.7.1	• Week 32 • Week 48
HRQoL, function, and appetite	Change from baseline in: • PROMIS SF Pain Interference 4a v1.1 • PGIS - Physical Function due to Weight • PGIC - Physical Function due to Weight • Domains of SF-36 v2, Acute • Appetite VAS	Section 4.7.1	• Week 32 • Week 48
Glucose regulation	Change from baseline in: • Fasting insulin • Fasting glucose • HbA1c • HOMA2-IR • HOMA2-B	MMRM, efficacy estimand, Section 4.2.1.	• Week 32 • Week 48

Biomarkers	Change from baseline in mechanistic biomarkers.
	See OXA1 Protocol Section 10.5 for detailed list of biomarkers and their calculations.
		• Longitudinal: MMRM, efficacy estimand, Section 4.2.1.
		• Endpoint:
		  ○ Primary data lock: ANHECOVA, efficacy estimand, Section 4.2.3
		  ○ Final data lock: MMRM, efficacy estimand, Section 4.2.1
		• Biomarkers with substantial missingness will instead be summarized using descriptive statistics.
			• Week 32
			• Week 48

Dose and/or exposure and key safety and efficacy measures, as applicable	Dose-response and exposure-response analyses for key safety and efficacy endpoints, as applicable	Described in a separate PK/PD analysis plan (see Section 6.2 [Appendix 2])	• Week 32 • Week 48

Measure	Endpoint	Analysis method	Timepoint
Body composition as measured by MRI	Change and percent change from baseline in: • Abdominal visceral adipose tissue (VAT) volume [L] • Abdominal subcutaneous adipose tissue (SAT) volume [L] • Total abdominal adipose tissue (VAT+SAT) volume, also called total abdominal fat (TAF) volume [L] • Total body adipose tissue (TAT) volume [L] • Total thigh fat-free muscle volume [L] • Total spinal erectors L1-L5 fat-free muscle volume [L] • Total body lean tissue (TLT) volume [L] • Mean anterior thigh muscle fat infiltration [%] • Mean spinal erectors L1-L5 muscle fat infiltration [%] Change from baseline in: • Visceral adipose tissue ratio [%] (VAT volume divided by TAF volume)	• Primary data lock: ANHECOVA, efficacy estimand, Section 4.2.3 • Final data lock: MMRM, efficacy estimand, Section 4.2.1	• Week 32 • Week 48
Liver fat content (LFC) as measured by MRI-PDFF	Proportion with baseline LFC ≥5% and baseline LFC ≥10% who achieve: • LFC < 5% (normal range) • ≥50% relative reduction in LFC	Logistic regression, efficacy estimand, Section 4.2.2	• Week 32 • Week 48

Abbreviations: ABPM = ambulatory blood pressure monitoring; ANHECOVA = analysis of heterogeneous variance; BP = blood pressure; MMRM= mixed model for repeated measures; ELF = enhanced liver fibrosis; HbA1c = glycated hemoglobin A1c; HOMA2-B = updated (1998) homeostasis model assessment of beta-cell function; HOMA2-IR = updated (1998) homeostasis model assessment of insulin resistance; HRQoL = health-related quality of life; MRI = magnetic resonance imaging; NASH = non-alcoholic steatohepatitis; NIS4 = NASH Index Score 4; PROMIS SF = Patient-Reported Outcome Measurement Information System Short Form; PGIC = Patient Global Impression of Change; PGIS = Patient Global Impression of Severity; SF-36 v2 = Short Form-36 Version 2; VAS = visual analog scale.

4.7.1. 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 (return device at Visit 0) and Visit 32 (return device at Visit 40) to collect activity data for at least 2 weeks. The device is then given to the participant at Visit 40, and they are instructed to start wearing the device at Visit 44 (return device at Visit 48).

The exploratory endpoints related to actigraphy are to compare the effect of LY3305677 versus placebo in change from baseline at Weeks 32 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, Week 32, and Week 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 32 and 48 in physical activity will be analyzed according to the efficacy estimand using an MMRM model and as detailed in Section 4.2.1. 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.

4.7.2. Health Outcomes

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

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

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

4.7.2.1. PROMIS Short Form Pain Interference 4a v1.1

The Patient-Reported Outcome Measurement Information System (PROMIS) (Health Measures [WWW]) is a set of patient-completed measures that evaluates and monitors physical, mental, and social health in adults and children. It can be used with the general population and with individuals living with chronic conditions.

The PROMIS Short Form Pain Interference 4a assesses consequences of pain on relevant aspects of one's life. This includes the extent to which pain interferes with day-to-day, social, and household activities. The PROMIS Short Form Pain Interference 4a consists of 4 items that asks participants to rate their pain interference over the past 7 days on a 5-point scale ranging from "1 – not at all" to "5 – very much." Individual item scores are totaled to obtain a raw score, with higher scores indicating more interference. Raw scores can be converted to a t-score, which is

standardized to have a mean of 50 and a standard deviation (SD) of 10; higher scores indicate better levels of function and/or better health. The PROMIS Physical Function user manual and scoring instructions are located online (PROMIS Health Organization 2023).

The following analyses will be conducted using FAS and DPS1:

- descriptive summaries by treatment group, and
- analysis by MMRM as described in Section 4.2.1.

4.7.2.2. Patient Global Impression of Change (PGIC) – Physical Function due to Weight

The Patient Global Impression of Change – Physical Function due to Weight (PGIC – physical function weight) questionnaire is designed to assess the participant's 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."

The following analyses will be conducted using FAS and DPS1:

- Shift table summary for physical activity response categories. Statistical significance will not be assessed for shift tables.

4.7.2.3. Patient Global Impression of Severity (PGIS) – Physical Function due to Weight

The Patient Global Impression of Severity - Physical Function due to Weight (PGIS – 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."

The following analyses will be conducted using FAS and DPS1:

- Shift table summary for physical activity response categories. Statistical significance will not be assessed for shift tables.

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

The Short Form-36 version 2 Health Survey acute form (SF-36 v2 acute), one-week recall version is a 36-item generic, participant-completed measure designed to assess the following 8 domains:

- Physical functioning,
- Role-physical,
- Bodily pain,
- General health,
- Vitality,
- Social functioning,
- Role-emotional, and
- Mental health.

The Physical Functioning domain assesses limitations due to health "now" while the remaining domains assess functioning "in the past week." Each domain is scored individually and information from these 8 domains is further aggregated into two health component summary scores: Physical Component Summary and Mental Component Summary. Items are answered on Likert scales of varying lengths (3-point, 5-point, or 6-point scales). Scoring of each domain and both summary scores are norm based and presented in the form of t-scores, with a mean of 50 and SD of 10; higher scores indicate better levels of function and/or better health (Maruish 2011).

The following analyses will be conducted using FAS and DPS1:

- descriptive summaries by treatment group, and
- analysis by MMRM as described in Section 4.2.1.

4.7.2.5. Appetite Visual Analog Scale (VAS)

To explore the effects of LY3305677 on meal intake and appetite sensation, participants will be asked to rate their appetite sensations using a 100-mm VAS for parameters of hunger, fullness, satiety, and prospective food consumption in the fasted state. The VAS is a validated tool to assess appetite sensation parameters (Flint et al. 2000). The VAS is presented as a 10-cm (100-mm) line, anchored by verbal descriptors, usually "extremely" and "not at all." Participants are required to rate their subjective sensations on four 100-mm scales combined with questions similar to the following:

Overall appetite score is calculated as the average of the 4 individual scores: satiety + fullness + (100-prospective food consumption) + (100-hunger) / 4 (van Can et al. 2014). A higher overall appetite score indicates less appetite, and a lower score indicates more appetite.

The following analyses will be conducted using FAS and DPS1:

- descriptive summaries by treatment group, and
- analysis by MMRM as described in Section 4.2.1.

Table OXA1.4.3. Description and Derivation of Primary and Secondary Endpoints and Select Exploratory Endpoints for Efficacy Measures

Measure	Description	Variable	Derivation/Comment	Handling Missing Components
Body weight	Per CWMM Protocol Section 10.8: • Body weight measurements should be done in a consistent manner using a calibrated electronic scale capable of measuring weight in kilograms to 1 decimal place. • All weights for a given participant should be measured using the same scale, whenever possible, at approximately the same time in the morning after evacuation of bladder contents. • Body weight will be measured in fasting state at all visits. If the participant is not fasting, the participant should be called in for a new visit within the visit window to have the fasting body weight measured.	Body weight (kg)	As measured	Single item, missing if missing
		Change from baseline in body weight (kg)	Calculated as: postbaseline body weight (kg) – baseline body weight (kg)	Missing if baseline or postbaseline value is missing
		Percent change from baseline in body weight (%)	Calculated as: (postbaseline body weight [kg] – baseline body weight [kg]) / (baseline body weight [kg]) * 100%	Missing if baseline or postbaseline value is missing
		≥x% body weight reduction from baseline Where, x = 5, 10, 15, 20, 25, 30	Response = yes if at least a x% reduction in body weight from baseline, that is, percent change from baseline in body weight ≤ -x%	Missing if baseline or postbaseline value is missing
BMI	BMI: Round to one decimal point. For example, a BMI of 26.6 kg/m^2 should **not** be rounded to 27.0 kg/m^2.	BMI (kg/m^2)	BMI will be calculated as: $\text{BMI (kg/m}^2) = \text{Weight (kg)} / (\text{Height [m]})^2$	Missing if weight or height is missing
		Change from baseline in BMI (kg/m^2)	Calculated as: postbaseline BMI (kg/m^2) – baseline BMI (kg/m^2)	Missing if baseline or postbaseline value is missing
Waist circumference	Per CWMM Protocol Section 10.8: • Waist circumference should be measured in centimeters in the horizontal plane and at the midpoint between the lower margin of the last palpable rib and the top of the iliac	Waist circumference (cm)	The average of 2 measures	If there is at least one measurement, take the average of all non-missing measurements; otherwise, set it to be missing

Measure	Description	Variable	Derivation/Comment	Handling Missing Components
	crest. • Measurements must be taken at the end of a normal expiration using a non-stretchable measuring tape provided. The tape should lie flat against the skin without compressing the soft tissue. The waist circumference should be measured twice, rounded to the nearest 0.5 cm. The measuring tape should be removed between the 2 measurements. Both measurements will be recorded in the CRF. If the difference between the 2 measurements exceeds 1 cm, this set of measurements should be discarded and the 2 measurements repeated.	Change from baseline in waist circumference (cm)	Calculated as: Postbaseline waist circumference (cm) – baseline waist circumference (cm)	Missing if baseline or postbaseline value is missing
Blood pressure	Per CWMM Protocol Section 10.8: • have the participant sit quietly for about 5 minutes before vital signs measurements are taken. • for each parameter, take 2 measurements from the same arm, preferably the nondominant arm. • measure the recordings at least 1 minute apart. blood pressure must be taken with an automated BP instrument with the appropriate size cuff on the upper arm.	Systolic blood pressure, SBP (mmHg)	The average of 2 measures	If there is at least one measurement, take the average of all non-missing measurements; otherwise, set it to be missing
		Change from baseline in SBP (mmHg)	Calculated as: Postbaseline SBP (mmHg) – baseline SBP (mmHg)	Missing if baseline or postbaseline value is missing
		Diastolic blood pressure, DBP (mmHg)	The average of 2 measures	If there is at least one measurement, take the average of all non-missing measurements; otherwise, set it to be missing
		Change from baseline in DBP (mmHg)	Calculated as: Postbaseline DBP (mmHg) – baseline DBP (mmHg)	Missing if baseline or postbaseline value is missing

Measure	Description	Variable	Derivation/Comment	Handling Missing Components
Lipid parameter	The tests will be performed by the central laboratory or Lilly designated laboratory.	Triglycerides	As provided Log transformation before the analysis	Single item, missing if missing
		Percent change from baseline in triglycerides (%)	Calculated as: log(postbaseline triglyceride) – log (baseline triglyceride) then will transform back to percent change.	Missing if baseline or postbaseline value is missing
		Total cholesterol	As provided Log transformation before the analysis	Single item, missing if missing
		Percent change from baseline in total cholesterol (%)	Calculated as: log(postbaseline total cholesterol) – log(baseline total cholesterol) then will transform back to percent change	Missing if baseline or postbaseline value is missing
		Non-HDL-C	As provided Log transformation before the analysis	Single item, missing if missing
		Percent change from baseline in non-HDL-C (%)	Calculated as: log(postbaseline non-HDL-C) – log(baseline non-HDL-C) then will transform back to percent change	Missing if baseline or postbaseline value is missing
		LDL-C	As provided Log transformation before the analysis.	Single item, missing if missing
		Percent change from baseline in LDL-C (%)	Calculated as: log(postbaseline LDL-C) – log(baseline LDL-C) then will transform back to percent change	Missing if baseline or postbaseline value is missing
		HDL-C	As provided Log transformation before the analysis	Single item, missing if missing
		Percent change from baseline in HDL-C (%)	Calculated as: log(postbaseline HDL-C) – log(baseline HDL-C) then will transform back to percent change	Missing if baseline or postbaseline value is missing

Measure	Description	Variable	Derivation/Comment	Handling Missing Components
Ambulatory blood pressure	Per OXA1 Protocol Section 8.2.12.1 The ABPM device will be attached to the nondominant arm, and participants will be instructed to wear the monitor for a 24- to 27-hour period. Ambulatory blood pressure measurements: • Should be collected on a typical workday, not on a non-working day • Will be recorded every 30 minutes during daytime hours (0700 to 2200 hours) • Will be recorded every 60 minutes during nighttime hours (2200 to 0700 hours). A 24-hour session of ambulatory monitoring produces technically acceptable measurements if ≥70% of the readings are valid.	Systolic blood pressure, SBP (mmHg)	The average will be pre-calculated across the first 24 hours if at least 70% of the measurements are valid.	Missing if <70% of measurements are valid.
		Change from baseline in SBP (mmHg)	Calculated as: Postbaseline SBP (mmHg) – baseline SBP (mmHg)	Missing if baseline or postbaseline value is missing
		Diastolic blood pressure, DBP (mmHg)	The average will be pre-calculated across the first 24 hours if at least 70% of the measurements are valid.	Missing if <70% of measurements are valid.
		Change from baseline in DBP (mmHg)	Calculated as: Postbaseline DBP (mmHg) – baseline DBP (mmHg)	Missing if baseline or postbaseline value is missing
		Heart rate (HR) (beats/min)	The average will be pre-calculated across the first 24 hours if at least 70% of the measurements are valid.	Missing if <70% of measurements are valid.
		Change from baseline in HR (beats/min)	Calculated as: Postbaseline HR (beats/min) – baseline HR (beats/min)	Missing if baseline or postbaseline value is missing

Measure	Description	Variable	Derivation/Comment	Handling Missing Components
Liver fat measures	Per OXA1 Protocol Section 8.1.3 MRI measurements at baseline should be performed between Visit 402 and Visit 0 (randomization) in OXA1, before the participant receives the first dose of study intervention. The postbaseline MRI measurements at Week 32 should occur within ±2 weeks. Week 48 measurements should be scheduled and performed anytime between Visits 44 and 48, and ED measurements may be performed within ±7 days of the date of the study visit. Both MRI measurements should be performed after a fast of at least 8 hours.	Mean liver fat content (LFC, %). Also known as mean hepatic fat fraction (HFF).	As provided: use the variable for mean hepatic fat fraction corresponding to liver (not the Couinaud segments). The liver value is the mean hepatic fat fraction across the whole liver. Log transformation before the analysis	Single item, missing if missing
		Change from baseline in mean liver fat content (%)	Calculated as: log(postbaseline mean LFC) – log(baseline mean LFC)	Missing if baseline or postbaseline value is missing
		Percent change from baseline in mean liver fat content (%)	Calculated as: log(postbaseline mean LFC) – log(baseline mean LFC) then will transform back to percent change	Missing if baseline or postbaseline value is missing
Body composition measures*	Per OXA1 Protocol Section 8.1.3 MRI measurements at baseline should be performed between Visit 402 and Visit 0 (randomization) in OXA1, before the participant receives the first dose of study intervention. The postbaseline MRI measurements at Week 32 should occur within ±2 weeks. Week 48 measurements should be scheduled and performed anytime between Visits 44 and 48, and ED measurements may be performed within ±7 days of the date of the study visit. Both MRI measurements should be performed after a fast of at least 8 hours.	Abdominal visceral adipose tissue volume (VATVOL) in L	As provided using location of ABDOMINAL CAVITY Log transformation before the analysis	Single item, missing if missing
		Change from baseline in VATVOL (L)	Calculated as: log(postbaseline VATVOL) – log(baseline VATVOL)	Missing if baseline or postbaseline value is missing
		Percent change from baseline in VATVOL (%)	Calculated as: log(postbaseline VATVOL) – log(baseline VATVOL) then will transform back to percent change	Missing if baseline or postbaseline value is missing
		Abdominal subcutaneous adipose tissue volume (SATVOL) in L	As provided using location of ABDOMINAL CAVITY Log transformation before the analysis	Single item, missing if missing

Measure	Description	Variable	Derivation/Comment	Handling Missing Components
		Change from baseline in SATVOL (L)	Calculated as: log(postbaseline SATVOL) – log(baseline SATVOL)	Missing if baseline or postbaseline value is missing
		Percent change from baseline in SATVOL (%)	Calculated as: log(postbaseline SATVOL) – log(baseline SATVOL) then will transform back to percent change	Missing if baseline or postbaseline value is missing
		Total abdominal adipose tissue volume, also known as total abdominal fat volume (TAFVOL) in L	Calculated as VATVOL + SATVOL Log transformation before the analysis	Single item, missing if missing
		Change from baseline in TAFVOL (L)	Calculated as: log(postbaseline TAFVOL) – log(baseline TAFVOL)	Missing if baseline or postbaseline value is missing
		Percent change from baseline in TAFVOL (%)	Calculated as: log(postbaseline TAFVOL) – log(baseline TAFVOL) then will transform back to percent change	Missing if baseline or postbaseline value is missing
		Total adipose tissue volume (TATVOL) in L	As provided using location of BODY Log transformation before the analysis	Single item, missing if missing
		Change from baseline in TATVOL (L)	Calculated as: log(postbaseline TATVOL) – log(baseline TATVOL)	Missing if baseline or postbaseline value is missing

Measure	Description	Variable	Derivation/Comment	Handling Missing Components
		Percent change from baseline in TATVOL (%)	Calculated as: log(postbaseline TATVOL) – log(baseline TATVOL) then will transform back to percent change	Missing if baseline or postbaseline value is missing
		Total thigh fat-free muscle volume (TTFFMVOL) in L	As provided using location of THIGH Log transformation before the analysis	Single item, missing if missing
		Change from baseline in TTFFMVOL (L)	Calculated as: log(postbaseline TTFFMVOL) – log(baseline TTFFMVOL)	Missing if baseline or postbaseline value is missing
		Percent change from baseline in TTFFMVOL (%)	Calculated as: log(postbaseline TTFFMVOL) – log(baseline TTFFMVOL) then will transform back to percent change	Missing if baseline or postbaseline value is missing
		Total spinal erectors L1-L5 fat-free muscle volume (TSEFFMVOL) in L.	As provided using location of BACK Log transformation before the analysis	Single item, missing if missing
		Change from baseline in TSEFFMVOL (L)	Calculated as: log(postbaseline TSEFFMVOL) – log(baseline TSEFFMVOL)	Missing if baseline or postbaseline value is missing
		Percent change from baseline in TSEFFMVOL (%)	Calculated as: log(postbaseline TSEFFMVOL) – log(baseline TSEFFMVOL) then will transform back to percent change	Missing if baseline or postbaseline value is missing

Measure	Description	Variable	Derivation/Comment	Handling Missing Components
		Total body lean tissue volume (TLTVOL) in L	As provided using location of BODY Log transformation before the analysis	Single item, missing if missing
		Change from baseline in TLTVOL (L)	Calculated as: log(postbaseline TLTVOL) – log(baseline TLTVOL)	Missing if baseline or postbaseline value is missing
		Percent change from baseline in TLTVOL (%)	Calculated as: log(postbaseline TLTVOL) – log(baseline TLTVOL) then will transform back to percent change	Missing if baseline or postbaseline value is missing
		Mean anterior thigh muscle fat infiltration (MATMFI) as %	As provided using location of THIGH Log transformation before the analysis	Single item, missing if missing
		Change from baseline in MATMFI (%).	Calculated as: log(postbaseline MATMFI) – log(baseline MATMFI)	Missing if baseline or postbaseline value is missing
		Percent change from baseline in MATMFI (%).	Calculated as: log(postbaseline MATMFI) – log(baseline MATMFI) then will transform back to percent change	Missing if baseline or postbaseline value is missing
		Mean spinal erectors L1-L5 muscle fat infiltration (MSEMFI) as %	As provided using location of BACK Log transformation before the analysis	Single item, missing if missing

Measure	Description	Variable	Derivation/Comment	Handling Missing Components
		Change from baseline in MSEMFI (%).	Calculated as: log(postbaseline MSEMFI) – log(baseline MSEMFI)	Missing if baseline or postbaseline value is missing
		Percent change from baseline in MSEMFI (%).	Calculated as: log(postbaseline MSEMFI) – log(baseline MSEMFI) then will transform back to percent change	Missing if baseline or postbaseline value is missing
		Visceral adipose tissue ratio (VATRAT) as %	As provided using location of ABDOMINAL CAVITY Log transformation before the analysis	Single item, missing if missing
		Change from baseline in VATRAT (%).	Calculated as: log(postbaseline VATRAT) – log(baseline VATRAT)	Missing if baseline or postbaseline value is missing
Fatty liver index (FLI)*	A measure of hepatic steatosis (Bedogni et al., 2006) that uses the following parameters in its calculation: • Triglycerides (mg/dL) • BMI (kg/m^2) • Gamma-glutamyl transferase (GGT in U/L) • Waist circumference (cm)	FLI (unitless)	Calculated as: FLI = exp(Z)/(1+exp(Z)) x 100 where Z=0.953*log(Triglycerides in mg/dL) + 0.139*BMI (kg/m^2) + 0.718*log(GGT in U/L)+0.053*(Waist Circumference in cm) - 15.745 Log transformation before the analysis	Single item, missing if missing
		Change from baseline in FLI (unitless)	Calculated as: log(postbaseline FLI) – log(baseline FLI)	Missing if baseline or postbaseline value is missing
		Percent change from baseline in FLI (%)	Calculated as: log(postbaseline FLI) – log(baseline FLI) then will transform back to percent change	Missing if baseline or postbaseline value is missing

Measure	Description	Variable	Derivation/Comment	Handling Missing Components
Anion gap	Calculated via the three-component formula which involves sodium (Na^+), chloride (Cl^-), and bicarbonate (HCO_3^-)	Anion gap (AG) in mmol/L	Calculated as: AG (mmol/L) = [Na^+] (mmol/L) – [Cl^-] (mmol/L) – [HCO_3^-] (mmol/L)	Single item, missing if missing
		Change from baseline in AG (mmol/L)	Calculated as: postbaseline AG – baseline AG	Missing if baseline or postbaseline value is missing
		Percent change from baseline in AG (%)	Calculated as: (postbaseline AG – baseline AG)/(baseline AG) x 100%	Missing if baseline or postbaseline value is missing

Abbreviations: BMI = body mass index; BP = blood pressure; CRF = case report form; GGT = gamma-glutamyl transferase; DBP = diastolic blood pressure; FLI = fatty liver index; HDL-C = high-density lipoprotein- cholesterol; HFF = hepatic fat fraction; LDL-C = low-density lipoprotein-cholesterol; LFC = liver fat content; MRI = magnetic resonance imaging; SBP = systolic blood pressure; VLDL-C = very low-density lipoprotein-cholesterol.

*Log-transformations were proposed based on examination of preliminary baseline data, historical data, or available literature. This SAP offers the flexibility of not log-transforming the data if the future data suggests otherwise.

4.8. Safety Analyses

All safety analyses will be conducted on the safety analysis set and DPS3.

4.8.1. Extent of Exposure

Listing of exposure to LY3305677 and placebo will be provided by treatment group. For all randomized patients, 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 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 weeks,
- 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.

4.8.2. 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 was defined in Section 4.1.1. The 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. Events with missing severity during the baseline period 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.

Table OXA1.4.4 lists the planned summaries for AEs.

Table OXA1.4.4. Summary Tables Related to Adverse Events

Analysis	Method	Analysis/Data Point Sets
Overview of AEs, including • TEAE • SAE • death, and • 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

Analysis	Method	Analysis/Data Point Sets
AEs leading to study treatment interruption by PT within SOC	Fisher's exact	SAS + DPS3
Listing of AEs		Entered
Listing of SAEs		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; MedDRA = Medical Dictionary for Regulatory Activities; PT = MedDRA Preferred Term; SAS = safety analysis set; 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.

4.8.3. Safety Laboratory Measures

Please see the CWMM-SAP Section 4.5.3 for a description of the analyses.

4.8.4. Vital Signs and Physical Characteristics

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

4.8.5. Electrocardiograms

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

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

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

4.9. Other Analyses

4.9.1. 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).

With the exception of race, 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.

For subgroups with 10% within a subgroup and a sufficient number of participants to fit the full model, the cMMRM model as described in Section 4.2.1 will be used. For subgroups with 10% within a subgroup, but insufficient number of participants to fit the full model, a simpler cMMRM will be fit using predictors of pooled treatment arm, visit, baseline value (of the dependent variable), and interaction between treatment and visit.

Divisions by race for subgroup analyses will maintain at least the 5 race categories specified by FDA guidance (FDA 2016), including White, Native Hawaiian or Pacific Islander, Black or African American, Asian, and American Indian or Alaska Native, even if a subgroup makes up less than 10% of the total study population. Because some subgroups by race are anticipated to make up less than 10% of the study population, only summary statistics will be provided for subgroup analyses by race.

4.10. Database Locks and Interim Analyses

The primary outcome database lock for OXA1 will occur when all participants have completed 32 weeks of treatment or ED visit. All applicable (Week 32) primary efficacy objectives, secondary efficacy objectives (excluding antidrug antibodies), select exploratory objectives, and safety objectives will be assessed at this time. The final database lock and final analyses for this study will be performed after all randomly assigned participants have completed the study. Unblinded data and results (including primary study results at 32 weeks) will not be shared with the study sites to maintain blinding at the sites while the study is still ongoing. Further details are specified in the unified master blinding/unblinding plan, OXA1 blinding/unblinding plan and in the AC charter. No interim analysis is planned before the primary database lock.

Early access to the PK and PD data for OXA1 before the primary database lock may be conducted to allow population PK/PD analysis and model development. If applicable, this early access will be detailed in the Unblinding Plan and the Population PK/PD Analysis Plan.

4.11. Changes to Protocol-Planned Analyses

No changes to the analyses specified in Protocol OXA1.

5. Sample Size Determination

Approximately 165 participants will be randomly assigned in a 3:2:3:3 ratio with

- 45 participants allocated to placebo,
- 30 to LY3305677 3/6 mg,
- 45 to LY3305677 10 mg, and
- 45 to LY3305677 16 mg.

Assuming a 20% dropout rate, this results in approximately 132 total completers with 24 completers on the LY3305677 3/6 mg arm and 36 completers for the placebo, each of 10 mg and 16 mg LY3305677 arms.

An upper limit of 60% enrollment of women will be used to ensure a sufficiently large sample of men.

Sample size determination is based on the evaluation of superiority to placebo that will be conducted for each of the three LY3305677 doses at a two-sided significance level of 0.05 using a two-sample t-test.

The LY3305677 group mean percentage change in body weight from baseline at Week 32 compared to placebo is assumed to be -8% assuming a common SD of 10%.

The chosen sample size provides at least 91% power to establish superiority of at least one of the LY3305677 doses compared to placebo. No adjustment for multiplicity will be performed.

Approved on 30 May 2025 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, and
- Additional Safety Analysis Search Criteria

6.2. Appendix 2: Statistical Analysis Plan for PK/PD analyses

The SAP for the PK/PD analyses is in a separate document in Veeva Vault RIM under document number VV-CLIN-130785.

Approved on 30 May 2025 GMT

7. References

Andersen SW, Millen BA. On the practical application of mixed effects models for repeated measures to clinical trial data. *Pharm Stat*. 2013;12(1):7–16.

Bedogni G, Bellentani S, Miglioli L, et al. The Fatty Liver Index: a simple and accurate predictor of hepatic steatosis in the general population. *BMC Gastroenterology*. 2006;6:33.

Diggle PJ, Liang KY, Zeger SL. *Analysis of Longitudinal Data*. Oxford University Press; 1994.

[FDA]. United States Food and Drug Administration. Collection of race and ethnicity data in clinical trials. Guidance for industry and Food and Drug Administration staff. October 26, 2016. Accessed May 09, 2025. https://www.fda.gov/media/75453/download

Flint A, Raben A, Blundell JE, Astrup A. Reproducibility, power and validity of visual analogue scales in assessment of appetite sensations in single test meal studies. *Int J Obes Relat Metab Disord.* 2000;24(1):38-48.

Health Measures. Patient-Reported Outcome Measurement Information System (PROMIS®) website. http://www.healthmeasures.net/explore-measurement-systems/promis

Mansournia MA, Nazemipour M, Naimi AI, et al. Reflection on modern methods: demystifying robust standard errors for epidemiologists. *Int J Epidemiol.* 2021;50(1):346-351.

Maruish ME. *User's Manual for the SF-36v2 Health Survey.* Quality Metric Incorporated; 2011.

PROMIS Health Organization. *User Manual and Scoring Instructions PROMIS Physical Function*. 2011. https://www.healthmeasures.net/images/PROMIS/manuals/Scoring_Manual_Only/PROMIS_Physical_Function_User_Manual_and_Scoring_Instructions_08Sept2023.pdf

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.* 2023;15:214-224.

van Can J, Sloth B, Jensen CB, et al. Effects of the once-daily GLP-1 analog liraglutide on gastric emptying, glycemic parameters, appetite and energy metabolism in obese, non-diabetic adults. *Int J Obes (Lond)*. 2014;38(6):784-793.

Ye T, Shao J, Yi Y, Zhao Q. Toward better practice of covariate adjustment in analyzing randomized clinical trials. *J Am Stat Assoc.* 2023;118(544):2370-2382.

Signature Page for VV-CLIN-128213 v3.0

Approval	PPD	Statistician 30-May-2025 13:40:26 GMT+0000

Signature Page for VV-CLIN-128213 v3.0