I9X-MC-MTAE Statistical Analysis Plan Version 4

Assessment of Safety, Tolerability, and Efficacy of LY3372689 in Early Symptomatic Alzheimer's Disease

NCT05063539

Approval Date: 01-Oct-2024

Title Page

Protocol Title: Assessment of Safety, Tolerability, and Efficacy of LY3372689 in Early Symptomatic Alzheimer's Disease

Protocol Number: I9X-MC-MTAE

Compound Number: LY3372689

Short Title: Assessment of Safety, Tolerability, and Efficacy of LY3372689 in Early Symptomatic Alzheimer's Disease

Sponsor Name: Eli Lilly and Company

Legal Registered Address: Indianapolis, Indiana, USA 46285

Regulatory Agency Identifier Number(s)

IND: 141582

EndraCT: 2021-000170-29

Table of Contents

Version history

This Statistical Analysis Plan (SAP) Version 3 for Study I9X-MC-MTAE is based on the protocol (d) dated 18 April 2024. SAP Version 1 was approved prior to the first visit at which a subject received study drug.

Table 1 SAP Version History Summary

Version	Approval Date	Change	Rationale
1	15 Jul 2022	Not Applicable	Original version
2	08Dec2023	Updated the statistical methods that will be analyzing clinical efficacy endpoints, including graphical testing scheme.	Revisited the analysis plan with better knowledge for the amount of data available during the common close period given the learnings from actual enrollment speed and early discontinuation.
3		Section 3: removed Biomarker Evaluable Set and the non-programmable item in per-protocol definition. Section 4.5: added a few more variables to be included in baseline demographic table. Section 4.8: added effect of interest for DPM analyses; updated sensitivity analysis for more specific details regarding timepoints being included. Section 4.9.2 to Section 4.9.3: updated with more details. Added Section 4.11.2 and Section 4.11.3: analyses for Digital Clock Drawing Test. Digit Symbol Substitution Test, time to progression based on CDRGS	for clarity, completeness, accuracy, and flexibility in less priority analyses.
		Section 4.11.1: Updated Analysis of Blood-based biomarkers	Updated per amendment (c).

		Section 4.12.1: removed 'prior use approved DMT subgroup analysis' Removed covariate adjustment for 'prior or current use of any approved disease-modifying therapy for AD'	Little value given very few (<5) enrolled participants enrolled in this subgroup.
		Section 4.8.3: Removed NCS3, per-protocol, completer analyses. Removed Bayesian Analysis with Shared Control Removed Health Outcome and Quality of Life analyses.	No longer of interest or not applicable.
4		Section 2: modified the language to add clarity.	For clericity
		Section 4.9.2.1: Removed the pre-specified probability thresholds associated with the graphical testing scheme to control the two-sided Type I error at 0.05 level due to the rationale provided on the right. Instead, the methodological approach for how probability thresholds will be calculated is pre-specified.	Subsequent Simulation has revealed the thresholds in the graphical testing scheme may be sensitive to the underlying assumption of the placebo trajectory and variance-covariance matrix with the Bayesian DPM. The methodological approach that was used to calculated the probability thresholds was also modified to maintain the consistency with the data generation mechanism as the sample size determination (as described in Section 5).
		Correction to the V3 History where it stated 'Section 4.9.2 to Section 4.9.3: updated with more details. Added key secondary analyses to graphical testing scheme.' The key secondary analysis was not added to graphical testing scheme in the V3 amendment.	Error correction.

1. Introduction

1.1. Objectives, Endpoints, and Estimands

Objectives	Endpoints[a]
Primary	
To assess the effect of LY3372689 vs. placebo on clinical progression in participants with early symptomatic AD with demonstrated presence of moderate[b] levels of tau pathology	iADRS change from baseline through end time point (76-124 weeks)
Key Secondary	
To assess the effect of LY3372689 vs. placebo on clinical progression in full study population (moderate + high[c] levels of tau pathology) with early symptomatic AD	iADRS change from baseline through end time point (76-124 weeks)
Other Secondary	
To evaluate safety and tolerability of LY3372689	Standard safety assessments • Spontaneously reported AEs • Clinical laboratory tests • Vital sign and body weight measurements • 12-lead ECGs • Physical and neurological examinations • C-SSRS • MRI (treatment-emergent radiological findings)
To assess the effect of LY3372689 vs. placebo on clinical progression in full study population and moderate tau sub population with early symptomatic AD	Change in cognition and function from baseline iADRS through end time point (76-124 weeks) as measured by: • ADAS-Cog13 • ADCS-iADL • CDR-SB • MMSE
To assess the effect of LY3372689 vs. placebo on brain region volumes	Change in volumetric MRI measures from baseline through end time point (76 - 124 weeks)

To assess the effect of LY3372689 vs. placebo on brain tau deposition	Change in brain tau deposition from baseline through end time point (76-100 weeks) as measured by Flortaucipir F18 PET scan (76 - 100 weeks)
To assess the PK of LY3372689	Plasma concentrations of LY3372689
Tertiary/Exploratory	
To assess the effect of LY3372689 vs placebo on blood-based biomarkers	Change from baseline through end time point (76-124 weeks) on blood-based biomarkers of AD pathology and neurodegeneration
To assess the effect of LY3372689 vs placebo on additional assessments of cognition	Change from baseline through end time point (76-124 weeks) as measured by: • Digital Clock Drawing Test • DSST
To explore the relationship between Tau burden and clinical outcomes	Tau burden • Blood-based biomarkers, and • Flortaucipir PET Clinical outcomes • ADAS-Cog13 • ADCS-iADL • CDR-SB • iADRS • MMSE
To explore the relationship between blood-based biomarkers and imaging	Blood-based biomarkers of AD pathology and neurodegeneration Imaging • Flortaucipir PET • MRI
To assess the efficacy of LY3372689 to prolong time to clinical progression	CDR global score

Abbreviations: ADAS-Cog$_{13}$ = Alzheimer's Disease Assessment Scale – Cognitive subscale; ADCS-iADL = Alzheimer's Disease Cooperative Study – instrumental Activities of Daily Living; AD = Alzheimer's disease; AE = Adverse Event; CDR-SB = Clinical Dementia Rating Scale- Sum of Boxes; C-SSRS = Columbia-Suicide Severity Rating Scale; iADRS = integrated Alzheimer's Disease Rating Scale; ECG = electrocardiograms; MMSE = Mini Mental State Examination; MRI – Magnetic Resonance Imaging; PET = Positron emission tomography; PK = Pharmacokinetics.

a Based on common close design, participants final endpoint time will be between 76-124 weeks.

b Moderate tau pathology is defined as those who meet tau PET inclusion criteria for evidence of tau pathology consistent with Alzheimer's disease (Fleisher et al. 2020), but do not have widespread and high levels of cortical tau pathology.

c High levels of cortical tau pathology by flortaucipir PET is defined by the top quartile of quantitative standardized uptake value ratios in a population of amyloid positive Alzheimer's participants and cognitively normal older controls (Pontecorvo et al. 2019).

Primary Estimand

The primary estimand is the overall treatment effect of LY3372689 versus placebo as measured by percent slowing of disease using the integrated Alzheimer's Disease Rating Scale (iADRS), in all participants with early symptomatic Alzheimer's disease (AD) and moderate cortical tau pathology burden while remaining on treatment up to 124 weeks after randomization. The 'initiation of approved prescription AD medication that is a disease-modifying therapy' and 'discontinuation of LY3372689' are considered as intercurrent events and are both addressed by the intention-to-treat (ITT) principle unless otherwise specified. An ITT analysis is an analysis of data by the groups to which subjects are assigned by random allocation, even if the subject does not take the assigned treatment, does not receive the correct treatment, or otherwise does not follow the protocol. No other intercurrent events are considered.

Secondary Estimand(s)

The estimand for the key secondary objective is the overall treatment effect of LY3372689 versus placebo as measured by percent slowing of disease progression using iADRS, in all participants with early symptomatic AD with elevated tau pathology burden defined by flortaucipir PET imaging while remaining on treatment up to 124 weeks after randomization. The initiation of approved prescription AD medication that is a disease-modifying therapy is considered an intercurrent event. Estimands for other secondary efficacy endpoints (e.g., Mini-Mental State Examination [MMSE], Alzheimer's Disease Assessment Scale - Cognitive subscale 13 [ADAS-Cog$_{13}$], Clinical Dementia Rating Scale - Sum of Boxes [CDR-SB], and Alzheimer's Disease Cooperative Study - instrumental Activities of Daily Living [ADCS-iADL]) are defined in a similar manner when analyzed by Bayesian Disease Progression Model (DPM).

For all efficacy endpoints (e.g., iADRS, MMSE, ADAS-Cog$_{13}$, CDR-SB, and ADCS-iADL), tau pathology biomarkers as measured by flortaucipir F18 PET SUVr, and brain region volumes measured by vMRI, when analyzed by mixed model for repeated measures (MMRM), the estimand is the treatment effect difference, LY3372689 versus placebo, as measured by change from baseline to end point (76 – 124 weeks) on the endpoint of interest, in participants with early symptomatic AD with moderate and/or the full study population with elevated tau pathology while remaining on treatment up to 124 weeks after randomization.

1.2. Study Design

Study MTAE is a multicenter, randomized, double-blind, placebo-controlled, Phase 2 study of LY3372689 in participants with early symptomatic AD. Participants who meet entry criteria will be randomized in a 1:1:1 (1 LY: 1 LY: 1 PBO) randomization of the following treatment groups:

- LY3372689: low dose once daily
- LY3372689: high dose once daily, and
- Placebo.

The maximum total duration of study participation for each participant, including screening and the post-treatment follow-up periods, is up to approximately 137 weeks:

- Lead-In: complete any time prior to Complete Screening visit
- Complete Screening: -63 days to -1 day prior to randomization
- Double-Blind: 76 weeks or up to 124 weeks, and
- Follow-Up: approximately 4 weeks after completion of double-blind period.

This study will utilize a 76-week common close design. Under the common close design, all enrolled participants will remain on double-blind randomized treatment and complete assessments until the last enrolled participant that has not discontinued from treatment has achieved 76 weeks of assigned treatment. The maximum duration of treatment is 124 weeks.

1.3. Method of Assignment to Treatment

This is a double-blind study, with design to maintain blinding to treatment. To preserve the blinding of the study, a minimal number of Lilly personnel will see the randomization table and treatment assignments before the study is complete.

Participants who meet all criteria for enrolment will be assigned a study (patient) number at Visit 601 or Visit 1and randomized to double-blind treatment at Visit 2. Participants will be randomized to LY3372689 or Placebo in a 1:1:1 ratio. For between-group comparability, patient randomization will be stratified by investigative site and tau burden (moderate versus high), and by prior or current use of any approved disease-modifying therapy for AD.

Assignment to treatment groups will be determined by a computer-generated random sequence using an interactive web response system (IWRS). Randomization into one stratum may be discontinued at the discretion of the sponsor. LY3372689 will be made available in blister packaging and will be given by a blinded nurse or other qualified blinded personnel, as described in the pharmacy manual. Site personnel will confirm that they have located the correct packages by entering a confirmation number found on the label into the IWRS.

For the first approximately 30 participants, the IWRS will be programmed to guarantee balance between the arms for the first interim analysis for safety; this is referred to as the burn-in period. After the burn-in period, participant randomization will then follow the dynamic allocation (minimization) method of Pocock and Simon (1975) to balance the treatment arms using investigative site as a factor. This is to ensure balanced participant assignment between treatment arms within each site at the end of the study.

2. Statistical Hypotheses

The primary efficacy analysis will be a Bayesian DPM conducted on the primary outcome iADRS. The primary efficacy analysis population will be those with moderate levels of tau pathology at baseline. A critical success factor (CSF) will be established of the following form:

- Probability (at least 25 % slowing of disease progression with LY3372689 relative to placebo) > 0.6.

The Bayesian DPM will be utilized to calculate the posterior probability of at least 25% slowing. If the posterior probability exceeds the probability threshold for at least one LY3372689 versus placebo comparison, the trial will have been considered to have met its primary endpoint.

To control the Type I error at two-sided 0.05 level in the graphical testing scheme (Section 4.9.2.1), additional CSFs will be established for iADRS using Bayesian DPM. The methodological approach used to find the CSFs will be pre-specified in the SAP prior to unblinding and the exact CSFs will be documented in the CSR. The testing scheme that will be used to control the false positive rate for the primary endpoint and important secondary endpoints will be described in Section 4.9.2.1.

The hypotheses for important secondary objectives are defined in a similar manner that LY3372689 is superior to placebo with regard to:

- clinical progression in the study population with early symptomatic AD and evidence of elevated tau pathology, including both the moderate levels and/or the full study population, through end time point (76 – 124 weeks), as measured by both percent slowing and change from baseline on MMSE, ADAS-Cog$_{13}$, CDR-SB, and ADCS-iADL
- brain tau deposition as measured by flortaucipir PET scan to end point (76 – 100 weeks)
- brain region volumes measured by vMRI to end time point (76 – 124 weeks)

2.1. Multiplicity Adjustment

To control the family wise Type I error rate, the graphical testing scheme described in Section 4.9.2.1 will be used to provide a strong control of Type I error rate at level α=0.05 (Bretz et al. 2009, 2011) for the primary and important secondary endpoints.

3. Analysis Sets

For the purposes of analysis, populations are defined in Table 3.1 and Table 3.2. These tables also list the study measures that will be summarized and/or analyzed for each analysis set.

Table 3.1. Analysis Set

Analysis Set	Description
Entered	All participants who sign the ICF.
Enrolled/ITT	All entered participants who are randomized and assigned to study treatment, regardless of whether they take any doses of study treatment, or if they take the correct treatment. Participants will be analyzed according to the treatment group to which they were assigned.
Efficacy evaluable	all participants with a baseline measurement and at least 1 complete postbaseline measurement for any efficacy assessment[1]
Safety	All participants randomly assigned to IP and who take at least 1 dose of IP and who do not discontinue the study for the reason "Lost to Follow-up" at the first postbaseline visit.
Pharmacokinetic analysis	Same as Safety Population.

Abbreviations: ICF = informed consent form; IP = investigational product.

[1] efficacy assessment including ADAS-Cog$_{13}$, ADCS-iADL, CDR-SB, MMSE.

The use of ITT or efficacy evaluable analysis set depends on the type of dependent variable in the analysis model. The ITT analysis set will be used for analysis methods that use measurements from both baseline and each scheduled post-baseline visit (according to the Schedule of Activities [SoA]) during the treatment period as the dependent variable. The efficacy evaluable analysis set will be used for analysis methods that use change from baseline as the dependent variable. The safety analysis set is used to analyze the endpoints and assessments related to safety.

Table 3.2. Safety and Efficacy Measures Summarized and/or Analyzed for Each Analysis Population

Population	Variables Assessed
Entered	Listings
Randomized/ITT	Tables and listings for patient characteristics, baseline severity, and patient disposition
Efficacy evaluable	Tables, listings, and figures of the following: iADRS, CDR-SB, ADAS-Cog$_{13}$, ADCS-iADL, MMSE, Digit Symbol Substitution Test (DSST), digital clock drawing test, tau burden measured by blood-based biomarkers, brain tau deposition measured by flortaucipir F18 PET scan, volumetric MRI measurements.

Safety	Tables, listings, and figures of the following:compliance, adverse events, laboratory results, vital signs, weight, ECG, safety MRIs, CSSRS

4. Statistical Analyses

4.1. General Considerations

The protocol calls for a Data Monitoring Committee (DMC) charged with making decisions regarding participant safety and study futility. This analysis plan describes analyses for the final clinical study report and all interim analyses for the DMC. Statistical analysis of this study will be the responsibility of the sponsor or its designee.

As Study MTAE is a Phase 2 study, the appropriate estimand is the efficacy of LY3372689 being assessed under the paradigm of all participants taking study drug as intended. Two intercurrent events are considered: initiation or change to standard of care disease-modifying therapy and discontinuation of the study drug prior to completing the 76-124 weeks of treatment.

The primary analysis is to use a Bayesian DPM analysis of the iADRS to compare the overall percent cognitive and functional slowing/worsening between LY3372689 versus placebo at end point (76 – 124 weeks). The same analysis set will be analyzed using Mixed-Effect Model Repeated Measure (MMRM) to serve as the sensitivity analysis. Similarly, MMRM will serve as the sensitivity analysis for each analysis of clinical progression that assessed via the DPM.

All efficacy analyses will follow the intention-to-treat (ITT) principle unless otherwise specified. An ITT analysis is an analysis of data by the groups to which participants are assigned by random allocation, irrespective of actual treatment received or participant's compliance with the protocol. even if the participant does not take the assigned treatment, does not receive the correct treatment, or otherwise does not follow the protocol. Consistent with ITT studies the estimand for Study MTAE is the Treatment Policy estimand.

With the 76-week common close design and given the backloaded enrollment progress in reality, the observations collected during the common close period such as Week 88 and onward could be too sparse for a proper data interpretation, especially for repeated measures models such as MMRM and natural cubic spline (NCS) models. The 'sparse' at a given timepoint/visit is defined as less than 10 datapoints per study arm. Unless otherwise noted, the full analyses set in terms of analysis visits for all longitudinal analysis models is data up to the timepoint where the post Week 76 visit that meets the requirement of > 10 datapoints per study arm at a given timepoint/visit in moderate tau population and the same will apply to full study population. Observations collected beyond *this* post Week 76 visit may be included for exploratory purpose.

For clinical efficacy endpoints (cognitive and functional scales) and biomarker endpoints assessed by repeated measures and disease progression models, observations collected at nonscheduled visits will not be included in the analyses unless the nonscheduled visit coincided with a protocol scheduled visit (Anderson and Millen, 2013). For analyses using last observation carried forward (LOCF), the last non-missing postbaseline observation (scheduled or unscheduled) will be used to calculate change from baseline.

Unless otherwise noted, baseline is defined as the last measurement prior to the initiation of the first dose, and endpoint is the last non-missing postbaseline measurement within the time period for the given analysis. Baseline definitions for safety analyses is described in Table 4.1. When

change from baseline is assessed, subjects will only contribute to the analysis if both a baseline and a postbaseline measurement are available.

Unless otherwise noted, all pairwise tests of treatment effects will be conducted at a 2-sided alpha level of 0.05 (or a 1-sided 0.025 alpha level if appropriate); 2-sided confidence intervals (CIs) will be displayed with a 95% confidence level. All tests of interactions between treatment and other factors will be conducted at an alpha level of 0.05.

Unless otherwise noted, no formal statistical hypothesis testing will be made for the safety interim analyses. Any reported p-values that may be reported as part of standardized output are merely for information purposes only.

Bretz's graphical approach will be utilized to provide strong control of the study-wise Type I error rate for the primary and secondary hypotheses of clinical efficacy endpoints at an alpha level of 0.05 (Bretz et al. 2009, 2011). Details on the graphical approach and testing strategy will be specified in Section 4.9.1.1.

If any of the individual items for ADAS-Cog$_{13}$ or ADCS-ADL or CDR-SB are missing or unknown, every effort will be made to obtain the score for the missing item or items. If either ADAS-Cog$_{13}$ or ADCS-iADL is missing, iADRS score will be considered missing. For all other scales, if any item is missing, any total or sum involving that item will be considered missing. For all scales, if a limited number of individual test items are missing, imputation using prorating will be implemented to obtain a total score.

A database lock is expected to occur after all randomized participants have had a chance to complete their final study visit in the double-blind period of the study (Period 2a/2b). Efficacy and safety analyses will be conducted based on data collected during the double-blind period. Visit 801 is the safety follow-up visit where concomitant medications, AEs, physical evaluation (e.g., vital signs, ECG, etc.), laboratory tests, and C-SSRS are collected. Data collected during the safety follow-up visit (V801) will be summarized and analyzed accumulatively unless otherwise specified.

Any change to the data analysis methods described in the protocol will require an amendment only if it changes a principal feature of the protocol. Any other change to the data analysis methods described in the protocol and the justification for making the change will be described in the SAP and the clinical study report (CSR). Additional exploratory analyses of the data will be conducted as deemed appropriate.

Table 4.1. Baseline Definition for Safety Groups

Analysis Type	Baseline
1.1) Treatment-Emergent Adverse Events	The baseline period is defined as the start of screening and ends prior to the first dose of study treatment (typically at Week 0).
1.2) Treatment-Emergent Abnormal Labs, Vital Signs, and ECGs.	Baseline will be all scheduled and unscheduled measurements recorded during the baseline period as defined above (1.1).
1.3) Change from Last Baseline to Week *xx* and to Last Postbaseline for Labs, Vital Signs, and ECGs.	The last scheduled non-missing assessment recorded prior to the date of first dose of study treatment during the baseline period defined above (1.1).

4.2. Handling of Dropouts or Missing Data

4.2.1. Handling Missing Data from Participant Dropouts

A likelihood-based mixed effects model for repeated measures will be used to handle missing data. The model parameters are simultaneously estimated using restricted likelihood estimation incorporating all of the observed data. Estimates have been shown to be unbiased when the missing data are missing at random and when there is ignorable non-random missing data.

Repeated measures and disease progression model analyses will only use data from visits where the data was scheduled to be collected (Andersen and Millen 2013). When patients discontinue from the study early, there may be efficacy or safety data measurements at visits where variables were not scheduled to be collected. This data will be used in all other analyses.

4.2.2. Handling Missing Items in Calculating Total Scores

If any of the individual items for ADAS-Cog or ADCS-ADL are missing or unknown, every effort will be made to obtain the score for the missing item or items.

For ADAS-Cog$_{13}$, if 3 or fewer of a total of 13 items are missing, the total score (maximum =85) will be imputed as follows: the total from the remaining items will be multiplied by a factor that includes the maximum score for the missing items. For example, if the first item, "Word-Recall Task," which ranges from a score of 0 through 10 (maximum = 10), is missing, and the second item "Commands," which ranges from a score of 0 to 5 (maximum = 5), is missing, then the multiplication factor = 85/(85 - [10 + 5]) = 85/70 = 1.21. Thus, the total score for this example will be the sum of the remaining 11 items multiplied by 1.21. The imputed number will be rounded up to the nearest integer. If more than 3 or fewer items are missing, the total score for ADAS-Cog$_{13}$ at that visit will be considered missing.

For the ADCS-iADL, if <30% of the items are missing, the total score will be imputed. The sum of the non-missing items will be prorated to the sum of total items. The imputed number will be rounded up to the nearest integer. If the nearest integer is greater than the maximum possible score, the imputed score will be equal to the maximum score. If >30% of the items are missing, the total score for ADCS-iADL at that visit will be considered missing. The same imputation technique will be applied to the ADCS-ADL total score. Note that, depending on the specific item responses that are missing, it is possible to have an imputed total score for both the ADCS-iADL and the ADCS-ADL, an imputed total score for one but not the other, or both total scores missing.

The iADRS score is calculated as follows: iADRS score = [−1(ADAS-Cog$_{13}$) + 85] + ADCS-iADL (Wessels et al. 2015). If either ADAS-Cog$_{13}$ or ADCS-iADL is missing, iADRS score will be considered missing.

For the CDR-SB score, if only 1 box (of 6) of the CDR is missing, the sum of the boxes will be imputed by prorating the sum from the other 5 boxes. If the score from more than 1 box is not available, the CDR-SB at that visit will be considered missing.

4.2.3. Handling Missing Date Information

For previous medications and medical history if parts of dates are missing, the following imputations will be performed:

- For start dates
    a. if DAY is unknown, it will be set to 01
    b. if MONTH is unknown, it will be set to JAN (01)
    c. if Day and Month are unknown, it will be set to 01 and JAN (01)
    d. if year is unknown, then start date is missing

- For end dates
    a. if DAY is unknown, it will be set to 30
    b. if MONTH is unknown, it will be set to DEC (12)
    c. if Day and Month are unknown, it will be set to 30 and Dec (12)
    d. if year is unknown, then end date is missing

4.3. Multicenter Studies

This study will be conducted by multiple investigators at multiple sites internationally.

In the event that any investigator has an inadequate number of subjects (defined as 1 or 0 randomized subjects per treatment group) for the planned analyses, the following strategy will be implemented. Data from all such investigators will be pooled. The pooling will be done first within a country. If the resulting pool within a country is still inadequate (1 or 0 randomized subjects to 1 or more treatment arms), no further pooling will be performed. The pooled site variable for each patient will be the country code if site was pooled; otherwise, it will be the site number. A listing including country, investigator site with address, number of participants enrolled (randomized) by each site, and unique participant IDs will be presented.

4.4. Participant Dispositions

Because this is a long-term study in a participant population that is elderly with multiple comorbidities, participant withdrawal is of particular concern. Additional efforts will be undertaken to reduce participant withdrawals and to obtain information on participants who are initially categorized as lost to follow-up.

All participants who discontinue from the study as well as discontinue from the treatment phase will be identified, and the extent of their participation in the study will be reported. If known, a reason for their discontinuation will be given. The reasons for discontinuation will be collected when the participant's participation in the study ends and will be summarized by treatment group for all randomized participants. The percentage of participants discontinuing from each treatment group will be compared between groups using Fisher's Exact test. The comparisons will be done for the overall percentage of participants who discontinue and for select specific reasons for discontinuation.

The median time to discontinuation will also be compared between treatment groups using the Kaplan-Meier product limit estimator. For any-cause discontinuation as well as discontinuation

due to adverse event or death, comparisons of time to discontinuation will be conducted using the Kaplan-Meier product limit estimator and the associated log-rank test.

4.5. Participant Characteristics

Baseline characteristics will be summarized for the randomized population by treatment group, primary efficacy subpopulation, and full study population. Summaries will include descriptive statistics for continuous and categorical measures. Participant characteristics to be presented include:

- age, sex, race, ethnicity, population (Japanese, non-Japanese)
- age group: 60 to 64, 65 to 69, 70 to 74, 75 to 80, >80
- ApoE ε4 carrier status non-carrier: ε3/ε3, ε2/ε2, ε3/ε2; carrier: ε2/ε4, ε3/ε4, ε4/ε4 [homozygote [ε4/ε4]; heterozygote [ε2/ε4, ε3/ε4))
- ApoE ε4 genotype (ε2/ε4, ε3/ε4, ε4/ε4, no ε4)
- height, body weight, body mass index (weight (kg) / [height (m)]^2)
- years of education, employment status, marital status
- tobacco use, alcohol use, caffeine use
- time since onset of first AD symptoms, time since diagnosis
- family history of AD, having 1 or more first degree relatives with AD
- the use of approved AD disease modifying treatment at baseline (yes/no)
- the use of AChEIs/memantine at baseline (yes/no)
- tau PET burden: the baseline tau PET categories (intermediate, high) as defined by the tau SUVr from the composite neocortical region multi-block barycentric discriminant analysis (MUBADA) at screening
- blood-based biomarkers: p-tau 217, GFAP, NfL.
- cognitive and functional scales: baseline severity of impairment as measured by CDR-SB, CDR-GS (CDR global score), ADAS-Cog$_{13}$, ADCS-ADL total score and instrumental (ADCS-iADL) and basic subscores (ADCS-bADL), MMSE, DSST, and Digital Clock Drawing Test (DCTClock).
- Clinical staging as defined by MMSE score at screening: Mild AD (22-26, inclusive), MCI (27-30, inclusive).

Baseline characteristics and baseline severity will also be listed.

4.6. Prior and Concomitant Therapy

Prior medications are defined as those that stop before randomization (the day prior to the first administration of study drug, Visit 2). Concomitant medications are defined as those being taken on or after randomization (the day prior to the first administration of study drug, Visit 2). If the start or stop dates of therapies are missing or partial to the degree that determination cannot be made of whether the therapy is prior or concomitant, the therapy will be deemed concomitant. Additional concomitant therapy information is as described in Section 6.5 of the MTAE study protocol.

Prior and concomitant medications will be listed. A summary of concomitant medications will be presented as frequencies and percentages for each treatment group. Fisher's Exact test will be used to test for treatment differences between groups. A summary table will also be provided for concomitant anticholinergics that affect cognitive function and use of approved prescription AD medications.

Medications will be coded using the World Health Organization (WHO) drug dictionary.

4.7. Treatment Compliance

The compliance for the study medication at individual level will be evaluated over the entire double-blinded period at individual level. The proportion of participants who are significantly noncompliant as described in Section 6.4 of the MTAE study protocol will be summarized and compared among all treatment groups using Fisher's Exact test.

4.8. Primary Endpoint Analysis

Study MTAE has a 76-week common close design where all enrolled participants continue on study drug and complete study assessments until last enrolled participant that has not discontinued reaches 76 weeks of double-blind treatment. The maximum duration of treatment is 124 weeks. The analysis in this section will be conducted in the moderate tau study population.

4.8.1. Main Analytical Approach

The primary objective of this study is to test the hypothesis that LY3372689 slows cognitive and functional decline in early symptomatic AD participants with moderate cortical tau burden as measured by iADRS compared with placebo through end time point (76-124 weeks). This will be assessed using a DPM, evaluating possible slowing of disease progression with treatment of LY3372689 relative to placebo.

The main purpose of the DPM is to estimate a quantity known as the disease progression ratio (DPR), which measures the proportion of disease progression in active treatment arms relative to placebo arm. The iADRS score at baseline and each scheduled postbaseline visit (according to SoA) during the treatment period will be the dependent variable. The DPM is as follows:

$$Y_{ij} = 1_{j>0} \cdot e^{\theta_{T_i}} \sum_{v=1}^{j} \alpha_v + X_i\beta + \gamma_i + \varepsilon_{ij}, i = 1, 2, \ldots k; j = 0,1, 2, \ldots, l$$

where Y_{ij} denotes the clinical outcome at visit j for participant i, clinical outcome scores at baseline (prior to treatment) is Y_{i0}. The term $1_{j>0}$is the indicator function that is 0 for $j = 0$ (baseline visit) and 1 otherwise. For post-baseline visits, the multiplication term $e^{\theta_{T_i}}$ is the DPR associated with the treatment group of participant i, $T_i \in \{\text{Placebo}, \text{LY}_{\text{low}}, \text{LY}_{\text{high}}\}$. $\theta_{\text{Placebo}} = 0$, $e^{\theta_{\text{LY}_{\text{low}}}}$ and $e^{\theta_{\text{LY}_{\text{high}}}}$ represent the DPR of the low dose and the high dose in relative to placebo, respectively. α_v $(v \geq 1)$ denotes the mean change of placebo response from visit $v-1$ to visit v. The parameter γ_i represents a subject-specific random intercept for participant i which is assumed to center the baseline main effect $\gamma_i \sim N(\mu_{BL}, \sigma_{BL}^2)$, where the parameter μ_{BL} models the mean baseline score across arms. ε_{ij} is the error variance which is assumed to follow $\varepsilon_{ij} \sim$

$N(0, \sigma^2)$. X_i denotes subject specific baseline covariates that will be included in the model and β represent the associated fixed effects. Stratification variables used for randomization such as investigator site (pooled) will be included in the model. Baseline covariates such as age and concomitant AChEI/memantine use at baseline (yes/no) will also be included in the model.

The DPM model assumes the treatment effect of LY3372689 is proportional to placebo over the course of the study. This proportionality assumption is similar to the assumption made in proportional hazards modeling of time to event data. The DPM includes generally diffuse priors on all parameters in the model. The precision value for the prior distribution of all parameters is set to a small value; therefore, the prior distributions on all parameters have very little impact on the posterior distributions. No information or knowledge of the effect of LY3372689 from previous studies will be incorporated into the prior distributions, and the inference will be based on Study MTAE only. The specifications of the prior of the model parameters are as following:

A diffuse normal prior is assumed for μ_{BL} and a diffuse inverse gamma prior is considered for the variance of the participant-level random intercept σ_{BL}^2 :

$$\mu_{BL} \sim N(0, 10^4), \sigma_{BL}^2 \sim IG(0.005,0.005).$$

The same diffuse normal prior distribution is assumed for each change in the placebo mean disease progression from visit $v-1$ to visit v, α_v:

$$\alpha_v \sim N(0, 10^4)$$

A prior distribution of $N(0,0.25)$ is assumed for the log(DPR) parameter of each active treatment arm (LY_{low} and LY_{high}).

A diffuse prior of the following form is also assumed for β, the fixed effect of the baseline covariates.

$$\beta \sim N(0,10^4)$$

An inverse gamma prior $\sigma^2 \sim IG(0.005,0.005)$ is assumed.

A DPR less than 1 corresponds to a slowing of disease progression (i.e., favoring LY3372689); similarly, a DPR greater than 1 reflects some evidence of cognitive worsening (i.e., favoring the placebo). The null hypothesis is that the DPR of both active treatment arms equal to 1.

A DPR less than 1 corresponds to a slowing of disease progression (i.e., favoring LY3372689); similarly, a DPR greater than 1 reflects some evidence of cognitive worsening (i.e., favoring the placebo). The null hypothesis is that the DPR of both active treatment arm equal to 1.

The Bayesian inferences will be summarized including posterior distribution of DPR and posterior probabilities of meeting various DPR thresholds of interest (e.g., 0.75 which translates to 25% slowing of disease progression with LY3372689 group versus placebo). Bayesian posterior probability of at least one active treatment arms being superior to placebo by at least a margin of interest (change from baseline in iADRS score between treatments equivalent to a 25% slowing of disease progression relative to placebo) will also be calculated.

4.8.2. Sensitivity Analyses

The iADRS will be analyzed by mixed model with repeated measures (MMRM) and natural cubic splines (NCS) (Donhue et al., 2023) as the sensitivity analyses. For MMRM and NCS analyses, treatment effect based on data up to WK76 and based on data up to WK88 will be provided, respectively. For the latter, the 'overall' treatment effect averaging across Week 64, Week 76, and Week 88 will be calculated. Additional timepoints may be included for exploratory purpose.

4.8.2.1. Natural Cubic Spline (NCS) Analysis

A NCS model with 2 degrees of freedom (NCS2) (Donohue et al. 2023) will be used to estimate the difference in the change in iADRS score between the treatment arms. The NCS2 model has 3 knots over the observation time will be placed: 2 at the boundaries (minimum and maximum observation time), and 1 internal knot at the median observation time. The baseline estimates are restricted to be the same for all treatment arms. The model will be estimated using restricted maximum likelihood method (RMLE).

The iADRS score at baseline and at each of the scheduled post-baseline visits (according to SoA) will be included in model as a dependent variable. The analysis model will include the fixed effects of NCS basis expansion terms (two terms), NCS basis expansion term-by-treatment interaction (two terms), age at baseline, AChEI/memantine use at baseline (yes/no), and investigator site (pooled). Visit will be treated as a continuous variable with values equal to weeks between baseline and post-baseline exam dates, and the NCS basis function will be derived using these visits in weeks. An unstructured variance-covariance structure matrix will be used to model within-subject variance-covariance of residual errors across visits where visits are modeled as a categorical variable. If the unstructured variance-covariance structure matrix results in a lack of convergence, the following structures will be used in sequence:

- Heterogeneous Toeplitz covariance structure
- Heterogeneous autoregressive order 1 covariance structure
- Heterogeneous compound symmetry covariance structure, and
- Compound symmetry covariance structure

Mean change from baseline values, and the comparisons between change from baseline values by treatment arms will be estimated through the proper contrast set up. The Kenward-Roger approximation will be used to estimate the denominator degrees of freedom.

Percent slowing comparing to placebo group will be calculated as the least-squares (LS) estimates of differences in change from baseline between treatment groups, then divided by the LS estimates of mean change from baseline value from placebo group. A 95% CI for this percent slowing is calculated based on a Delta method (Beyene et al. 2005).

4.8.2.2. Mixed Model with Repeated Measures (MMRM) Analysis

The change from baseline score on the iADRS at each scheduled post-baseline visit (according to the SoA) during the treatment period will be the dependent variable. The analysis model includes the fixed effects of treatment, visit as a categorical variable, baseline score, baseline-by-visit

interaction, treatment-by-visit interaction, and investigator site (pooled). Baseline covariates such as age and concomitant AChEI/memantine use at baseline (yes/no) will be included in the model. An unstructured covariance matrix will be used to model the within-subject variance-covariance of residual errors across visits. If the unstructured covariance structure matrix results in a lack of convergence, the following tests will be used in sequence:

- Heterogeneous Toeplitz covariance structure
- Heterogeneous autoregressive order 1 covariance structure
- Heterogeneous compound symmetry covariance structure, and
- Compound symmetry covariance structure

The Kenward-Roger approximation will be used to estimate the denominator degrees of freedom.

The treatment group contrast in LS mean progression and its associated p-value and 95% CI will be calculated for the treatment comparison of LY3372689 versus placebo using the MMRM model specified above.

4.8.3. Supplementary Analyses

4.8.3.1. Time Progression Models for Repeated Measures (Time-PMRM) Analysis

Time-PMRM (Raket 2022) will be used to evaluate the possible slowing of the disease progression, as measured by time, with the treatment of LY3372689 in relative to placebo. The null hypothesis is that there is no slowing of the time progression of the disease in active treatment arms as compared to the placebo arm. The testing will be done by likelihood ratio test (LRT).

The iADRS score at baseline and at each of the scheduled post-baseline visits (according to the SoA) will be included in model as a dependent variable. The analysis model will include age at baseline AChEI/memantine use at baseline (yes/no), and investigator site (pooled). Weeks since randomization at each planned visit will be included as a continuous variable. The intercepts are constrained to be the same across treatment arms. An NCS with internal knots at each planned visit will be used to interpolate the disease progression between the planned visits for the placebo arm, and the trajectories of the active treatment arms will be estimated with the assumption that the mean disease progression of the active treatment arms at a given visit can be estimated by the mean disease progression of the placebo arm at another time point. The assumption of proportional time slowing will be tested. If the proportional time slowing assumption is not met, a model similar to the above but having individual time slowing parameters estimated separately at each post-baseline visit will be fitted. Model parameters will be estimated using maximum likelihood estimation (MLE).

4.8.3.2. Delta Adjustment Tipping Point Analysis

Sensitivity to departures from the MAR assumption in the MMRM analyses will be investigated using a tipping point analysis (Carpenter & Kenward 2013). This method is a sensitivity analysis in multiple imputation (MI) under the missing-not-at-random (MNAR) assumption that searches

for a tipping point that reverses the study conclusion. Departures from MAR in the active treatment arms will be assessed assuming that participants who discontinue the study have, on average, efficacy outcomes after discontinuation that are worse by some amount δ compared to other similar participants with observed data (i.e., compared to a value which would have been assumed under an MAR model). A series of analyses will be performed with increasing values of δ until the analysis conclusion of a statistically significant treatment effect no longer holds. The value of δ that overturns the primary results will represent a tipping point. An interpretation of clinical plausibility of the assumption underlying the tipping point will be provided.

Mean changes from baseline in iADRS scores will be analyzed based on data observed while the participant remains on study as well as data imputed using MI methodology for time points at which no value is observed. Imputed values in the active treatment arms will first be sampled from an MAR-based MI model and then δ-adjusted as described below.

Missing-at-random-based imputations will be generated for iADRS scores at each time point, and then a value of $\delta = \{\Delta\}$ will be added to all imputed values in the active treatment arms prior to analyzing multiply imputed data. This approach assumes that the marginal mean of imputed participant measurements is worse by δ at each time point after discontinuation compared to the marginal mean of participants with observed data at the same time point. Analyses will be conducted with values of δ starting from 0 with increments of 0.10 until the null hypothesis can no longer be rejected.

4.9. Secondary Endpoint(s) Analyses

The analyses described in Section 4.9.1 and Section 4.9.2 will be conducted in moderate tau and full study populations separately with only analyses for moderate tau population being included in the graphical testing scheme described in Section 4.9.2.1.

4.9.1. Key Secondary Efficacy Endpoint Analysis

The key secondary objective of this study is to test the hypothesis that LY3372689 slows cognitive and functional decline in early symptomatic AD participants with elevated cortical tau burden as measured by iADRS compared with placebo through end time point (76 -124 weeks). The model setup will be the same as described in Section 4.8.1 and Section 4.8.2, with the inclusion of the tau burden (moderate versus high) as defined by flortaucipir PET in the model as a baseline covariate. The time-PMRM (Section 4.8.3.2) analysis may be conducted to assess the treatment effect from time-saving perspective.

4.9.2. Other Secondary Efficacy Endpoints Analyses

Other secondary efficacy endpoints include CDR-SB, ADAS-Cog13, ADCS-iADL, and MMSE.

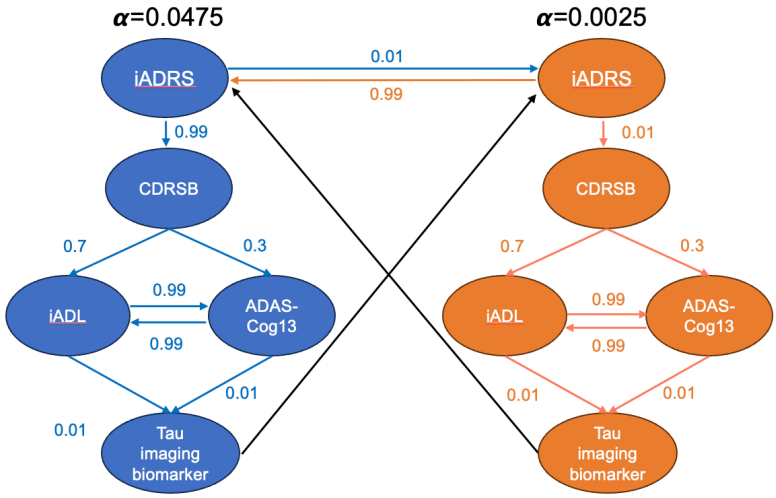
The key secondary objective of this study is to test the hypothesis that LY3372689 slows cognitive and functional decline as measured by CDR-SB compared with placebo at end time point (76-124 weeks) in full study population and moderate tau sub population with early symptomatic AD. The main analytical approach to secondary efficacy endpoints is NCS2 as described in Section 4.8.2.1. The treatment effect based on data up to WK76 and based on data

up to WK88 will be provided, respectively. For the latter, the 'overall' treatment effect averaging across Week 64, Week 76, and Week 88 will be calculated.

The secondary efficacy endpoints will also be analyzed by Bayesian DPM (Section 4.8.1) in both moderate tau subpopulation and the full study population, respectively. The MMRM (Section 4.8.2.2) and time-PMRM (Section 4.8.3.2) analyses may be conducted as supplementary analysis.

4.9.2.1. Gated Secondary Efficacy Analyses

Bretz's graphical approach will be utilized to provide strong control of the study-wise Type I error rate for the primary and secondary clinical efficacy hypotheses at alpha level of 0.05 (Bretz et al. 2009, 2011) in the moderate tau population. The statistical significance, outlined using alpha levels in a frequentist testing paradigm, will be determined based on the following multiplicity graph of hypotheses regarding the LY3372689 slowing the cognitive and/or functional decline and tau SUVr level of AD with demonstrated presence of moderate levels of tau pathology (the blue ovals represent the high dose arm and the ornage ovals represent the low dose arm):



iADRS = DPM analysis of iADRS
CDR-SB = NCS2 analysis of CDR-SB average across Week 64 to Week 88.
iADL = NCS2 analysis of iADL average across Week 64 to Week 88.
ADAS-Cog13 = NCS2 analysis of ADAS-Cog13 average across Week 64 to Week 88.
Tau imaging biomarker for brain tau deposition = MMRM analysis of tau imaging biomarker endpoint AD-signature volume of interest (VOI) using CereCrus as the reference region as measured by flortaucipir F18 PET SUVr value at Week 76.

The Bayesian DPM will be used to analyze the primary endpoint of iADRS, the NCS2 model will be utilized for secondary clinical efficacy endpoints including CDR-SB, iADL, ADAS-Cog13, and MMRM will be used to analyze the tau imaging biomarker endpoint AD-signature VOI using CereCrus as the reference region for brain tau deposition as measured by flortaucipir F18 PET standard uptake value ratio (SUVr) value.

The frequentist analyses for the clinical endpoints will utilize the endpoint averaging across Week 64, Week 76, and Week 88. The tau imaging biomarker endpoint will be analyzed based on the LS mean change from the baseline at Week 76.

As mentioned in Section 2, additional CSFs for iADRS using Bayesian DPM will be established to control the Type I error at two-sided 0.05 level in the graphical testing scheme (Section 4.9.2.1). To test the hypothesis of a disease progression benefit as measured by iADRS using the Bayesian DPM, the posterior probability of superiority in cognitive/functional slowing for each active treatment arm will be calculated. The superiority is defined by Pr(DPR < 0.75) > X, where the value of X is determined via simulation to control the test at the corresponding alpha level. If the posterior probability is above the threshold for at least one active treatment arm (which controls the overall false positive rate at approximately 2.5% one-sided), then a statistical significance claim of cognitive/functional slowing will be made. Success will be declared at a graphical node if the posterior probability of at least 25% slowing of disease progression as measured by the iADRS in high dose and the low dose is greater than a certain threshold:

- Probability (at least 25 % slowing of disease progression on iADRS in LY3372689 high dose relative to placebo) > probability threshold 1.
- Probability (at least 25 % slowing of disease progression on iADRS LY3372689 low dose relative to placebo) > probability threshold 2.

These two thresholds will be calculated after unblinding at PODBL to establish the accurate probability thresholds that correspond to each alpha level as shown in graphical testing scheme above. As the thresholds can be sensitive to the underlying assumptions of placebo trajectory and the variance-covariance matrix, the methodological approach to find the probability thresholds is pre-specified below. The established thresholds will be documented in the CSR.

1. The NCS2 results for the iADRS (Section 4.8.2.1) from the moderate tau population will provide the underlying assumptions that are needed for this simulation. The following will be utilized: the LS mean placebo trajectory and the number of participants with data from the placebo arm at each visit; the pooled variance covariance matrix as estimated from the NCS2.

2. To create a null distribution, simulate multivariate normal data for the number of participants from the moderate tau population representing the placebo arm, assuming all participants have the same decline as the placebo arm from the NCS model and the same variability. The amount of data simulated for each treatment arm will reflect the exact number of participants at each visit from the placebo arm.

3. The posterior probability of at least 25% slowing is calculated by fitting the Bayesian DPM to each 100,000 simulated dataset. The Bayesian DPM as described in Section 4.8.1 but without covariate adjustment. The threshold is the lowest value at which at least ($\alpha/2 \times 100$)% of the simulated posterior probabilities are larger. Note that the probability threshold 2 will be re-calculated if it makes all the way down to tau imaging biomarker in the high dose arm.

The tests for the rest of the graph: CDR-SB, iADL, and ADAS-Cog13 will be based on the p-value of the difference of averaged LSM change across Week 64 to Week 88 between the treatment arms; tau imaging biomarker endpoint AD-signature VOI will be based on the p-value of the difference of LSM change at Week 76 between the treatment arms.

4.9.3. Other Secondary Endpoints Analyses

4.9.3.1. Analysis of Volumetric MRI

Analyses of the following volumetric MRI (vMRI) parameters will be conducted (right + left for all but whole brain volume and ventricular volume):

- Hippocampal volume (mm^3)
- Atrophy of total whole brain volume (cm^3)
- Enlargement of Ventricular volume (cm^3)

To evaluate the changes in vMRI parameters, an MMRM model will be used to compare change from baseline to Week 76 between the treatment arms in the Efficacy Evaluable analysis set within moderate tau population and full study population, respectively.

The change from baseline at each post-baseline visit up to Week 76 will be the dependent variable. The model will include the fixed effects of treatment, visit as a categorical variable, treatment-by-visit interaction, and adjusted for baseline vMRI value and age at baseline. Baseline tau PET category will also be included as a fixed effect to the model when analyzing full study population. The null hypothesis is that the difference in LS mean between both the active treatment arm and placebo equal zero. Analyses including post-Week 76 visits may be included for exploratory purpose.

To assess the relationship of vMRI with efficacy endpoints up to Week 76, Spearman's rank correlation coefficient may be obtained on baseline and change from baseline on vMRI parameters versus change from baseline for iADRS, CDSR-SB, $ADAS\text{-}Cog_{13}$, ADCS-ADL, and MMSE. This will be performed using participants who have both the clinical outcome and vMRI result at Week 52 and Week 76, respectively. This analysis may be performed in moderate tau subpopulation and/or the full study population. The correlation analysis including post-Week76 timepoints such as Week 100 and Week 124 may be conducted for exploratory purpose.

4.9.3.2. Analysis of Flortaucipir PET Scan

To evaluate the change from baseline in tau imaging parameters, ANCOVA will be used to compare change from baseline in SUVr at Week 76 in the Efficacy Evaluable analysis set. The tau SUVr values from the 4 composite regions: frontal, parietal, lateral occipital, lateral temporal will be evaluated with CereCrus (cerebellar crusteneous) as the reference region. In addition, AD-signature volume of interest (VOI) using CereCrus as the reference region will be evaluated. AD-signature VOI was also known as multi-block barycentric discriminant analysis, MUBADA (Kotari et al., 2023, Shcherbinin et al., 2023) .

The tolerance window to remap the tau PET scan to the nearest scheduled scan is +/- 28 days. This remapping rule will be implemented to scans that are taken as the ET visit Flortaucipir F18 PET scan when applicable.

The change from baseline on the tau imaging parameters will be the dependent variable. The ANCOVA model will be adjusted by baseline SUVr value and age at baseline. Baseline tau PET category will be included as a fixed effect to the model applied to full study population. The null hypothesis is that the difference in LS mean change between both the active treatment arm and placebo equals zero. MMRM analysis incorporating tau PET scans at post-Week 76 timepoints may be included for exploratory purpose.

To assess the relationship of tau PET imaging biomarkers with efficacy endpoints at Week 76, Spearman's rank correlation coefficient may be obtained on baseline and change from baseline to Week 76 for the SUVr versus the change from baseline to Week 76 for iADRS, CDR-SB, ADAS-Cog_{13}, ADCS-ADL, and MMSE. This will be conducted using participants who have both the clinical outcome and SUVr result at Week 76, and will be analyzed by treatment group as well as pooling all treatment groups. This analysis may be performed in moderate tau subpopulation and/or the full study population. This correlation analysis including post-Week76 timepoints may be conducted for exploratory purpose.

4.9.4. Pharmacokinetic Analyses

LY3372689 plasma concentrations will be illustrated graphically and summarized descriptively.

If warranted and based on availability of data, the exposure-response relationship of LY3372689 concentrations to biomarkers, pharmacodynamic endpoints, efficacy endpoints, and/or safety endpoints may be explored. Data from other clinical studies evaluating LY3372689 may be combined with data from this study to support exposure-response analyses. Such analyses may be reported separately.

4.10. Safety Analyses

All safety analyses will group subjects based on the actual treatment received. If a patient receives a treatment different than the randomized treatment for the duration of the placebo-controlled period, then that different treatment is the actual treatment; otherwise, the actual treatment is the planned treatment.

4.10.1. Extent of Exposure

Days of exposure will be calculated for each participant:

date of last dose – date of first dose + 1.

Summary statistics will be provided for the total number of days and participant-years of exposure by treatment. Study drug treatment assignment will be listed.

4.10.2. Adverse Events

Treatment-emergent adverse events (TEAEs) will be defined as events that occurred or worsened between the first dose (treatment start date) and two days post last dose (treatment end date), inclusive. Should there be insufficient data for AE start date, stop date, and time to make this

comparison, the AE will be considered treatment-emergent. The Medical Dictionary for Regulatory Activities (MedDRA) lower-level term (LLT) will be used in the treatment-emergent computation. The maximum severity for each LLT during the baseline period will be used as baseline.

An overview of AEs, including the number and percentage of participants who died, suffered serious adverse events (SAEs), discontinued due to AEs and who suffered TEAEs, will be provided. Comparison between treatments will be performed using Fisher's Exact Test.

Summaries of AEs by decreasing frequency of PT within SOC will be provided for the following:

- Pre-existing conditions
- TEAEs
- TEAEs by maximum severity
- TEAEs occurring in greater than or equal to 2% of participants by PT
- SAEs
- Adverse events reported as reason for study treatment discontinuation
- Adverse events of special interest (AESI) include, but not limited to, the following:
  - CCI
  - CCI
  - CCI

These summaries will include number and percentages of participants with TEAEs. Treatment comparisons will be carried out using Fisher's Exact Test. A listing of AEs will be generated. Separated adverse events summaries will be provided for treatment period and follow-up period, respectively.

The following will be used to identify AESIs:

- Cardiac arrythmias SMQ:
  - Arrhythmia related investigations, signs and symptoms SMQ (20000051, narrow terms).
  - Cardiac arrhythmia terms (incl bradyarrhythmias and tachyarrhythmias) SMQ (20000050, narrow and broad terms).
- Ovarian neoplasms, malignant and unspecified SMQ,
  - Ovarian malignant tumours SMQ (20000200, narrow and broad excluding broad terms Inferior vena cava syndrome, Malignant ascites).
  - Ovarian tumours of unspecified malignancy SMQ (20000201, narrow and broad terms)/
- Lens disorders SMQ (20000155, narrow and broad terms).

- Glaucoma SMQ (20000146, only broad term Cataract).

4.10.3. Deaths, Other Serious Adverse Events, and Other Notable Adverse Events

An overview of AEs, including the number and percentage of participants who died or suffered SAEs during the study, discontinued due to AEs and who suffered TEAEs, will be provided. Comparison between treatments will be performed using Fisher's Exact Test.

4.10.4. Clinical Laboratory Evaluation

Laboratory measurements will be analyzed using continuous data (change from baseline) and categorical or ordinal data (proportion of treatment-emergent abnormalities). If there are multiple records of laboratory measurements at baseline or postbaseline visit, the last record will be used. Summaries and analyses of continuous data (change from baseline) will be performed using both conventional and International System of Units (SI units) and include scheduled visits only. Listing will include data collected at unscheduled visits.

Measures of central tendency for laboratory measurements and change from baseline will be summarized with boxplots. Boxplots will display results quarterly for the first 24 weeks and semi-annually for the rest of the study period (Visit 5, 8, 15, 21, 25, 29, if applicable) and for the last visit (LOCF). Summary tables of N, mean, median, quartiles, min, max, standard deviation, and p-value (for change scores) will be included. If there are considerable missing visits, the measures of central tendency may be based on MMRM analyses.

Change from baseline to postbaseline visit at which laboratory measurements are taken will be compared between treatment groups using an MMRM model on the Safety Dataset. For each lab analyte, the rank-transformation will be applied to the change from baseline for all participants and all visits prior to analysis. Similarly, an independent rank-transformation will be applied to the baseline values prior to analysis. The model for the fixed effects will include the following terms: categorical effects of treatment, visit as a categorical variable, and treatment-by-visit interaction as well as the continuous effects of ranked baseline value and age at baseline. This analysis will be done separately for each laboratory analyte.

Treatment differences in the proportion of participants with treatment-emergent high or treatment-emergent low or treatment-emergent abnormal laboratory values at (1) anytime and (2) each postbaseline visit will be assessed using Fisher's Exact test. Treatment-emergent high or low laboratory abnormality will be based on SI unit. For each laboratory analyte, only participants who were low or normal at all baseline assessments and have at least 1 postbaseline will be included in the denominator when computing the proportion of participant with treatment-emergent high. Similarly, only participants who were high or normal at all baseline assessments and have at least 1 postbaseline will be included in the denominator when computing the proportion of participant with treatment-emergent low. In addition, treatment differences in the proportion of participants who have normal baselines with a change to abnormal high or abnormal low values at any postbaseline visits will be summarized.

For urinalysis parameters, baseline to postbaseline shifts will be summarized at each visit. Likelihood ratio chi-square tests will be used to compare increase, no change, and decrease shifts in urinalysis parameters between treatment groups at each visit.

For hormone panel parameters (i.e., follicle-stimulating hormone (FSH), luteinizing hormone (LH), Inhibin B, Testosterone, Estrogen), baseline to postbaseline values as well as change from baseline at each post-baseline visits will be summarized.

For all laboratory analytes, frequencies of participants with notable changes (i.e., increases or decreases of a pre-specified amount unique to each analyte) from baseline to each postbaseline visit were also summarized for all participants and stratified by low, normal, or high at baseline.

The proportion of subjects with treatment-emergent clinically significant changes from a low value or normal value at all baselines at any time in alanine aminotransferase (ALT), alkaline phosphatase (ALP), and total bilirubin will be summarized by treatment group. Clinically significant changes of interest for hepatic laboratory values at any time are: ALT or AST ≥3 x upper limit of normal (ULN), total bilirubin ≥2 x ULN, ALP ≥2 x ULN, ALT ≥5 x ULN, ALT ≥10 x ULN, and total bilirubin ≥2 x ULN. Additionally, Hy's Law analysis will be conducted by comparing treatment groups with regard to the proportion of participants with (ALT ≥3 x ULN OR AST ≥3 x ULN) AND total bilirubin ≥2 x ULN. Comparisons between treatment groups will be carried out using Fisher's Exact test. When criteria are met for hepatic evaluation and completion of the hepatic safety case report form (CRF), investigators are required to answer a list of questions pertaining to the participant's history, relevant preexisting medical conditions, and other possible causes of liver injury. A listing of the information collected on the hepatic safety CRF will be generated.

4.10.5. Vital Signs and Other Physical Findings

Vital sign measurements and weight will be analyzed using continuous data (change from baseline) and categorical data (proportion of potentially clinically significant changes) using the safety dataset. Summary will include data collected at scheduled visits only. Data collected at unscheduled visits will be provided in the listing.

Summary statistics will be presented for observed values at baseline and for change from baseline results at each scheduled postbaseline visit. Sitting systolic and diastolic blood pressure and pulse (measurement after 5 minutes in the sitting position), orthostatic diastolic and systolic blood pressures, and orthostatic pulse (measurement after 5 minutes in the supine position minus that after 3 minutes in the standing position), temperature, and weight by treatment group for all participants in the safety population will be summarized. If there are multiple records of vital sign or weight measurements at baseline or postbaseline visit, the last record will be used.

Measures of central tendency for vital sign and weight raw measurements and change from baseline will be summarized with boxplots. Boxplots will display results quarterly for the first 24 weeks and semi-annually for the rest of the study period (Visit 5, 8, 15, 21, 25, 29) and for the last visit (LOCF). Summary tables of N, mean, median, quartiles, min, max, standard deviation,

and p-value (for change scores) will be included. If there are considerable missing visits, the measures of central tendency may be based on MMRM analyses.

With the large number of visits at which vital signs are scheduled to be measured, the MMRM model is not suitable for the change from baseline comparison of treatments due to computational challenges. Change from baseline to each postbaseline visit at which vital signs are taken will be assessed using an ANCOVA model with treatment as an independent factor and baseline value and age as covariates in the model. This analysis will be done separately for each vital sign parameter and weight.

The incidence of treatment-emergent abnormal high or low vital signs and weight will be presented by treatment group and visit. Treatment-emergent vital sign evaluations are defined for evaluations collected after the initiation of study medication. Abnormal criteria for postbaseline vital signs and weight are presented in Appendix 1. Any vital sign or weight meeting the criteria will be considered abnormal. Treatment differences in the proportion of participants with treatment-emergent abnormal high or low vital signs and weight will be assessed between treatment groups using Fisher's Exact test at (1) any time (2) postbaseline visit. Baseline will be considered as all measurements prior to the initiation of the first dose and all baseline assessments must not meet abnormality criteria to be included in the analysis.

For each vital sign at each postbaseline visit, only participants who had a baseline result and had a non-missing result at that postbaseline visit will be included in the denominator when computing the proportion of participants with treatment-emergent high, low, or abnormal values.

Summary and analyses of change from baseline in weight will be provided. The proportion of participants with a weight gain or loss of $\geq$7% of baseline body weight will be compared between treatment groups using Fisher's Exact test at each visit and at any time.

A listing of treatment-emergent abnormal vital signs and weight will also be presented.

4.10.6. Electrocardiograms

ECG measurements will be analyzed using continuous data (change from baseline) and categorical data (proportion of treatment-emergent abnormalities) using the safety dataset. Summary will include data collected at scheduled visits only. Data collected at unscheduled visits will be provided in the listing.

The ECG measurements are derived from three 10 second readings taken every 30 seconds. These 3 readings are to be averaged prior to analysis. Additionally, whenever ECG is measured in triplicate, the average of these readings will be used in the analysis. If there are multiple records after averaging ECG triplicates within a visit, the last record of averages will be used.

The analysis will be done for the following ECG measurements: heart rate, PR, QT, QTc, and RR intervals and QRS duration. All analyses of QTc will be carried out using the Fridericia correction (QTcF) method. Measures of central tendency for ECG raw measurements and change from baseline will be summarized with boxplots. Boxplots will display results quarterly for the first 24 weeks and semi-annually for the rest of the study period (Visit 5, 8, 15, 21, 25, 29) and

for the last visit (LOCF). Summary tables of N, mean, median, quartiles, min, max, standard deviation, and p-value (for change scores) will be included. If there are considerable missing visits, the measures of central tendency may be based on MMRM analyses.

Change from baseline to each postbaseline visit at which ECG measurements are taken may be assessed using an MMRM model. The model will include the fixed effects of treatment, visit as a categorical variable, and treatment-by-visit interaction as well as continuous effects of baseline ECG score and age at baseline. This analysis will be done separately for each ECG parameter.

QTc assessment based on C-QTc modeling evaluating QTc effects using MMRM may be conducted if deemed appropriate (Garnett C. et. al. 2017). The analysis model includes the fixed effects of treatment, concentration, time (nominal time), baseline QTcF as covariate, and subject and subject-by-concentration interaction as random effects. An unstructured covariance matrix will be used to model the within-subject variance-covariance of residual errors across visits. The handling of the lack of convergence with the unstructured covariance structure matrix is the same as described in Section 4.8.2.2. The Kenward-Roger approximation will be used to estimate the denominator degrees of freedom. Each dose arm effect, its relative effect to placebo on QTcF, and their corresponding 90% CI will be estimated via an appropriate linear combination of the coefficients using the geometric mean of the Cmax at the corresponding dose. Note the placebo concentration is zero. The same model may be applied to evaluate the relationship of the concentration and change from baseline in heart rate, PR interval, and QRS duration.

Incidence of treatment-emergent abnormal ECGs will be assessed by comparisons at (1) anytime and (2) each postbaseline visit between treatment groups with Fisher's Exact test. For analyses of treatment-emergent abnormal ECGs, baseline will be considered as all measurements prior to the initiation of the first dose. Abnormal ECG criteria and criteria for abnormal QTcF prolongation are presented in Appendix 2.

Treatment-emergent high ECG parameters (heart rate, PR interval, QRS duration, QT and QTcF intervals) are the values which are low or normal at all baseline visits and fall into the high abnormal categories postbaseline. Similarly, treatment-emergent low ECG parameters (heart rate, PR interval, QRS duration) are the values which are high or normal at all baseline visits and fall into the low abnormal categories above.

In addition, treatment differences in the proportion of participants who have normal baselines with a change to abnormal high or abnormal low values at any postbaseline visits will be summarized.

4.10.7. Safety MRIs

To evaluate white matter changes over time, a shift table will be created from the following categories:

- 0 = No lesions
- 1 = Focal lesions
- 2 = Beginning confluence of lesions
- 3 = Diffuse involvement of entire region

A listing of MRI data will also be presented.

4.10.8. Columbia-Suicide Severity Rating Scale

Suicidal ideation, suicidal behavior, and self-injurious behavior without suicidal intent occurring during treatment, based on the responses to Columbia-Suicide Severity Rating Scale (C-SSRS) consistent with the C-SSRS Scoring and Data Analysis Guide (CUIMC 2016), will be summarized by treatment. In particular, for each of the following events, the frequencies and percentages of participants with the event will be enumerated by treatment: completed suicide, non-fatal suicide attempt, interrupted attempt, aborted attempt, preparatory acts or behavior, active suicidal ideation with specific plan and intent, active suicidal ideation with some intent to act without specific plan, active suicidal ideation with any methods (no plan) without intent to act, non-specific active suicidal thoughts, wish to be dead, and self-injurious behavior without suicidal intent.

Although not suicide-related, the frequencies and percentages of participants with non-suicidal self-injurious behavior occurring during the treatment period will also be summarized by treatment.

In addition, the frequencies and percentages of participants who experienced at least one of various composite measures during treatment will be presented and compared. These include suicidal behavior (completed suicide, non-fatal suicidal attempts, interrupted attempts, aborted attempts, and preparatory acts or behavior), suicidal ideation [active suicidal ideation with specific plan and intent, active suicidal ideation with some intent to act without specific plan, active suicidal ideation with any methods (no plan) without intent to act, non-specific active suicidal thoughts, and wish to be dead], and suicidal ideation or behavior.

The frequencies and percentages of participants who experienced at least one of various comparative measures during treatment will be presented and compared. These include treatment-emergent suicidal ideation compared to recent history, treatment-emergent serious suicidal ideation compared to recent history, emergence of serious suicidal ideation compared to recent history, improvement in suicidal ideation at endpoint compared to baseline, and emergence of suicidal behavior compared to all prior history.

Specifically, the following outcomes are C-SSRS categories and have binary responses (yes/no). The categories have been re-ordered from the actual scale to facilitate the definitions of the composite and comparative endpoints, and to enable clarity in the presentation of the results.

Category 1 – Wish to be Dead
Category 2 – Non-specific Active Suicidal Thoughts
Category 3 – Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act
Category 4 – Active Suicidal Ideation with Some Intent to Act, without Specific Plan
Category 5 – Active Suicidal Ideation with Specific Plan and Intent
Category 6 – Preparatory Acts or Behavior
Category 7 – Aborted Attempt
Category 8 – Interrupted Attempt

Category 9 – Actual Attempt (non-fatal)
Category 10 – Completed Suicide

Self-injurious behavior without suicidal intent is also a C-SSRS outcome (although not suicide-related) and has a binary response (yes/no).

Composite endpoints based on the above categories are defined below.

Composite Measure	Description
Suicidal ideation	A "yes" answer at any time during treatment to any one of the five suicidal ideation questions (Categories 1-5) on the C-SSRS
Suicidal behavior	A "yes" answer at any time during treatment to any one of the five suicidal behavior questions (Categories 6-10) on the C-SSRS
Suicidal ideation or behavior	A "yes" answer at any time during treatment to any one of the ten suicidal ideation and behavior questions (Categories 1-10) on the C-SSRS.

Note: All study participants with a post-baseline C-SSRS assessment are included.

The following outcome is a numerical score derived from the C-SSRS categories. The score is created at each assessment for each participant and is used for determining treatment emergence.

- Suicidal Ideation Score: the maximum suicidal ideation category (1-5 on the C-SSRS) present at the assessment. Assign a score of 0 if no ideation is present.

Comparative endpoints of interest are defined below. "Treatment emergence" is used for outcomes that include events that first emerge or worsen. "Emergence" is used for outcomes that include events that first emerge.

Endpoint	Definition	Notes
Treatment-emergent suicidal ideation compared to recent history	An increase in the maximum suicidal ideation score during treatment (V3-V29) from the maximum suicidal ideation score at baseline (V2).	Recent history excludes "lifetime" scores from the Baseline/Screening C-SSRS scale.
Treatment-emergent serious suicidal ideation compared to recent history	An increase in the maximum suicidal ideation score to 4 or 5 on the C-SSRS during treatment (V3-V29) from not having serious suicidal ideation (scores of 0-3) at baseline (V2).	Recent history excludes "lifetime" scores from the Baseline/Screening C-SSRS scale.
Emergence of serious suicidal ideation compared to recent history	An increase in the maximum suicidal ideation score to 4 or 5 on the C-SSRS during treatment (V3-V29) from no suicidal ideation (scores of 0) at baseline (V2).	Recent history excludes "lifetime" scores from the Baseline/Screening C-SSRS scale.

Improvement in suicidal ideation at endpoint compared to baseline	A decrease in suicidal ideation score at endpoint (the last measurement during treatment; Visits V3-V29) from the baseline measurement (the measurement taken just prior to treatment, at V2).	This analysis should only be performed for a non-lifetime baseline measurement (i.e., having improvement from the worse event over a lifetime is not clinically meaningful). A specific point in time can be used instead of endpoint.
Emergence of suicidal behavior compared to all prior history	An occurrence of suicidal behavior (Categories 6-10) during treatment (V3 – V29) from no suicidal behavior (Categories 6-10) prior to treatment (V2).	Prior to treatment includes "lifetime" and "Since last _ months" scores from the Baseline C-SSRS scale and any "Since Last Visit" from the Since Last Visit C-SSRS scales taken prior to treatment.

Note: All study participants with baseline and post-baseline C-SSRS assessments are evaluated.

Participants who discontinued from the study with no postbaseline C-SSRS value will be considered unevaluable for analyses of suicide-related events. Only evaluable participants will be considered in the analyses. Fisher's Exact test will be used for treatment comparisons.

4.11. Exploratory Endpoint(s) Analysis

4.11.1. Analysis of Blood-based Biomarkers of AD Pathology and Neurodegeneration

To evaluate the change from baseline in blood-based biomarkers of AD pathology and neurodegeneration, the change from baseline up to Week 76 on blood-based biomarkers will be assessed by MMRM in the Efficacy Evaluable analysis set. The blood-based biomarkers includes:

- plasma Neurofilament Light chain (NfL),
- plasma Glial fibrillary acidic protein (GFAP),
- phosphorylated tau (P-tau217)

The values for these biomarkers may be log transformed to fit the normality assumption of the model. For each biomarker, the model will include fixed effects of treatment, visit as a categorical variable, treatment-by-visit interaction, baseline value, baseline value-by-visit interaction, age at baseline. Baseline tau PET category will also be included as a fixed effect to the model applied to full study population. The null hypothesis is that the difference in LS mean change between both active treatment arms and placebo equals zero. The values for the blood-based biomarkers of AD pathology and neurodegeneration may be log transformed to fit the normality assumption of the model. Post-Week 76 time points may be included in the analysis for exploratory purpose.

To assess the relationship of blood-based biomarkers of AD pathology and neurodegeneration versus efficacy endpoints, Spearman's rank correlation coefficient may be obtained on change from baseline up to Week 76 for the blood-based biomarkers versus change from baseline to for CDR-SB, iADRS, ADAS-Cog13, ADCS-iADL, and MMSE. Correlation analyses will be conducted using only participants who have the both the blood-based biomarker and clinical endpoints up to Week 76. This analysis may be performed in moderate tau subpopulation and/or

the full study population. Additional correlation analysis including post-Week 76 time points may be conducted for exploratory purpose.

The same analyses for other blood-based biomarkers may be conducted when the results become available.

4.11.2. Analysis of Digit Symbol Substitution Test and Digital Clock Drawing Test

To assess the effect of LY3372689 vs placebo on additional assessments of cognition as measured by Digit Symbol Substitution Test (DSST) as well as Digital Clock Drawing Test through end time point (76-124 weeks), the DPM analysis as described in Section 4.8.1 will be utilized. Additional analyses using NCS2 and MMRM may be conducted.

4.11.3. Analysis of Time to Clinical Progression

Any increase in CDR-GS from baseline at 2 consecutive visits during the double blinded study period is considered as meeting the criteria of time to substantial decline (MCID, Andrews et al. 2019; Wessels et al. 2022; Lansdall et al. 2023). A Cox proportional hazard model will be utilized to evaluate the hazards of progressing to the clinical worsening events as defined by CDR-GS by treatment arms in the Efficacy Evaluable analysis set in both moderate tau subpopulation and the full study population, respectively.

The analysis will be modeling time to first occurrence of the event as determined above. The ties will be handled using discrete method. The analysis will adjust for age at baseline, concomitant AChEI and/or memantine use at baseline (yes/no). The analysis will be conducted in both moderate tau subpopulation and full study population. The model will be stratified by investigator site (pooled). The model will also be stratified by baseline tau PET category when analyzing full study population. The hazard ratio (HR) for LY3372689-treated group versus placebo group, 95% CI and associated p- value will be provided.

To further evaluate the efficacy of LY3372689 in terms of delaying the substantial clinical worsening The time to substantial clinical progression as defined by iADRS and CDR-SB using the criteria of substantial decline (MCID, Andrews et al. 2019; Wessels et al. 2022; Lansdall et al. 2023) below may be conducted. A clinical worsening event is defined as meeting the criteria at 2 consecutive visits during the double blinded phase:

baseline clinical status / Endpoint	mild cognitive impair (MCI)	mild AD
iADRS	5 points decrease from baseline	9 points decrease from baseline
CDR-SB	1 point or more increase from baseline	2 points increase from baseline

4.11.4. Analyses of Cerebrospinal Fluid Biomarkers

The cerebrospinal (CSF)biomarkers: CSF total tau, CSF P-tau, and CSF NfL collected in the addendum will be summarized by visit. Plasma LY3372689 concentrations may be used to assess the CSF: plasma concentration ratio, or in exposure response relationships, as warranted.

4.12. Other Analyses

4.12.1. Subgroup analyses

To assess the effects of various demographic and baseline characteristics on treatment outcome, subgroup analyses for the primary endpoint, iADRS, will be conducted:

- tau burden as measured by tau PET at baseline– moderate or high; split into two within moderate tau subpopulation; grouped into thirds in full study population.
- ApoE ε4 Carrier Status – Carrier defined as ε2/ε4, ε3/ε4, or ε4/ε4 genotype; Non-Carrier defined as all other genotypes

The primary endpoint will be modeled using NCS2 model, additional analyses may be conducted utilizing DPM and MMRM. For NCS2 and MMRM analyses, data up to WK88 will be included and the 'overall' treatment effect averaging across Week 64, Week 76, and Week 88 will be provided. This general model will include terms for baseline, treatment, treatment-by-visit, subgroup by treatment, subgroup by visit, and treatment-by-visit by subgroup, investigator site (pooled), and tau burden (moderate versus high) as defined by flortaucipir PET when analyzing the full study population for ApoE ε4 Carrier subgroup analyses. Baseline covariates such as age and concomitant AChEI/memantine use (yes/no) will also be included in the model. Redundant terms will be dropped from the model in those cases where the subgroup of interest is overlapping with this general model. In order to run these analyses, at least 20 participants are required in each strata-treatment combination.

4.12.2. Protocol Violations

Listings of participants with significant protocol violations will be provided for the randomized population. The following list of significant protocol violations will be determined from the clinical database and from the clinical/medical group:

- Informed consent violation detected as a missing date of informed consent.
- Did not have an assessment of either the ADAS-Cog or iADL at any of the visits at which the scales were scheduled to be assessed.
- Non-compliance of treatment – define incomplete dose as participant who does not consume 80% to 100% of the prescribed daily dose during this study

The following list of significant protocol violations will be determined by clinical/medical group:

- Protocol violations of inclusion/exclusion criteria.
- Had a study dosing algorithm violation (such as if participants randomized to treatment A were given treatment B or participants randomized to treatment A never received the assigned study drug.)
- Unqualified raters for ADAS-Cog or ADL.

Other protocol violations reported through the monitoring process will be reviewed by the study team and if judged to be significant, will be added to the final reported listing.

4.12.3. Clinical Trial Registry Analyses

Additional analyses will be performed for the purpose of fulfilling the Clinical Trial Registry (CTR) requirements.

Analyses provided for the CTR requirements include the following:

Summary of adverse events, provided as a dataset which will be converted to an XML file. Both Serious Adverse Events and 'Other' Adverse Events are summarized: by treatment group, by MedDRA preferred term.

- An adverse event is considered 'Serious' whether or not it is a TEAE.
- An adverse event is considered in the 'Other' category if it is both a TEAE and is not serious. For each Serious AE and 'Other' AE, for each term and treatment group, the following are provided:
    - the number of participants at risk of an event
    - the number of participants who experienced each event term
    - the number of events experienced.
- Consistent with www.ClinicalTrials.gov requirements, 'Other' AEs that occur in fewer than 5% of participants/subjects in every treatment group may not be included if a 5% threshold is chosen (5% is the minimum threshold).

AE reporting is consistent with other document disclosures for example, the CSR, manuscripts, and so forth.

4.13. Interim Analyses and Data Monitoring Committee

An external DMC is authorized to evaluate results from unblinded interim analyses for the assessment of safety and efficacy/biomarker and to recommend any modifications to the study (including stopping the study). The DMC will have the responsibility to review accumulating unblinded study data and make recommendations to protect the safety of participants. Study sites will receive information about interim results ONLY if they need to know for the safety of trial participants.

Each member of the DMC is a recognized expert in the fields of Alzheimer's disease, neurology, cardiology, or biostatistics. All members will be external to the sponsor. The approved DMC charter enumerates the roles of the DMC members, the frequency with which it meets, the structure of their meetings, operational details.

4.13.1. Safety Interim Analyses

The objective of the first safety interim analysis is to assess the viability of continuing with at least 1 dose of LY3372689 with regard to safety. The initial DMC review of safety data from MTAE is planned to occur when approximately 30 participants have reached 12 weeks of

exposure to study treatment. No statistical adjustments will be made to account for this interim since it is a safety only interim and there is no possibility of stopping for efficacy. The following analyses will be included in safety interim analyses:

- Summary of disposition: Rates of enrollment and participant discontinuations, including reasons for discontinuation
- Summary of demographic characteristics
- Summary of prior and concomitant therapy
- Summary of exposure
- Summary of participant compliance
- Listing and Summary of laboratories (categorical and mean change from baseline)
- Listing and Summary of vital signs and weight (categorical and mean change from baseline)
- Listing and Summary of electrocardiographic/ECGs (categorical and mean change from baseline)
- Listing and Summary of C-SSRS data
- Summary of MMSE data
- Listing and Summary of safety MRI data
    - Listing of all significant treatment-emergent MRI findings.
    - Number of participants with significant treatment-emergent MRI findings.
- Listing and Summary of adverse events, non-serious adverse event, SAEs, TEAEs, discontinuation because of adverse events (DCAEs), adverse events of special interest (AESI, as described in Section 4.9.2.1.2), adverse events necessitating unblinding at the site or by the sponsor.

At least one efficacy and/or biomarker interim analysis may be conducted for Study MTAE; for example, when 50% of randomized subjects have had a chance to complete 52 weeks of treatment (Visit 15), and data will be used to evaluate whether to stop the study for futility. Operational details and a quantitative framework to provide information for these decisions will be documented in a later version of this Clinical Trial Statistical Analysis Plan.

Unblinding details are specified in a separate unblinding plan document.

4.13.2. Efficacy Interim Analysis

An efficacy interim analysis may have been conducted prior to the conclusion of the trial to make an investment recommendation which could potentially accelerate or modify the developmental program for the asset. The details of the interim analysis would have been provided in a separate Interim Efficacy Statistical Analysis Plan and study team members would not have been involved in the interim efficacy analysis planning, execution, or data review. Any Lilly personnel unblinded to efficacy reports are documented in the Blinding or Unblinding Log of Key Study Personnel.

4.14. Changes to Protocol-Planned Analyses

No changes to the planned protocol analyses have been made.

5. Sample Size Determination

Approximately 330 participants will be enrolled and randomized in a 1:1:1 ratio to the 3 treatment arms (placebo and 2 active arms of LY3372689). It is expected that approximately 240 participants (i.e., 75%) will complete the double-blind Period 2a 76-week treatment period of the study (approximately 80 per treatment arm). Using a DPM model over the entire 124-week treatment period, this sample size will provide approximately 96.6% power CCI [redacted] If the active LY3372689 treatment arm progression rate is not different than the placebo treatment arm rate, the probability of passing the efficacy criterion specified above (i.e., false positive) is approximately 2.4% CCI [redacted] The simulation for the power calculation and sample size determination was carried out in the statistical program R, version 3.6.3

6. Appendices

Appendix 1. Potentially Clinically Significant Changes in Vital Signs and Weight

Vital Sign Parameter (Unit)	Postbaseline Low Criteria	Postbaseline High Criteria
Sitting systolic blood pressure (mmHg)	Absolute value ≤90 and ≥20 decrease from baseline	Absolute value ≥160 and ≥20 increase from baseline
Sitting diastolic blood pressure (mmHg)	Absolute value ≤50 and ≥10 decrease from baseline	Absolute value ≥100 and ≥10 increase from baseline
Sitting pulse (bpm)	Absolute value <50 and ≥15 decrease from baseline	Absolute value >100 and ≥15 increase from baseline
Weight	≥7% decrease	≥7% increase
Vital Sign Parameter (Unit)	**Postbaseline Criteria for Abnormality**	
Orthostatic systolic blood pressure (mmHg)	≥20 mmHg decrease in systolic blood pressure (supine to standing) (i.e., supine minus standing ≥20)	
Orthostatic diastolic blood pressure (mmHg)	≥10 mmHg decrease in diastolic blood pressure (supine to standing) (i.e., supine minus standing ≥10 mm Hg)	
Orthostatic pulse (bpm)	≥30 increase in bpm (standing to supine) (i.e., standing minus supine ≥30)	
Temperature	Absolute value ≥38.3°C and ≥1.1°C increase from baseline (Absolute value ≥101°F and ≥2°F increase from baseline)	

Abbreviation: bpm = beats per minute.

Appendix 2. Potentially Clinically Significant Changes in ECGs

ECG Parameter	Low Criteria	High Criteria
Heart Rate	<50 bpm	>100 bpm
PR Interval	<120 msec	≥220 msec
QRS Duration	<60 msec	≥120 msec
QTcF Interval		
Males	<330 msec	≥450 msec
Females	<340 msec	≥470 msec
Males and females		> 500 msec

Abbreviations: bpm = beats per minute; ECG = electrocardiogram; QTcF = Fridericia-corrected QT interval.

Criteria for Prolonged ECG QTcF Interval

QTcF Delta Changes (msec): >30, >60, >75.

Abbreviation: QTcF = Fridericia-corrected QT interval.

7. References

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

Andrews JS, Desai U, Kirson NY, et al. Disease severity and minimal clinically important differences in clinical outcome assessments for Alzheimer's disease clinical trials. *Alzheimers Dement (N Y)*. 20195(1):354-363. https://doi.org/10.1016/j.trci.2019.06.005

Bretz F, Maurer W, Brannath W, Posch M. A graphical approach to sequentially rejective multiple test procedures. *Stat Med*. 2009;28(4):586-604.

Bretz F, Posch M , Glimm E, Klinglmueller F, Maurer W, Rohmeyer K. Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes, or parametric tests. *Biom J*. 2011;53(6):894-913.

Carpenter JR, Kenward MG. *Multiple Imputation and Its Application*. New York, NY: John Wiley & Sons Ltd; 2013.

Chambers JM and Hastie T. editors. *Statistical models in S*. Wadsworth & Brooks/Cole; 1992.Fleisher A, Pontecorvo M, Devous M, et al. Positron emission tomography imaging with [^{18}F]flortaucipir and postmortem assessment of Alzheimer disease neuropathologic changes. *JAMA Neurol.* 2020;77(7):829-839. doi: 10.1001/jamaneurol.2020.0528.

Donohue MC, Langford O, Insel PS, et al. Natural cubic splines for the analysis of Alzheimer's clinical trials [published online January 10, 2023]. *Pharm Stat*. https://doi.org/10.1002/pst.2285

Kotari V, Southekal S, Navitsky M, Kennedy IA, Lu M, Morris A, Zimmer JA, Fleisher AS, Mintun MA, Devous MD Sr, Pontecorvo MJ. Early tau detection in flortaucipir images: validation in autopsy-confirmed data and implications for disease progression. Alzheimers Res Ther. 2023 Feb 28;15(1):41. doi: 10.1186/s13195-023-01160-6. PMID: 36855201; PMCID: PMC9972744.

Lansdall CJ, McDougall F, Butler LM, et al. Establishing clinically meaningful change on outcome assessments frequently used in trials of mild cognitive impairment due to Alzheimer's Disease. *J Prev Alzheimers Dis*. 2023;10(1):9-18. https://doi.org/10.14283/jpad.2022.102

Marcus R, Peritz E, Gabriel KR. On closed testing procedures with special reference to ordered analysis of variance. *Biometrika*. 1976;63(3):655-660.

McLaughlin T, Feldman H, Fillit H, Sano M, Schmitt F, Aisen P, Leibman C, Mucha L, Ryan JM, Sullivan SD, Spackman DE, Neumann PJ, Cohen J, Stern Y. Dependence as a unifying construct in defining Alzheimer's disease severity. *Alzheimers Dement*. 2010;6(6):482-493.

Millen BA, Dmitrienko A. Chain procedures: a class of flexible closed testing procedures with clinical trial applications. *Stat Biopharm Res*. 2011;3(1):14-30.

Pocock SJ, Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics*. 1975 Mar;31(1):103-15. PMID: 1100130.

Pontecorvo M, Devous M, Kennedy I, et al. A multicentre longitudinal study of flortaucipir (^{18}F) in normal ageing, mild cognitive impairment and Alzheimer's disease dementia. *Brain*, 2019;142(6):1723-1735. doi: 10.1093/brain/awz090.

Raket LL. Progression models for repeated measures: estimating novel treatment effects in progressive diseases. *Stat Med*. 2022;41(28):5537-5557. https://doi.org/10.1002/sim.9581

Shcherbinin S, Morris A, Higgins IA, Tunali I, Lu M, Deveau C, Southekal S, Kotari V, Evans CD, Arora AK, Collins EC, Pontecorvo M, Mintun MA, Sims JR. Tau as a diagnostic instrument in clinical trials to predict amyloid in Alzheimer's disease. Alzheimers Dement (N Y). 2023 Aug 16;9(3):e12415. doi: 10.1002/trc2.12415. PMID: 37600216; PMCID: PMC10432878

Wessels AM, Siemers ER, Yu P, Andersen SW, Holdridge KC, Sims JR, Sundell K, Stern Y, Rentz DM, Dubois B, Jones RW, Cummings J, Aisen PS. A Combined Measure of Cognition and Function for Clinical Trials: The Integrated Alzheimer's Disease Rating Scale (iADRS). *J Prev Alzheimers Dis*. 2015;2(4):227-241.

Wessels AM, Rentz DM, Case M, et al. Integrated Alzheimer's Disease Rating Scale: clinically meaningful change estimates. *Alzheimers Dement (N Y)*. 2022;8(1):e12312. https://doi.org/10.1002/trc2.12312