H8A-MC-LZAZ (A4) Statistical Analysis Plan Version 3

Anti-Amyloid Treatment in Asymptomatic Alzheimer's Disease (A4 Study)

NCT02008357

Approval Date: 18-Jan-2023

1. Statistical Analysis Plan: H8A-MC-LZAZ/ADC-040-A4: Anti-Amyloid Treatment in Asymptomatic Alzheimer's Disease (A4 Study)

Solanezumab (LY2062430) Alzheimer's Disease

Multicenter, randomized, double-blind, placebo-controlled, Phase 3 study comparing solanezumab with placebo given as infusions once every 4 weeks over 4.5 years in approximately 1150 participants with preclinical AD with optional open-label extension of up to 4 years.

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Coordinating and Data Management Center
University of Southern California, Alzheimer's Therapeutic Research Institute (ATRI)
San Diego, California USA 92121

Protocol H8A-MC-LZAZ (A4 Study)
Phase 3

Version 3 electronically signed and approved by Lilly and ATRI on date provided below.

Document ID: VV-CLIN-074356

Approved on 18 Jan 2023 GMT

2. Table of Contents

Section **Page**

Table of Contents

Table **Page**

Table of Contents

Figure **Page**

Table of Contents

3. List of Abbreviations

Term	Definition
Aβ	amyloid beta
AD	Alzheimer's disease
ADA	anti-drug antibody
ADCS	AD Cooperative Study
ADCS-PI	AD Cooperative Study Prevention Instrument
ADL	activities of daily living
ADNI	Alzheimer's Disease Neuroimaging Initiative
AE	adverse event
AIBL	Australian Imaging, Biomarkers and Lifestyle Study
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
ANOVA	analysis of variance
APOE4	apolipoprotein subtype E allele 4
APOE ε4	apolipoprotein subtype E allele 4
ARIA-E	amyloid-related imaging abnormality-edema
ARIA-H	amyloid-related imaging abnormality-hemorrhage
AST	aspartate aminotransferase
ATRI	Alzheimer's Therapeutic Research Institute
ATRI-CC	ATRI-Coordinating Center
BMI	body mass index
C3	computerized cognitive composite
C-Path	Critical Path
C-SSRS	Columbia-Suicide Severity Rating Scale
CDR-SB	Clinical Dementia Rating - Sum of Boxes

Term	Definition
CFI	Cognitive Function Index
cLDA	constrained longitudinal data analysis
CN	conventional
COVID	coronavirus disease
COVID-19	coronavirus disease 2019
CRF	case report form
CSF	cerebrospinal fluid
CTR	Clinical Trial Registry
DBL	direct bilirubin levels
DPM	disease progression model
DPR	disease progression ratio
DSMB	Data Safety Monitoring Board
ECG	electrocardiogram
FCSRT	free and cued selective reminding test
GGT	gamma-glutamyl transferase
HABS	Harvard Aging Brain Study
ITT	intent-to-treat
IPTW	inverse probability of treatment weighting
IV	intravenous
Lilly	Eli Lilly and Company
LY	LY2062430
LZAZ	H8A-MC-LZAZ
MAR	missing-at-random
MedDRA	Medical Dictionary for Regulatory Activities
MI	multiple imputation
mITT	modified intent-to-treat

Term	Definition
MNAR	missing-not-at-random
MMRM	mixed-effects model repeated measures
MMSE	Mini-Mental State Examination
MRD	minimal residual disease
MRI	magnetic resonance imaging
MSM	marginal structural models
NAb	neutralizing antibody
NAb+	neutralizing antibody positive
NCS	Natural Cubic Spline
NIA	National Institute on Aging
OLE	Open-label extension
p-tau	phosphorylated tau
PACC	Preclinical Alzheimer Cognitive Composite
PC	placebo-controlled
PET	positron emission tomography
PR	PR interval on the ECG
PRO	participant-reported outcome
PT	preferred term
QRS	the interval that denotes depolarization of the right and left ventricles of the heart
QR	QR interval on the ECG
QT	QT interval on the ECG
QTc	corrected QT interval
QTcF	Fridericia's corrected QT interval
RUI-B	Resource Use Inventory, Brief
SAE	serious adverse event
SAP	statistical analysis plan

Term	Definition
SD	standard deviation
SI	International System of Units
SUVr	standardized uptake value ratio
TBL	total bilirubin levels
TE	treatment-emergent
TE ADA	treatment-emergent ADA
TE ADA+	treatment-emergent ADA positive
TE ADA-	treatment-emergent ADA negative
TEAE	treatment-emergent adverse event
trt	treatment group
ULN	upper limit of normal
usubjid	subject id, unique identifier of each participant
vMRI	volumetric magnetic resonance imaging
WHO	World Health Organization
y	measurement of the PACC

4. Revision History

Statistical Analysis Plan (SAP) Version 1 was approved 27 November 2013 and SAP version 2 was approved 14 January 2019.

SAP Version 3 was approved prior to any unblinded efficacy analyses involving data from Study LZAZ (A4) being conducted and includes the following revisions:

- Added language introducing COVID-19 pandemic interruption and hiatus
- Primary outcome model was changed from MMRM to cLDA-NCS model
- APOE4 carrier status, a stratification variable for randomization in the study, was added to the primary analysis model and secondary/sensitivity analyses.
- Added sensitivity analyses of PACC with MMRM and disease progression model (DPM)
- Analyses for CDR Global Score added to analysis plan.
- Addition of analyses to look at the effect of increasing dose.
- Further description of the Cogstate Composite Scores
- Removed Section 7.9.2.2. Constrained Longitudinal Data Analysis of Slopes as the NCS model will address this question.

SAP Version 2 was approved prior to any unblinded efficacy analyses involving data from Study LZAZ (A4) being conducted and included the following revisions:

- Added delayed-start efficacy analyses for the open-label extension period that was added as part of the Study LZAZ (A4) protocol amendment (c).
- Added safety analyses for the open-label extension
- Modified the description of the analyses to account for the option of increasing the solanezumab dose to 1600 mg
- Changed the study duration from 3.25 years to 4.5 years
- Changed the Gatekeeping Strategy for multiple comparisons based on the FDA feedback to assess the study's weight of evidence
- Changed the sensitivity analyses
- Added subgroup analyses for dose

5. A constrained Longitudinal Data Analysis (cLDA) Natural Cubic Spline (NCS) analysis of the primary outcome measure, the Preclinical Alzheimer Study Objectives

5.1. Primary Objective

The primary objective of this study is to test the hypothesis that solanezumab, administered as an IV infusion every 4 weeks for 4.5 years, will slow cognitive decline as compared with placebo in participants with preclinical AD. The primary objective will be assessed using a constrained Longitudinal Data Analysis (cLDA) Natural Cubic Spline (NCS) analysis (referred to as NCS henceforth) of the primary outcome measure, the Preclinical Alzheimer Cognitive Composite (PACC), a composite that is primarily weighted on episodic memory and executive function tests, in which the specific hypothesis is that the cognitive decline at 240 weeks will be significantly less for solanezumab than for placebo.

The PACC will be composed of 4 components: the total score from the free and cued selective reminding test (FCSRT), the delayed paragraph recall on the logical memory IIa test from the Wechsler Memory Scale, the digit-symbol substitution test from the Wechsler adult intelligence scale-revised, and the MMSE total score. Specifically, for the FCSRT, the total will be the sum of the free and cue score plus the free recall score, resulting in a total from 0 to 96. Each component score is converted to a z-score by subtracting the baseline mean for that component and dividing by the baseline standard deviation for that component. The PACC is the sum of the four resulting z-scores.

5.2. Secondary Objectives

The secondary objectives of the study are as follows:

- To test the hypothesis that solanezumab will slow the decline of perceived cognitive function and performance of everyday activities, as compared with placebo, using NCS analysis of the Cognitive Function Index (CFI).
- To assess whether decline in activities of daily living begins by the end of the treatment period, and if so, whether an effect of solanezumab, compared with placebo, can be detected using NCS analysis of the ADCS-Activities of Daily Living (ADL) Prevention Questionnaire.
- To assess the relationship between treatment effect and time using the PACC.
- To test the hypothesis that solanezumab reduces brain amyloid burden, as compared with placebo, as assessed using florbetapir positron emission tomography (PET) imaging.
- To assess effects of solanezumab on CSF and plasma concentrations of total tau peptides and phosphorylated tau peptides (p-tau).
- To assess effects of solanezumab on CSF concentrations of Aβ.
- To investigate the effect of treatment with solanezumab on volumetric MRI.

- To assess the safety of solanezumab versus placebo treatment, including AEs and immunogenicity.

The secondary objectives of the open-label period of the study are as follows:

- To assess the persistence of effect of solanezumab treatment in participants with preclinical AD. That is, to test the hypothesis that participants originally randomized to receive placebo and later switched to solanezumab at the start of the open-label period do not "catch up" to participants originally randomized to receive solanezumab in the placebo-controlled period, using a randomized/delayed-start analysis. The main efficacy objective of the open -label period will be an interim randomized/delayed-start analysis of the PACC at the time of database lock for the placebo-controlled period using data available up to 2 years in the open-label period.
- To test the hypothesis that solanezumab will continue to slow the decline associated with preclinical AD during open-label treatment, comparing participants initially randomized to solanezumab with participants initially randomized to placebo in the placebo-controlled treatment period, using randomized/delayed-start analysis of the MMSE, CDR-SB, C3, CFI, and the ADCS-ADL Prevention Questionnaire.

5.3. Exploratory Objectives

The exploratory objectives of the study are as follows:

- To assess the effect of treatment with solanezumab as demonstrated using the MMSE.
- To assess the utility of a novel computerized battery, the Computerized Cognitive Composite (C3), in predicting and tracking clinical decline and response to solanezumab.
- To assess the effects of treatment on healthcare resource utilization.
- To determine the best predictors of clinical decline based on the PACC.
- To develop novel sensitive outcome measures to improve the efficiency of future secondary prevention trials, including exploratory measures of self--reported assessment of cognitive and interpersonal functioning.
- To assess the effect of treatment with solanezumab as demonstrated using the Clinical Dementia Rating (CDR) scale.
- To assess the effect of treatment with solanezumab as demonstrated using the C-Path Consortium's Participant-Reported Outcome Questionnaire (C-Path PRO).
- To investigate the impact of solanezumab on markers of synaptic dysfunction on functional connectivity MRI.
- To test the hypothesis that solanezumab reduces brain tau burden, as compared with placebo, as assessed using flortaucipir positron emission tomography (PET) imaging.

- To assess solanezumab-associated changes in levels of plasma Aβ species and CSF Aβ species. The hypothesis that solanezumab, unlike placebo, alters amyloid-plaque associated forms of Aβ will be assessed by 1) demonstrating presence of plaque-associated Aβ species in the plasma, and 2) confirming that concentrations of CSF free (unbound to antibody) $A\beta_{1-42}$ are increased or unchanged and CSF free $A\beta_{1-40}$ levels are decreased with solanezumab treatment.
- To explore whether baseline markers of neurodegeneration (volumetric MRI or CSF tau or p-tau) are related to cognitive decline and response to treatment with solanezumab.
- To explore the impact of disclosure of amyloid status on questionnaires probing perception of amyloid imaging and concern about developing AD.
- To explore biomarker assessments collected at baseline and at the end of the placebo-controlled period as potential predictors of treatment effect during the open-label period.
- To assess preclinical AD during open-label treatment by comparing participants initially randomized to solanezumab in the placebo-controlled period with participants initially randomized to placebo in the placebo-controlled treatment period, using randomized/delayed-start analysis of the RUI-B.

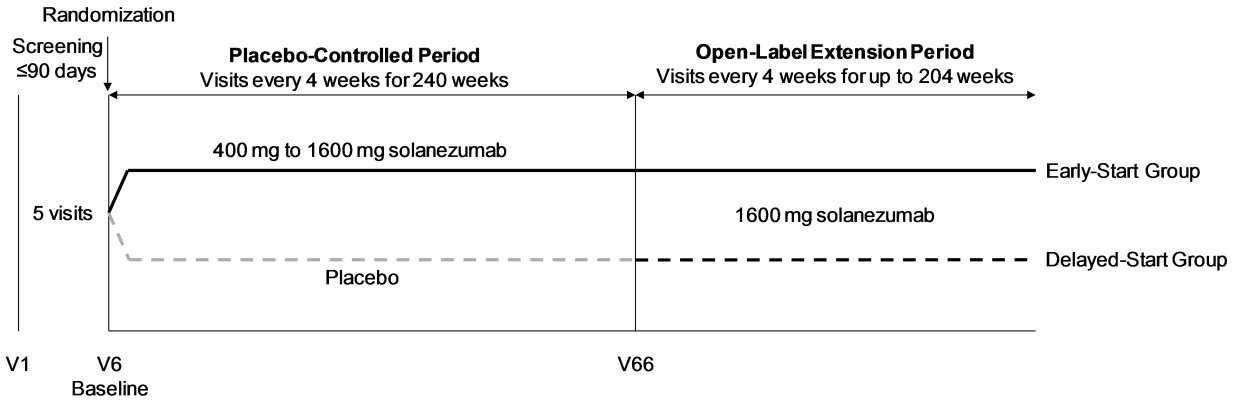
6. Study Design

6.1. Summary of Study Design

Study A4 is a multicenter, randomized, double-blind, placebo-controlled, Phase 3 study comparing solanezumab with placebo for 240 weeks in approximately 1150 participants ages 65 to 85 years old with preclinical AD, defined as having evidence of brain amyloid pathology without clinically evident cognitive impairment at screening. Participants who met entry criteria were randomized in a 1:1 ratio (approximately 575 per treatment arm) to solanezumab or placebo once every 4 weeks. Participants were randomized within the site by education – high (13 or more years) or low (12 or less years) and the presence of one or more apolipoprotein E gene (APOE) ε4 alleles (yes, no). The primary hypothesis being tested is that the cognitive decline at the end of the placebo-controlled period will be significantly less for solanezumab than for placebo. Participants completing the placebo-controlled period of the study can opt to participate in the open-label period of the study that will last for up to 204 weeks or until the primary analyses of the placebo-controlled period of the study are completed and reviewed.

At the beginning of the COVID-19 pandemic, a hiatus was implemented to allow participants to pause study activities without missing visits and resume when COVID safety mitigations were in place (for example, vaccines, home infusions, etc.). The hiatus extended the end of the placebo-controlled period for up to 9 months and the placebo-controlled period will be administratively closed in December 2022.

Figure A4.SAP.1 illustrates the study design.



Abbreviation: V = visit.

Figure A4.SAP.1. Illustration of study design for Clinical Protocol A4.

6.2. Determination of Sample Size

In order to develop an appropriate measure of cognitive decline in a preclinical population, we tested several combinations of measures using data from longitudinal studies in clinically normal populations, including the ADNI; the AIBL; the ADCS-PI trial; and the HABS.

Sample size and power at 4.5 years (240 weeks) for the PACC were estimated using data from ADNI and HABS. Estimates were obtained by applying appropriate assumptions regarding treatment difference, SD, and attrition for a 4.5-year study in a preclinical AD population. In ADNI, the difference in the composite change between participants with and without elevated brain amyloid at 240 weeks was 2.13 (SD=2.85). Similarly, in HABS, the difference in the composite change between participants with and without elevated brain amyloid at 240 weeks was 2.66 (SD=3.08). Given the ADNI-derived estimate of SD at 240 weeks of SD=2.85 and 30% attrition, the total N=1150 provides 80% power (5% 2-sided alpha) to detect a treatment difference of 0.532 points or 0.532/2.13 = 24.9% of the amyloid group difference. Similarly, based on the HABS estimate of SD at 240 weeks of SD=3.08, the study has 80% power to detect a treatment difference of 0.570 points or 0.570/2.66 = 21.4% of the amyloid group difference.

For the open-label period, the estimated mean PACC difference between cognitively normal participants with and without elevated brain amyloid at 336 weeks is 2.95 (SD=3.77) based on ADNI and 4.74 (SD=3.99) based on HABS. Accounting for the administrative censoring that will be induced by the common close design of the open-label period, and assuming an overall attrition rate of 30%, about 266 PACCs are expected to be available at 336 weeks. The visit-to-visit correlation is estimated to be about 0.5. Under these pilot estimates, 266 PACCs at 336 weeks would provide 80% power to detect a randomized group (early-start vs delayed-start treatment with solanezumab) difference of about 0.9 PACC points. This change of 0.9 PACC points reflects 30% of the difference between amyloid-positive and -negative participants based on ADNI, and 20% of that difference based on HABS.

Sample size for the study was determined assuming an MMRM analysis prior to the interruption caused by the COVID-19 pandemic. Applying the NCS approach in simulated trials with a COVID-19 pandemic interruption suggests the study has 94% power to detect an effect size of 0.75 PACC points at 240 weeks.

6.3. Method of Assignment to Treatment

Participants who met all criteria for enrollment were randomized in a 1:1 ratio to double-blind treatment at baseline (Visit 6). For between-group comparability, participants were randomized within site by education – high (13 or more years) or low (12 or less years), and by the presence of one or more APOE ε4 alleles (yes, no). Assignment to treatment groups was determined by a computer-generated random sequence.

7. A Priori Statistical Methods

7.1. General Considerations

Unless otherwise noted, all tests of treatment effects will be conducted at a 2-sided alpha level of 0.05; 2-sided confidence intervals 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.

All analyses will follow the modified intent-to-treat (mITT) principle unless otherwise specified. An ITT analysis is an analysis of data by groups to which the participants are assigned by random allocation, even if the participant does not take the assigned treatment, does not receive the correct treatment, or otherwise does not follow the protocol. An mITT analysis is an ITT analysis for all participants who have a baseline and at least 1 postbaseline measure. All safety analyses will be done using the safety population. Analysis populations are described below (Section 7.5).

When change from baseline is assessed, participants will be included in the analysis only if both a baseline and a post-baseline measure are available. Unless otherwise defined, a baseline measure is the last non-missing observation collected prior to the first infusion of study medication. End point is the last non-missing post-baseline measurement.

For MMRM analyses, observations collected at non-scheduled visits (see Section 7.2) will not be included in the analyses. For NCS and MMRM analyses, unless otherwise specified, an unstructured covariance structure will be used. If the unstructured covariance structure matrix results in a lack of convergence, the following covariance structures will be used in sequence: heterogeneous Toeplitz covariance structure, heterogeneous autoregressive covariance structure, heterogeneous compound symmetry covariance structure and homogeneous compound symmetry covariance structure.

7.2. Handling of Dropouts or Missing Data

If any of the individual items for any of the clinical scales are missing or unknown, every effort will be made to obtain the score for the missing item or items. For the primary outcome (PACC), if one of the four components is missing, the total score will be imputed. The sum of the other three components will be prorated to the sum of total components. 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 more than one of the components is missing, the PACC at that visit will be considered missing.

For the ADL Prevention Questionnaire, 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 at that visit will be considered missing.

The same imputation technique will be applied to the CFI, CDR-SB and MMSE.

For the FCSRT, the total score will be imputed if only 1 of the three trials of cued scores is missing – if any of the free recall scores are missing or if more than 1 cued recall scores are missing, the total will not be imputed and will be considered missing. The imputation will rely on the sensitivity to cueing measure, which is defined as the number of correct cued responses divided by the total number of cued responses across the two completed trials. For instance, if a participant was cued on 10 items and correctly answered on 8 of them, the sensitivity to cueing measure would be 0.8. This value will then be multiplied by the number of items the participant should have been cued on for the missing trial and the resulting value will be rounded to the nearest integer.

For all other scales, if any item is missing, any total or sum involving that item will be considered missing.

Analyses that treat time as a continuous variable (e.g. NCS and slope) will utilize every post-baseline observation in the placebo-controlled phase. Repeated measures analyses that treat time as a categorical variable (MMRM) will only use data from visits at which the data were scheduled to be collected (Andersen and Millen 2013). Amendment (e) of the protocol defined the acceptable window for efficacy scales to be 8 weeks. Therefore, if a scale is missing from a protocol-defined visit where the scale was to be assessed but present at a subsequent visit that is not more than 8 weeks after the protocol-defined visit and prior to the next protocol-defined visit at which the scale is to be assessed, then the late-assessed scale result will be "carried back" and used at the protocol-defined visit.

When participants discontinue from the study early, there may be efficacy or safety data measurements at visits at which the variables were not scheduled to be collected. These data will be used in all other analyses.

7.3. Multicenter Studies

This study will be conducted at multiple centers (investigative sites). No adjustments will be made for multiple sites in the analyses.

7.4. Multiple Comparisons/Multiplicity

The primary analysis is the NCS analysis of the PACC (Section 7.9.1). All other analyses are secondary to this analysis.

7.5. Analysis Populations

The primary and secondary efficacy measures will be analyzed using the mITT population unless otherwise specified. In addition, the PACC will be analyzed using the per-protocol population to verify the robustness of the results. Summaries and analyses for safety measures will be based on the safety population.

Table A4.SAP.1 defines each of the analysis populations used in this study. Table A4.SAP.2 lists the study measures that will be summarized and/or analyzed for each population.

Tabulations of the number and percentage of participants included in each analysis set, by treatment group and overall, will be provided.

Table A4.SAP.1. Analysis Populations for Study A4

Population Name	Description of Population
All Participants Entered	All participants who signed informed consent.
Intent-to-Treat Population (ITT)	All randomized participants.
Modified Intent-to-Treat Population (mITT)	All randomized participants who have a baseline observation and at least one post-baseline observation
Placebo-Controlled Completers Population	All ITT participants who have completed the placebo-controlled study period based on disposition data.
Safety Population	All ITT participants with at least 1 complete or partial infusion of study medication
Per-Protocol Population	All participants in the ITT population who also: • signed the informed consent form • had an assessment of the PACC at each scheduled visit they completed • had no violations of inclusion/exclusion criteria • had no 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) • had less than 20% of infusions incomplete (incomplete infusion = less than 75% of volume infused) • did not miss any more than 10 infusions during the placebo-controlled period • did not miss 3 or more consecutive infusions for reasons other than medical • for placebo-controlled period completers, the treatment duration for the study (that is, Visit 6 through Visit 66) was not more than 240 weeks + 10 days

Abbreviation: PACC = Preclinical Alzheimer Cognitive Composite.

Table A4.SAP.2. Efficacy and Safety Measures Summarized and/or Analyzed for Each Analysis Population

Population Name	Type(s) of Reports
All Participants Entered	Listings of Participant ID, Site/Investigator ID, Informed Consent date
Intent-to-Treat Population (ITT)	Tables, Listings, Figures of the following: participant disposition, participant characteristics at baseline, pre-existing conditions, significant historical diagnoses, concomitant medications.
Modified Intent-to-Treat Population (mITT)	Tables, Listings, Figures of the following: PACC, ADL-Prevention Questionnaire, CFI, CDR-SB, CDR Global Score, C3, MMSE, RUI-B, florbetapir parameters, plasma parameters, vMRI parameters, CSF parameters.
Completers Population	Tables, Listings, Figures of the following: PACC, ADL-Prevention Questionnaire, CFI, CDR-SB, C3, MMSE, RUI-B, florbetapir Parameters, plasma parameters, vMRI parameters, CSF parameters.
Safety Population	Tables, Listings, Figures of the following: exposure and infusion, adverse events, laboratory results, vital signs, weight, ECG, MRI, Immunogenicity, Psychological Well-Being, C-SSRS
Per-Protocol Population	Tables, Listings, Figures of the following: participant characteristics at baseline, PACC.

Abbreviations: ADL = activities of daily living; C3 = computerized cognitive composite; C-SSRS = Columbia-Suicide Severity Rating Scale; CDR-SB = Clinical Dementia Rating - Sum of Boxes; CFI = Cognitive Function Index; CSF = cerebrospinal fluid; ECG = electrocardiogram; ID = identification; MMSE = Mini-Mental State Examination; MRI = magnetic resonance imaging; PACC = Preclinical Alzheimer Cognitive Composite ;RUI-B = Resource Use Inventory, Brief; vMRI = volumetric magnetic resonance imaging.

7.6. Participant Characteristics

Baseline characteristics will be summarized for the ITT and per-protocol populations by treatment group and overall. Summaries will include descriptive statistics for continuous and categorical measures. Fisher's exact test or Pearson's chi-square test will be used for treatment-group comparisons of categorical data. For continuous data, analysis of variance (ANOVA), with independent factor for treatment will be used. Participant characteristics to be presented include:

- age
- gender
- race
- height
- body weight
- body mass index (weight (kg) / [height (m)]^2)
- tobacco use
- alcohol use
- caffeine use

- years of education
- work status
- Baseline score as measured by PACC, ADL Prevention Questionnaire, CFI, CDR-SB, C3, MMSE, RUI-B.
- APOE4 carrier status (carrier [ε2/ε4, ε3/ε4, ε4/ε4], noncarrier [ε3/ε3, ε2/ε2, ε3/ε2]), and
- APOE4 genotype (ε2/ε4, ε3/ε4, ε4/ε4, no ε4), and
- Baseline Amyloid Centiloids

7.7. Participant Disposition

The following analyses of participant disposition will be conducted for the ITT population

- The percentage of participants discontinuing from each treatment group will be compared between groups using Fisher's exact test.
- The median time to discontinuation will also be compared between treatment groups using the Kaplan-Meier product limit estimator.

Comparisons using Fisher's exact test will be done for the overall percentage of participants who discontinue and also for each specific reason for discontinuation collected in the CRF. For any-cause discontinuation as well as discontinuation due to adverse event (AE) or death, comparisons of time-to-discontinuation will be conducted using the Kaplan-Meier product limit estimator and the associated log-rank test. For any-cause discontinuation, for participants who discontinue, time to discontinuation will be defined as the date of study disposition minus randomization date. For participants who did not discontinue, they will be considered censored and time will be defined as completion date of the double-blind period minus randomization date. For discontinuation due to adverse event or death, for participants who discontinue due to one of those reasons, time to discontinuation will be defined as the date of study disposition minus randomization date. All other participants will be considered censored and time will be defined as either completion date of the double-blind period minus randomization date or study disposition date (but not for AE or death) minus randomization date.

7.8. Concomitant Medications

Prior medications are defined as those that stop before randomization (Visit 6). Concomitant medications are defined as those being taken on or after randomization (Visit 6). 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.

If the start or stop dates of medications are missing or partial to the degree that determination cannot be made of whether the medication is prior or concomitant, the medication will be deemed concomitant. Medications will be coded using the World Health Organization (WHO) drug dictionary.

7.9. Efficacy Analyses

7.9.1. Analysis of Primary Outcome (Placebo-Controlled Period)

7.9.1.1. NCS Analysis of Primary Outcome

The primary objective of this study is to test the hypothesis that intravenous infusion of solanezumab will slow cognitive decline in preclinical AD as compared with placebo.

The original mITT analysis (that is, all participants as randomized to treatment who have a baseline and at least 1 postbaseline measure) for the primary outcome will be retained and conducted at the end of the treatment period, although the dose was increased during the study and dosing and assessments were interrupted due to COVID-19; in other words, the primary analysis will include all data for the solanezumab group as the treatment group versus placebo. As with the original analysis plan, the hypothesis for this analysis will be tested against a two-sided alpha level of .05.

The details of this analysis including the analytic method and model are presented below.

An NCS analysis with 2 degrees of freedom will be used to assess the difference between treatment groups in change from baseline for the PACC score at 240 weeks. The composite score at baseline and each post-baseline visit will be the dependent variable. Time will be treated as a continuous variable with values equal to the years between baseline and follow-up exam dates. We will assume an unstructured covariance structure, where the time values are mapped back to the corresponding visits. The NCS basis expansion will assume an interior knot at the median of observation times, and boundary knots at zero years and the maximum follow-up (Hastie 1992). The fixed effects will include the following terms: (i) NCS basis expansion terms (two terms), (ii) NCS basis expansion terms-by-treatment interaction (two terms), (iii) PACC test version administered, (iv) age, (v) education, (vi) APOE4 Carrier Status (yes/no), and (vii) baseline florbetapir cortical SUVr. The model is constrained to not allow a difference between treatment group means at baseline.

The null hypothesis is that the treatment difference between solanezumab and placebo for the PACC score at 240 weeks is equal to zero. The primary analysis will be carried out in the mITT population that is randomized and has a baseline score and at least 1 post-baseline assessment. If the unstructured 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.

The Satterthwaite approximation will be used to estimate the denominator degrees of freedom. The model with unstructured covariance structure will be fit in R using the lme4::lmer function.

Other correlation structures will be fit with the nlme::gls function. The emmeans R package will be used to calculate the Satterthwaite approximation for denominator degrees of freedom.

7.9.1.2. Sensitivity Analyses

7.9.1.2.1. MMRM Analysis of Primary Outcome

An MMRM analysis will be used to assess the difference between treatment groups in change from baseline for the PACC score at 240 weeks (Visit 66). Change from baseline at each visit when the composite score is scheduled will be the dependent variable. Visit will be treated as a categorical variable with values equal to the visit numbers at which the composite measure was scheduled. We will assume an unstructured covariance structure. The model for the fixed effects will include the following terms: (i) baseline PACC score, (ii) treatment, (iii) visit, (iv) treatment-by-visit interaction, (v) age, (vi) education, (vii) APOE4 Carrier Status (yes/no), and (viii) baseline florbetapir cortical SUVr.

The null hypothesis is that the treatment difference between solanezumab and placebo for the PACC score at 240 weeks is equal to zero. The primary analysis will be carried out in the mITT population that is randomized and has a baseline score and at least 1 post-baseline assessment. If the unstructured covariance structure matrix results in a lack of convergence, the same sequence of covariance structures as above will be attempted. Additionally, a Bayesian disease progression model (DPM) (Sparks 2021) with a proportional treatment effect will be conducted to assess statistical differences in the rate of decline of the PACC score between the solanezumab group and the placebo group. The analysis is testing the hypothesis that the disease progression ratio (DPR), defined as the ratio of the rate of decline of the solanezumab arm divided by the rate of decline of the placebo arm, is less than 1.

$$H_0: DPR=1$$

$$H_1: DPR<1$$

The key assumption of the DPM is that it assumes that the treatment effect of solanezumab is proportional to placebo over the course of the study. The proportionality assumption is similar to the assumption made in proportional hazards modeling of time to event data. The model includes diffuse priors on all parameters; therefore, the prior distributions have very little impact on the posterior distributions. No information or knowledge of the effect of solanezumab from previous studies will be incorporated into the prior distributions.

The DPM is as follows:

$$Y_{ij} = \gamma_i + e^{\theta_{T_i}} \sum_{v=0}^{j} \alpha_v + \varepsilon_{ij}$$

where Y_{ij} denotes the PACC score at visit j for participant i; the PACC score for a participant at baseline (prior to treatment) is Y_{i0}. The value γ_i (i=1, 2, …, k) represents a participant specific random effect. The parameter T_i denotes the treatment arm for participant i, where T_i has a value

of 1 if a participant is randomized to solanezumab, and a value of 0 if the participant is randomized to placebo. The parameter α_v is the change in mean PACC score for placebo from visit *v*-1 to *v*, and ε_{ij} is the error term. The DPR for solanezumab relative to placebo is provided by the parameter e^{θ}. Covariates of the model include PACC test version, age at baseline, education, APOE4 Carrier Status (yes/no), and baseline florbetapir cortical SUVr.

To examine the hypothesis of a disease progression benefit, we calculate the posterior probability of the alternative hypothesis for various thresholds (for example, 0% slowing, 25% slowing relative to placebo, etc.). A 95% credible interval (from the 2.5th to the 97.5th percentiles) and posterior mean and median disease progression ratio will be presented.

7.9.1.2.2. Assessing Missing at Random (MAR) in MMRM

Sensitivity to departures from the missing-at-random (MAR) assumption will be investigated using a tipping point analysis (Carpenter and Kenward 2013). This method is a sensitivity analysis in multiple imputation under the missing-not-at-random (MNAR) assumption that searches for a tipping point that reverses the study conclusion. Departures from MAR in the solanezumab treatment group 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 (ie, 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 PACC scores will be analyzed based on data observed while the participant remains on study as well as data imputed using multiple imputation (MI) methodology for time points at which no value is observed. Imputed values in the solanezumab treatment group will first be sampled from an MAR-based multiple imputation model and then δ-adjusted as described below.

Missing-at-random-based imputations will be generated for PACC scores at each time point, and then a value of δ will be added to all imputed values in the solanezumab treatment group 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. The multiple datasets will be analyzed with the NCS approach described for the primary analysis, and estimated treatment effects will be appropriately pooled.

7.9.1.2.3. Assessing Effect of Dose Change

Sensitivity analyses of the effect of dose change on clinical outcomes will be conducted. An NCS analysis will be performed separately on the PACC and other clinical measures after censoring all observations after the dose increase (that is, after the first dose of 800 mg). In other words, clinical measures will be included only through the last available value prior to first dose

escalation for each participant. The model and variables included in the NCS analysis will be the same as specified in Section 7.9.1.1. The null hypothesis is that the treatment difference between solanezumab and placebo for the PACC score is equal to zero. Since the timing (visit) of dose escalation will vary for each participant, a specific endpoint for comparison between treatment groups will not be available. Therefore, the hypothesis will be tested using the p-value obtained for the overall treatment effect.

7.9.1.2.4. Assessing Effect of Exposure

Additionally, we will implement a sensitivity analysis to estimate the potential differences in the cumulative treatment exposure over time due to both the dose increase and the COVID-19 hiatus. In addition to the two spline basis expansion terms for time from baseline in the primary analysis NCS model, this model will also include spline basis expansion terms for the cumulative exposure to solanezumab over time. The fixed effects will include the following terms: (i) NCS basis expansion terms for time (two terms), (ii) NCS basis expansion terms for cumulative exposure (two terms), (iii) PACC test version administered, (iv) age, (v) education, (vi) APOE4 Carrier Status (yes/no), and (vii) baseline florbetapir cortical SUVr. Treatment group is not explicitly included in the model, but the cumulative exposure for the placebo group is set to zero. Based on this model, we will be able to visualize the modeled mean for the PACC over time with 95% confidence intervals for (i) placebo group, (ii) the low-dose per-protocol exposure, and (iii) the high-dose per-protocol exposure. We will also be able to visualize the modelled mean PACC at 240 weeks for the observed range of cumulative dose exposure and identify the minimum exposure at which the 95% confidence interval excludes the mean for the placebo group.

To further investigate exposure, participants will be grouped into one of three exposure groups based on the tertiles of the distribution of cumulative exposure to either solanezumab or placebo: low exposure, medium exposure, and high exposure. Within these exposure groups, solanezumab-treated participants will be compared to placebo-treated participants using the NCS model including effects for (i) NCS basis expansion terms, (ii) exposure group (iii) NCS basis expansion terms-by-treatment-by-exposure group interaction, (iv) PACC test version administered, (v) age, (vi) education, (vii) APOE4 Carrier Status (yes/no), and (viii) baseline florbetapir cortical SUVr. This model will constrain the baseline treatment group means to be the same within exposure group, but allow different baseline means between the exposure groups. The interaction will be tested by likelihood ratio test comparing the full model to the reduced model without any exposure group interaction effects.

7.9.2. Analysis of Additional Efficacy Outcomes (Placebo-Controlled Period)

7.9.2.1. MMRM Analysis of Additional Clinical Outcomes

Similar to the primary analysis, the following efficacy outcomes will be assessed using an MMRM analysis: MMSE, ADCS-ADL Prevention Questionnaire, CDR-SB, CFI, and C3. For each efficacy measure, the change from baseline score at each scheduled postbaseline visit (according to the Study Schedule) during the treatment period will be analyzed using MMRM. Also similar to the primary analysis, all data from all participants regardless of dose increase will

be included for the solanezumab and placebo groups. Change from baseline at each visit when the scale/measure is scheduled will be the dependent variable. Visit will be treated as a categorical variable with values equal to the visit numbers at which the composite measure was scheduled. We will assume an unstructured covariance structure. The model for the fixed effects will include the following terms: (i) baseline score for scale/measure, (ii) treatment, (iii) visit, (iv) treatment-by-visit interaction, (v) age, (vi) education, (vii) APOE4 Carrier Status (yes/no), and (viii) baseline florbetapir cortical SUVr.

7.9.2.2. CDR Global Score

Participants will be classified for clinical worsening based on the CDR Global Score. As all participants begin the trial with a CDR Global Score of 0, clinical worsening will be defined as two consecutive CDR Global Scores above 0, or an endpoint CDR Global Score above 0.

Association between treatment and clinical worsening at any time during the double-blind treatment phase will be tested using Fisher's Exact test.

Time to clinical worsening will be assessed through a Kaplan-Meier analysis. For participants who meet the definition of clinical worsening, time to clinical worsening will be defined as the time from randomization to the first non-zero assessment of the CDR Global Score that contributes to the clinical worsening. Participants who do not meet the definition of clinical worsening will be censored at the time of their last CDR Global Score assessment. Treatments will be compared using the log-rank test.

Additionally, time to clinical worsening will also be measured with a Cox proportional hazards model. Definitions of time to event will be identical to those used for the Kaplan-Meier analyses. Time to clinical worsening will be modelled with fixed effects for baseline PACC score, treatment, age, education, APOE4 Carrier Status (yes/no), and baseline florbetapir cortical SUVr. The coefficient associated with treatment will be tested using the Wald Chi-squared statistic.

For both the Kaplan-Meier and the Cox proportional hazard models, APOE4 Carrier Status (yes/no) will also be investigated for differences in time to clinical worsening.

7.9.2.3. Completer Analysis

Separate NCS analysis of the subset of participants who complete the placebo-controlled period will be done for each of the following:

- PACC
- CFI
- ADL-Prevention Questionnaire
- MMSE
- CDR-SB, and
- C3.

The model and approach will be the same as described above for the primary analysis.

7.9.3. Open-Label Period

7.9.3.1. General Considerations

When change from baseline is assessed, participants will be included in the summary only if both a baseline and a postbaseline measure are available. For efficacy analyses (noninferiority and superiority analyses), the baseline will be the last non-missing observation collected prior to the initiation of treatment in the placebo-controlled period, generally, Visit 6.

7.9.3.2. Delayed-Start Analysis of PACC

The main efficacy analysis of the open-label period is to test for disease progression between treatment groups from the end of the placebo-controlled period at Week 240 to the time point in the open-label study period at Week 336 as assessed by the PACC. Comparisons will be made between participants who started solanezumab at the beginning of the placebo-controlled period (early start) and those who started solanezumab at the beginning of the open-label period (delayed start).

In the following description, the difference between solanezumab and placebo at the end of the placebo-controlled period will be denoted Δ_1; the difference between early-start and delayed-start solanezumab at the time point of interest in the open-label period will be denoted Δ_2. Note that Δ_1 is the effect obtained from the MMRM analysis of the placebo-controlled period. The primary delayed-start efficacy measure will be analyzed using estimates from MMRM models, in which three specific hypotheses are tested:

1. Δ_1 is statistically significantly greater than 0.
2. Δ_2 is statistically significantly greater than 0.
3. 90% one-sided confidence limit of $\Delta_2 - 50\% \Delta_1$ is greater than 0.

The change from baseline (prior to the initiation of treatment in the placebo-controlled period) at each visit during both the placebo-controlled period and the open-label period (at the visits when the PACC is assessed) will be the dependent variable. The model for the fixed effects will include terms for: baseline PACC, treatment, visit, treatment-by-visit interaction, education, APOE4 Carrier Status (yes/no), baseline florbetapir cortical SUVr, and age at baseline. Visit will be considered a categorical variable with values equal to the visit numbers at which the PACC was assessed. An unstructured covariance matrix will be used to model the within-participant variance–covariance errors. 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
- compound symmetry covariance structure

The first hypothesis will be tested using an estimate of Δ_1 obtained from an MMRM model fit using data from all randomized participants through week 240 (the end of the placebo-controlled period). The second hypothesis will be tested using an estimate of Δ_2 obtained from an MMRM including data from all randomized participants through week 336 (96 weeks of open-label period).

The noninferiority margin for this hypothesis is specified as 50% of the treatment difference observed at the end of the placebo-controlled period. The hypothesis will be tested by constructing a 90% one-sided confidence interval of the difference in least-squares means at the last visit in the delayed-start period. If the lower limit of the confidence interval rules out the difference which would have been obtained if 50% of the observed difference had been lost, the disease progression for the treatment groups will be considered parallel in the open-label period, suggesting the treatment effect is independent of symptomatic effects and that solanezumab has a persistent effect on disease course.

7.9.3.3. Delayed-Start Analyses of Other Efficacy Scales

Several other efficacy measures (ADCS-ADL Prevention Questionnaire, CDR-SB, C3, and CFI) will be analyzed separately using the delayed-start analysis described in Section 7.9.3.2. Additionally, the Clinician Diagnostic Impression (CDI) will be summarized with proportion of participants in each of the categories by treatment at visits 66 (OLE week 0), 78 (OLE week 48), 90 (OLE week 96), 102 (OLE week 144), and 114 (OLE week 192).

7.9.3.4. Other Analyses of the Open-Label Period

Additional descriptive analyses of the OLE data will include (1) an extension of the primary analysis NCS model respecting the blinded-phase randomized group assignments; and (2) an extension of the exposure adjusted model (Section 7.9.1.2.4).

7.10. Analyses of Cogstate/C3 Outcome Measures

The cognitive tests administered in the study include the following computerized tests from the Cogstate Battery. Each of these tests has been described in detail in the literature and hence they are described only briefly below. The order of administration is reflected in the order in which each test is described.

Detection (DET; Psychomotor Function)

The Detection test is a measure of psychomotor function and uses a well-validated simple reaction time paradigm with playing card stimuli. In this test, the playing cards all depict the same joker. The subject is asked to press the **Yes** key as soon as the card in the center of the screen turns face up. The software records the number of correct responses and expresses the number of correct responses as a proportion of the trials completed (accuracy). Data distributions for this proportion correct score are normalized using an arcsine transformation (asin prop correct). The software also records the speed of reaction times (RT) for correct responses in milliseconds (msec). Distributions of RT are then normalized through application of a logarithmic base 10 ($\log_{10}$RT) transformation.

Identification (IDN; Attention)

The Identification test is a measure of visual attention and uses a well-validated choice reaction time paradigm with playing card stimuli. In this test, the playing cards are all either red or black jokers. The subject is asked whether the card displayed in the center of the screen is red. The subject responds by pressing the **Yes** key when the joker card is red and **No** when it is black. The software records the number of correct responses and speed of reaction times for correct responses and expresses these as accuracy (asin prop correct) and speed (log_{10}RT).

One Card Learning (OCL; Visual Learning)

The One Card Learning test is a measure of visual learning and uses a well-validated pattern separation paradigm with playing card stimuli. In this test, the playing cards are identical to those found in a standard deck of 52 playing cards (without the joker cards). The subject is asked whether the card displayed in the center of the screen was seen previously in this test. The subject responds by pressing the **Yes** or **No** key. The software records the number of correct responses and speed of reaction times for correct responses and expresses these as accuracy (asin prop correct) and speed (log_{10}RT).

One Back (ONB; Working Memory)

The One Back test is a measure of working memory and uses a well-validated n-back paradigm with playing card stimuli. In this test, the playing cards are identical to those found in a standard deck of 52 playing cards (without the joker cards). The subject is asked whether the card displayed in the center of the screen is the same as the card presented immediately previously. The subject responds by pressing the **Yes** or **No** key. Because no card has been presented yet on the first trial, a correct first response is always **No**. The software records the number of correct responses and speed of reaction times for correct responses and expresses these as accuracy (asin proportion correct) and speed (log_{10}RT).

Behavioral Pattern Separation Object Test (BPSO; Recognition Memory)

The Behavioral Pattern Separation Object (BPSO) test is a measure of recognition memory. The BPSO test is divided into two sub-tests: 1) "Indoor-Outdoor" (BPET) and 2) "Old, Similar, New" (BPXT). This first test presents a series of photos of common objects to the subject who is asked to decide whether each object is primarily used indoors or outdoors. Once the “Indoor-Outdoor” test (BPET) has been completed, the "Old, Similar, New" test (BPXT) will begin. A series of images of common objects will be presented to the subject who is asked to decide if the object is the same as one they saw before, similar to one they saw before, or an entirely new object altogether. The software computes a BPSO metric, which is calculated as the difference of the probability of a "Similar" response to a "Similar" image and the probability of a "Similar" response to a "New" image (Stark et al. 2013). The probability of a ‘Similar’ response to a ‘Similar’ image is compute by summing the ‘Similar’ responses to a ‘distractor’ stimulus, divided by the sum of total responses to a ‘distractor’ stimulus. The probability of a "Similar" response to a "New" image is computed by summing the total number of ‘Similar’ responses when the stimulus was a ‘New’ item and dividing it by the sum of the total responses to a ‘New’

item. The other outcome calculated for this test is the 'percentage correct' for 'Old' and 'New' responses. The probability of a response to an 'Old' item and the probability of response to a 'New' item are summed and divided by 2, to compute the percentage correct outcome. These outcomes are described in Stark et al.,

Face Name Associative Memory Exam (FNAME; Associative Memory)

The Face Name Associative Memory Exam is a measure of associative memory using visual stimuli. The subject is shown a series of faces and names, and is asked to remember the face-name pair. Each face is then presented again, and the subject must recall and input the name that was associated with that face. The correct number of face-name pairs is recorded as an initial learning score. After a delay the subject is shown each face again, along with three names listed underneath it. The subject must select the name that was initially paired with the face. The software records the number of correct responses and speed of reaction times for correct responses and expresses these as accuracy (asin prop correct) and speed (log_{10}RT).

For each Cogstate computerized cognitive test, a single primary outcome measure was selected prior to data analysis from each test in the battery to minimize experiment-wise error rates. Each primary outcome measure was selected because it has been shown to be optimal for the detection of change.

A supporting document created by Cogstate, also called a statistical analysis plan, but limited to aspects related to the Cogstate battery, will be filed with the study documents. It contains further technical details and references related to the Cogstate battery.

The main outcome from the computerized test battery will be the C3 composite.

- C3 Composite Score (6 items): BPSO, FNAME (matching), DET, IDN, OCL, ONB (speed): computed only if z-scores are available for a minimum of four tests.

Exploratory composite outcomes will include the composite

- Learning/Working Memory Composite: OCL and ONB (accuracy): computed only if z-scores are available for both tests
- One Card Learning/One back speed accuracy composite: OCL and ONB (speed and accuracy): computed only if z-scores are available for both tests
- Psychomotor/Attention Composite: DET and IDN: computed only if z-scores are available for both tests
- Attention Domain Composite: DET, IDN, and ONB (speed): computed only if z-scores are available for all three tests
- C3 Composite Score (3 items) – Abbreviated: BPSO, FNAME (matching), OCL: computed only if z-scores are available for all three tests

7.11. Analyses of Resource Utilization Data

Resource utilization as measured by Resource Use Inventory, Brief version (RUI-B) will be compared across treatment groups in the placebo-controlled period. Continuous data will be analyzed using an NCS model and categorical data will be analyzed using Fisher's exact test.

7.12. Analyses of Biomarker Data

7.12.1. Analyses of florbetapir (Amyloid PET Imaging) Data

7.12.1.1. Placebo-Controlled Period

Analyses of the standardized uptake value ratio (SUVr) of all florbetapir parameters including the following parameters will be conducted:

- Composite SUVr and centiloids and
- SUVr in frontal, temporal, parietal, and cingulate brain regions.

The full list of florbetapir parameters is provided in Appendix 1.

The conversion of SUVr value to centiloid value is based on the following formula: Florbetapir centiloids = 183.07 * Florbetapir SUVr – 177.26.

Analysis of the change from baseline to endpoint of each florbetapir parameter will be done. The analysis will be done using an analysis of covariance (ANCOVA) model with fixed effects of baseline florbetapir result, treatment, APOE4 Carrier Status (yes/no), and age at baseline. The null hypothesis is that the difference in least square mean between the solanezumab group versus placebo equals zero.

Because of the challenges associated with scheduling and obtaining a PET scan, some endpoint PET scans were obtained after the initiation of open-label dosing. Analyses specified above will be conducted in which these observations are removed.

Additionally, due to the disassociation of visit and time caused by the COVID disruption, an analysis treating time as a continuous variable will be conducted. Each florbetapir parameter will be regressed on time, treatment, APOE4 Carrier Status (yes/no), and age at baseline, with the constraint that the two treatment groups have the same y-intercept. Slopes between the two treatment groups will be compared.

To assess the relationship of florbetapir parameters and efficacy outcomes with treatment, Spearman's rank correlation coefficient will be obtained on change from baseline to endpoint for the florbetapir composite SUVr with change from baseline to endpoint for PACC, ADL-Prevention Questionnaire, CFI, CDR-SB and C3. These correlations will be obtained separately within each treatment group and for all participants.

7.12.1.2. Open-Label Period

Florbetapir is not collected during the open-label study period.

7.12.2. Analyses of flortaucipir (Tau PET Imaging) Data

7.12.2.1. Placebo-Controlled Period

Analyses of the standardized uptake value ratios (SUVr) of all flortaucipir scans will be conducted. An NCS analysis will be used to assess the difference between treatment groups in change from baseline in SUVr at 240 weeks. Time will be treated as a continuous variable with values equal to the years between baseline and follow-up PET scans. We will assume an unstructured covariance structure. The NCS basis expansion will assume an interior knot at the median of observation times, and boundary knots at zero years and the maximum follow-up. The fixed effects will include the following terms: (i) NCS basis expansion terms (two terms), (ii) NCS basis expansion terms-by-treatment interaction (two terms), (iii) baseline flortaucipir SUVr, (iv) APOE4 Carrier Status (yes/no), and (v) age. The model is constrained to not allow a difference between treatment group means at baseline.

The list of composite brain regions for flortaucipir is provided in Appendix 1. The SUVr for each composite region will be defined as the voxel-weighted counts from the included component regions divided by the counts for the cerebellum crus region. The composite region called Middle will be considered the primary area of interest, with the other 7 regions being considered secondary.

The null hypothesis is that the treatment difference between solanezumab and placebo for the SUVr at 240 weeks is equal to zero. The analysis will be carried out in the population of participants that are randomized and have a baseline score and at least 1 post-baseline assessment. If the unstructured covariance structure matrix results in a lack of convergence, the same order of covariance structures specified for the primary analysis will be used.

To assess the relationship of flortaucipir parameters and efficacy outcomes with treatment, Spearman's rank correlation coefficient will be obtained on change from baseline to endpoint for each flortaucipir parameter with change from baseline to endpoint for PACC, ADL-Prevention Questionnaire, CFI, CDR-SB and C3. These correlations will be obtained separately within each treatment group and all participants.

7.12.2.2. Open-Label Period

Changes in the standardized uptake value ratio (SUVr) during open-label treatment of all flortaucipir composite brain regions specified in Appendix 1 will be described.

7.12.3. Analysis of Plasma Aβ and Plasma pTau-217

7.12.3.1. Placebo-Controlled Period

To evaluate the change in plasma Aβ analytes (including assayed plasma $A\beta_{1-40}$ and assayed plasma $A\beta_{1-42}$) and plasma pTau-217 after treatment, an ANCOVA will be used to compare change from baseline to each of the post-baseline visits at which plasma analytes are assayed (week 12 and week 240). This analysis will be done separately for each plasma parameter. The model for the fixed effects will include the following terms: (i) baseline plasma value, (ii)

treatment, (iii) APOE4 Carrier Status (yes/no), and (iv) age. The null hypothesis is that the difference in least square mean between the solanezumab group and placebo equals zero.

Additionally, due to the disassociation of visit and time caused by the COVID disruption, an analysis treating time as a continuous variable will be conducted. Each plasma parameter will be regressed on time, treatment, APOE4 Carrier Status (yes/no), and age at baseline, with the constraint that the two treatment groups have the same y-intercept. Slopes between the two treatment groups will be compared.

To assess the relationship of plasma parameters and efficacy outcomes with treatment, Spearman's rank correlation coefficient will be obtained on change from baseline to endpoint for each plasms Aβ parameter with change from baseline to endpoint for PACC, ADL-Prevention Questionnaire, CFI, CDR-SB and C3. Additionally, the correlation between plasma pTau-217 and both amyloid PET and tau PET will be examined. These correlations will be obtained separately within each treatment group and for all participants

7.12.3.2. Open-Label Period

Plasma Aβ analytes are not collected during the open-label study period.

7.12.4. Analyses of vMRI Data

Analyses of all the vMRI parameters including the following parameters will be conducted:

- Right hippocampal volume
- Left hippocampal volume
- Right hippocampal volume + Left hippocampal volume
- Right entorhinal volume
- Left entorhinal volume
- Right entorhinal volume + Left entorhinal volume
- Right cortical gray matter volume + Left cortical gray matter volume
- Right lateral ventricle + Left lateral ventricle

7.12.4.1. Placebo-Controlled Period

Analysis of the change from baseline to endpoint of each vMRI parameter will be done. The analysis will be done using an analysis of covariance (ANCOVA) model with fixed effects of baseline vMRI value, treatment, age at baseline, education, APOE4 Carrier Status (yes/no), and baseline florbetapir cortical SUVr. The null hypothesis is that the difference in least square mean between the solanezumab group versus placebo equals zero.

Additionally, due to the disassociation of visit and time caused by the COVID disruption, an analysis treating time as a continuous variable will be conducted. Each vMRI parameter will be regressed on time, treatment, APOE4 Carrier Status (yes/no), and age at baseline, with the constraint that the two treatment groups have the same y-intercept. Slopes between the two treatment groups will be compared.

To assess the relationship of vMRI parameters and efficacy outcomes with treatment, Spearman's rank correlation coefficient will be obtained on change from baseline to endpoint for each vMRI parameter with change from baseline to endpoint for PACC, ADL-Prevention Questionnaire, CFI, CDR-SB and C3. These correlations will be obtained separately within each treatment group and all participants.

7.12.4.2. Open-Label Period

Volumetric MRIs are not collected during the open-label study period.

7.12.5. Analyses of CSF Data

Analyses of CSF biomarkers will be done for the subset of participants who have a lumbar puncture.

The CSF parameters analyzed will include the following:

- total $A\beta_{1-40}$
- total $A\beta_{1-42}$
- free $A\beta_{1-40}$
- free $A\beta_{1-42}$
- total tau, and
- p-tau.

7.12.5.1. Placebo-Controlled Period

Analysis of the change from baseline to endpoint for each CSF parameter will be done. The dependent variable for each CSF parameter is its change from baseline to endpoint. The analysis of the change from baseline will use an analysis of covariance (ANCOVA) model with fixed effects of (i) baseline CSF, (ii) treatment, (iii) APOE4 Carrier Status (yes/no), and (iv) age at baseline. The null hypothesis is that the difference in least square mean between the solanezumab group versus placebo equals zero.

Additionally, due to the disassociation of visit and time caused by the COVID disruption, an analysis treating time as a continuous variable will be conducted. Each CSF parameter will be regressed on time, treatment, APOE4 Carrier Status (yes/no), and age at baseline, with the constraint that the two treatment groups have the same y-intercept. Slopes between the two treatment groups will be compared.

To assess the relationship of CSF parameters and efficacy outcomes with treatment, Spearman's rank correlation coefficient will be obtained on change from baseline to endpoint for each CSF parameter with change from baseline to endpoint for PACC, ADL-Prevention Questionnaire, CFI, CDR-SB and C3. These correlations will be obtained separately within each treatment group and all participants.

7.12.5.2. Open-Label Period

CSF biomarkers are not collected during the open-label study period.

7.13. Safety Analyses

Two different safety datasets will be created. The safety placebo-controlled (safety-PC) dataset consists of visits 1-66 and allows for comparisons between solanezumab-treated participants and placebo-treated participants; the safety open-label dataset consists of visits 66-117 during which all participants are treated with solanezumab. Participants will be analyzed according to the treatment group to which they were randomized. All safety measures will be summarized for each dataset in which they were collected.

Safety will be assessed by summarizing and analyzing AEs, laboratory analytes, vital signs, electrocardiograms (ECGs), magnetic resonance imaging (MRI), immunogenicity and additional safety data collected during the treatment period.

Safety analyses for the treatment period will include comparisons between solanezumab and placebo. All hypotheses will be tested at a 2-sided 0.05 significance level. No adjustments for multiple comparisons will be made.

7.13.1. Analyses of Adverse Events

Adverse events will be coded according to established Medical Dictionary for Regulatory Activities (MedDRA) terms and summarized by MedDRA System Organ Class and Preferred Term.

7.13.1.1. Analyses of TEAEs, SAEs, Discontinuation Due to AEs

Treatment-emergent adverse events (TEAEs) will be defined as events that first occurred or worsened after first infusion of study drug (generally Visit 6). Should there be insufficient data for AE start date or stop date to assess for treatment emergence, the adverse event will be considered treatment emergent.

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

Summaries of AEs by decreasing frequency of preferred term within system organ class will be provided for the following:

- Preexisting conditions
- TEAEs
- TEAEs by maximum severity
- TEAEs occurring in greater than or equal to 5% of participants by preferred term
- SAEs ,and
- AEs reported as reason for discontinuation.

These summaries will include numbers and percentages of participants with TEAEs. Treatment comparisons will be carried out using Fisher's exact test.

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

7.13.1.2. Subgroup Analyses of TEAEs

Subgroup analyses of TEAEs will also be done. The subgroups will include age at baseline (≥65 and <75 years, ≥75 and <85 years), gender, and race (dichotomized based on distribution of race at baseline). For each subgroup variable, a Breslow-Day test of homogeneity of odds-ratios between treatment groups will be performed on the incidence of each TEAE.

7.13.1.3. Analyses of TEAEs by Dose

For participants treated with solanezumab, summaries of treatment-emergent adverse events will be displayed with events assigned to the dose (400 mg versus 1600 mg for participants assigned to solanezumab, with events occurring at the 800 mg level assigned to the 1600 mg group, and low dose placebo and high dose placebo for participants assigned to placebo based on the blinded dose level) that the participant was taking at the time of the event. At the first occurrence of a high dose for a participant, all subsequent observations will also be considered as high dose. Events will be reported as frequencies and rates, adjusted for the amount of time until the event for participants who experienced the event or total time on the specified dose for those who did not experience the event. Treatment-emergence will be defined separately for each dose level, with the time period immediately preceding the dose as the baseline period. An adverse event that first appears or worsens in severity subsequent to the dose will be considered treatment-emergent.

7.13.1.4. Analyses of Specific Clusters of TEAEs

In addition, the proportion of participants within the following specific clusters of TEAEs will be summarized and treatment comparisons will be conducted using Fisher's exact test:

- infusion-related reactions including anaphylaxis and urticiaria, overall and broken out by immediate (same day as infusion) versus delayed
- suicidal ideation or behaviors
- hemorrhagic stroke
- cardiac ischaemic events
- cardiac arrhythmias, and
- amyloid-related imaging abnormality – hemorrhage (ARIA-H, also known as microhemorrhage).

The listing of preferred terms of the TEAEs corresponding to each specific cluster above is provided in Appendix 2.

7.13.2. Analyses of Laboratory Data

Laboratory measurements (chemistry, hematology, and urinalysis) will be analyzed using continuous data (change from baseline) and categorical data (proportion of treatment-emergent

abnormalities). If there are multiple records of laboratory measurements at baseline or post-baseline visit, the last record will be used.

Summaries and analyses for continuous data will be performed separately using conventional units (CN units) and also International System of Units (SI units). Summaries and analyses for categorical data will be performed using the International System of Units (SI units). The list of laboratory parameters analyzed is provided in Appendix 3.

7.13.2.1. Analyses of Continuous Data

Change from baseline to each postbaseline visit at which laboratory samples are collected will be assessed using an ANCOVA model with treatment as independent factors and baseline value and age as covariate in the model. This analysis will be done separately for each laboratory analyte.

7.13.2.2. Analyses of Categorical Data

7.13.2.2.1. Treatment-Emergent High, Low, or Abnormal Analyses

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 laboratory measurements are values which are low or normal at the baseline visit and fall into the high category at post-baseline visit(s). Similarly, treatment-emergent low laboratory measurements are values which are high or normal at the baseline visit and fall into the low category at post-baseline visit(s). Treatment-emergent abnormal laboratory measurements are values which are normal at baseline and fall into the high category or low category at postbaseline visit(s).

Table A4.SAP.3. describes which participants will be included in the denominator and numerator when assessing treatment-emergent abnormalities.

Table A4.SAP.3. Denominator and Numerator for Treatment-Emergent Laboratory Analyses

Treatment-Emergence Category	Denominator	Numerator
Treatment-Emergent High	Participants who are normal or low at baseline	Participants who are high postbaseline (among participants who were low or normal at baseline)
Treatment-Emergent Low	Participants who are normal or high at baseline	Participants who are low postbaseline (among participants who were high or normal at baseline)
Treatment-Emergent Abnormal	Participants who are normal at baseline	Participants who are low or high postbaseline (among participants who were normal at baseline)

7.13.2.2.2. Shift Tables and Additional Analyses

For urinalysis categorical parameters, baseline normal to post-baseline abnormal shifts will be summarized at any time. Fisher's exact test will be used to compare shifts in urinalysis parameters between treatment groups.

The proportion of participants with treatment-emergent changes from a low value or normal value at all baseline at any time in the following analytes will be summarized by treatment group. Changes of interest are:

- ALT: The number and percentage of participants with a measurement greater than or equal to 1 time (1X), 3 times (3X), 5 times (5X), 10 times (10X), and 20 times (20X) the performing lab upper limit of normal (ULN)
- AST: The number and percentage of participants with a measurement greater than or equal to 1 time (1X), 3 times (3X), 5 times (5X), 10 times (10X), and 20 times (20X) the performing lab ULN
- ALP: The number and percentage of participants with a measurement greater than or equal to 2 times (2X), and 3 times (3X) the performing lab ULN
- TBL: The number and percentage of participants with a measurement greater than or equal to 2 times (2X), 5 times (5X), and 8 times (8X) the performing lab ULN
- DBL: The number and percentage of participants with a measurement greater than or equal to 2 times (2X) and 5 times (5X) the performing lab ULN
- GGT: The number and percentage of participants with a measurement greater than or equal to 2 times (2X) the performing lab ULN

Comparisons between treatment groups will be carried out using Fisher's exact test.

Additionally, plots of TBL versus ALT values and TBL versus AST values will be produced.

7.13.3. Analyses of Vital Signs and Weight

Vital sign measurements (including weight) will be analyzed using continuous data (change from baseline) and categorical data (proportion of treatment-emergent abnormalities).

If there are multiple records of vital sign or weight measurements at baseline or post-baseline visit(s), the last record will be used.

The analyses will be done for the following vital sign measurements and weight:

- systolic blood pressure
- diastolic blood pressure
- pulse
- temperature, and

- weight.

7.13.3.1. Analyses of Continuous Data

Change from baseline to each post-baseline visit at which the vital signs are taken will be assessed using an ANCOVA model with treatment as independent factors and baseline value and age as covariate in the model. This analysis will be done separately for each vital sign parameter and weight.

7.13.3.2. Analyses of Categorical Data

Treatment differences in the proportion of participants with treatment-emergent high, treatment-emergent low or treatment-emergent abnormal vital sign measurement will be assessed between treatment groups using Fisher's exact test at (1) anytime (2) each post-baseline visit.

Treatment-emergent high vital sign measurements are the values which are low or normal at the baseline visit and fall into the high category at post-baseline visit(s). Similarly, treatment-emergent low vital sign measurements are the values which are high or normal at the baseline visit and fall into the low category at post-baseline visit(s). Treatment-emergent abnormal vital sign measurements are the values which are normal at the baseline visit and fall into the high or low category at post-baseline visit(s). The criteria for post-baseline low and post-baseline high are provided in Table A4.SAP.4.

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 greater than or equal to 4 percent of baseline body weight will be compared between treatment groups using Fisher's exact test at each visit and at any time. In addition, categories of BMI (underweight = < 18.5, normal = ≥ 18.5 - < 25, overweight = ≥ 25 to < 30, obese = ≥ 30) will be compared between treatment groups using Fisher's exact test at (1) anytime and (2) each post-baseline visit.

Table A4.SAP.4. Criteria for Abnormal Vital Signs and Weight

Vital Sign Parameter (Unit)	Postbaseline Low Criteria	Postbaseline High Criteria
Systolic Blood Pressure (mmHg)	Absolute value ≤90 and ≥20 decrease from baseline	Absolute value ≥160 and ≥20 increase from baseline
Diastolic Blood Pressure (mmHg)	Absolute value ≤50 and ≥10 decrease from baseline	Absolute value ≥100 and ≥10 increase from baseline
Pulse (bpm)	Absolute value <50 and ≥15 decrease from baseline	Absolute value >100 and ≥15 increase from baseline
Weight	≥4% decrease	≥4% increase
Vital Sign Parameter	**Postbaseline Criteria for Abnormality**	
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.

7.13.4. Analyses of Electrocardiograms

Electrocardiogram (ECG) measurements will be analyzed using continuous data (change from baseline) and categorical data (proportion of treatment-emergent abnormalities).

Since ECG is measured in triplicates, the average of triplicates will be used for baseline and postbaseline visit(s). If there are multiple records after averaging ECG triplicates within a visit, the last record of averages will be used.

The analyses will be done for the following ECG measurements:

- heart rate
- PR interval
- QRS duration
- QT interval, and
- QTc interval.

All analyses of QTc will be carried out using the Fridericia correction (QTcF) method. These summaries will include data from each visit at which ECG measures are performed.

7.13.4.1. Analyses of Continuous Data

Change from baseline to each postbaseline visit at which ECG measurements are taken will be assessed using an ANCOVA model. The model for the fixed effects will include terms for the following independent effects: baseline ECG score, treatment. This analysis will be done separately for each ECG parameter.

7.13.4.2. Analyses of Categorical Data

Incidence of treatment-emergent abnormal ECGs will be compared between treatment groups by Fisher's exact test at (1) anytime and (2) each post-baseline visit.

Abnormal ECG criteria are presented in Table A4.SAP.5 and criteria for abnormal QTcF prolongation are presented in Table A4.SAP.6.

Table A4.SAP.5. Criteria for Abnormal ECG Parameters

ECG Parameter	Low Criteria	High Criteria
Heart Rate	≤ 40 bpm	≥ 120 bpm
PR Interval	≤ 120 msec	≥ 220 msec
QT Interval		> 500 msec
QRS Duration	< 60 msec	≥ 120 msec

Abbreviations: bpm = beats per minute; ECG = electrocardiogram; msec = millisecond.

Table A4.SAP.6. Criteria for Identifying Participants with a Prolonged ECG QTcF Interval

Criterion Number	Criterion
1	In adult males, QTcF > 450 msec; in adult females, QTcF > 470 msec
2	QTcF ≥ 500 msec
3	Increase of > 0 msec and ≤ 30 msec relative to baseline
4	Increase of > 30 msec and ≤ 60 msec relative to baseline
5	Increase of > 60 msec relative to baseline
6	Increase of > 60 msec relative to baseline and QTcF ≥ 500 msec

Abbreviations: ECG = electrocardiogram; msec = millisecond; QTcF = Fridericia's corrected QT interval.

Treatment-emergent high ECG parameters (heart rate, PR interval, QRS duration, QT and QTcF intervals) are the values which are low or normal at the baseline visit and fall into the high abnormal categories at post-baseline visit(s) in Table A4.SAP.5 and Table A4.SAP.6 postbaseline. Similarly, treatment-emergent low ECG parameters (heart rate, PR interval, QRS duration) are the values which are high or normal at the baseline visit and fall into the low abnormal categories at postbaseline visit(s) above.

7.13.5. Analyses of MRI Data

To evaluate any changes in MRI data following treatment, Fisher's exact test will be used to compare frequencies of responses in the MRI parameters.

Frequencies and percentages of the following amyloid-related imaging abnormality – edema (ARIA-E, also known as vasogenic edema) and ARIA – hemorrhage (ARIA-H

, also known as microhemorrhage) parameters will be summarized:

- ARIA-E:
  - Status compared to the previous MRI(s) (increased, unchanged, partial resolution, or complete resolution)
  - ARIA-E by APOE4 genotype
- ARIA-H:
  - Number of ARIA-H (≤4, 5–10, >10, or no presence),
  - Presence of superficial siderosis,

- Baseline to each visit (at which MRI is done) changes (increase in size of pre-existing ARIA-H, increase in number of ARIA-H, no change, partial resolution, or complete resolution), and
- ARIA-H by APOE4 genotype.

To investigate the effect of dose increases on MRI safety parameters, ARIA-E and ARIA-H rates for the first MRI subsequent to the dose increase will be displayed.

7.13.5.1. Analysis of MRI Data by Dose

For participants treated with solanezumab, summaries of ARIA-E and ARIA-H events will be displayed with events assigned to the dose (400 mg versus 1600 mg for participants assigned to solanezumab, with events occurring at the 800-mg level assigned to the 1600-mg group, and low-dose placebo and high-dose placebo for participants assigned to placebo based on the blinded dose level) that the participant was taking at the time of the event. At the first occurrence of a high dose for a participant, all subsequent observations will also be considered as high dose.

7.13.6. Analyses of Immunogenicity Data

7.13.6.1. Immunogenicity Definitions

Figure A4.SAP.2. provides an overview of the immunogenicity assay process. At a high level, an individual sample is potentially examined multiple times, in a hierarchical procedure, to produce a sample solanezumab ADA assay result and potentially a sample solanezumab neutralizing antibody (NAb) assay result. The drug tolerance of each assay, the possible values of titers, and the cutpoints applied are operating characteristics of the assays and the hierarchical testing procedure of Figure A4.SAP.2.

It can be the case that the presence of high concentrations of solanezumab will affect the measurements of the presence of ADA or NAb, and conversely high levels of ADA or NAb may affect the measurement of LY concentration. Thus an LY drug concentration, assessed from a sample drawn at the same time as the ADA sample, plays a key role in clinical interpretation of a sample when the laboratory result is Not Detected, as shown in Figure A4.SAP.2.

The rest of this section defines the component concepts of Figure A4.SAP.2 in greater detail.

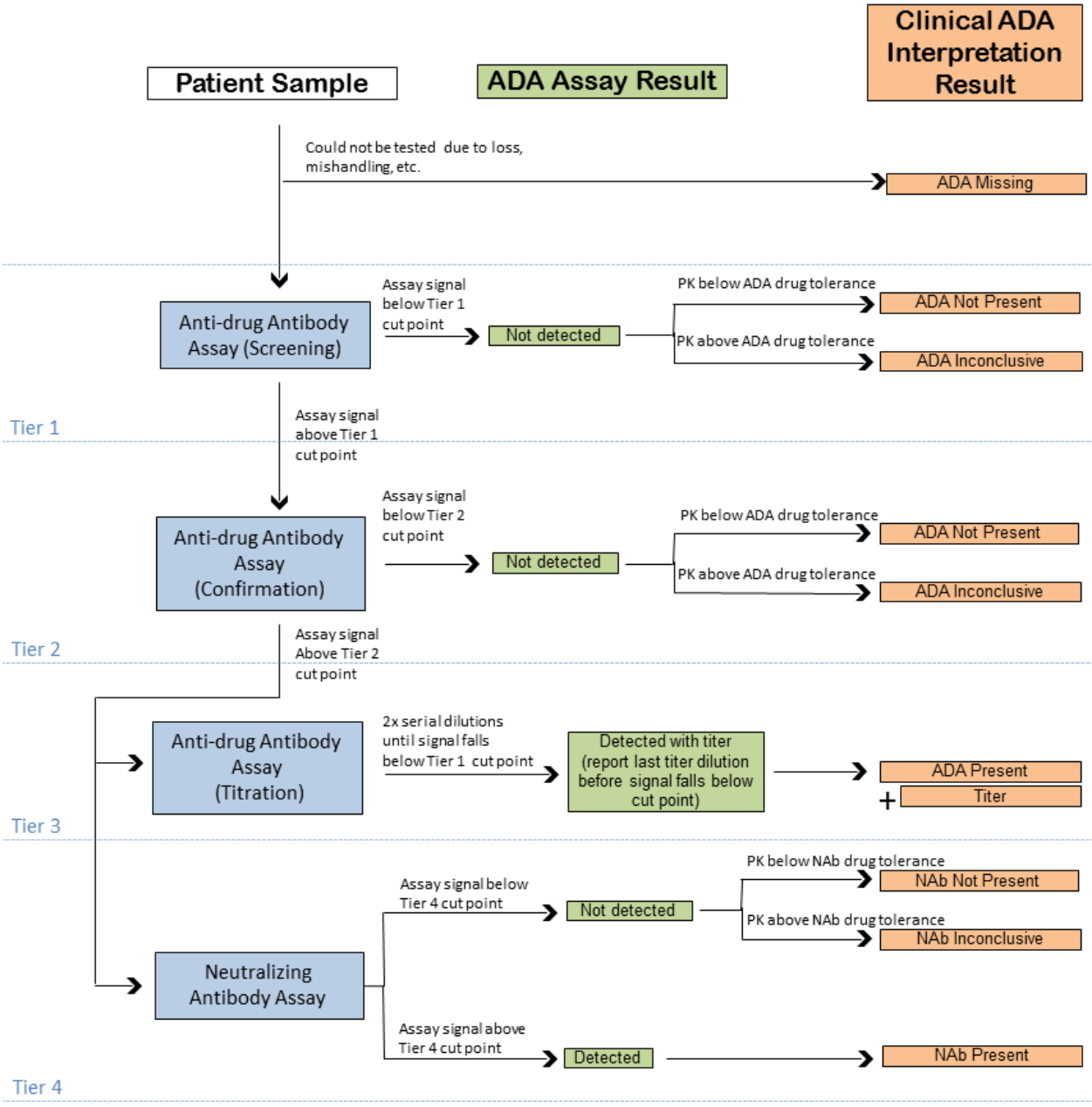


Figure A4.SAP.2. Flow Chart of ADA Assessment with Clinical Interpretation of the Various Result Possibilities

Abbreviations: ADA = anti-drug antibody; NAb = neutralizing antibody; PK = pharmacokinetic.

Definitions of ADA statuses (Table A4.SAP.7) and clinical interpretation results (Table A4.SAP.8) are provided below.

Table A4.SAP.7. Sample ADA Assay Results

Sample Laboratory Result	Explanation
Detected	ADA are detected and confirmed.
Not Detected	The raw result as reported from the laboratory indicates ADA not detected. The clinical interpretation of such results depends on other factors (see below).
NO TEST, QNS (quantity not sufficient), etc.	Sample exists but was un-evaluable by the assay

Abbreviation: ADA = anti-drug antibody.

Table A4.SAP.8. Sample Clinical ADA Interpretation Results

Sample Clinical Interpretation	Explanation
ADA Present	ADA assay result is Detected
ADA Not Present	ADA assay result is Not Detected and simultaneous drug concentration is at a level that has been demonstrated to not interfere in the ADA detection method (i.e., drug concentration is below the assay's drug tolerance level). If drug concentration is not available for a treatment-period sample, the sample is inconclusive (see below). For pre-treatment samples and participants receiving placebo, drug concentration is not assessed and is assumed to be below the assay's drug tolerance level.
ADA Not Detected, Drug Concentration Not Available	If drug concentration is expected per protocol but not available, the immunogenicity sample is "ADA Not Detected, Drug Concentration Not Available" for the purpose of participant listings. For the purpose of TE ADA computation (see below), these samples are taken to be ADA Not Present.
ADA Inconclusive	ADA assay result is Not Detected but drug concentration in the sample is at a level that can cause interference in the ADA detection method, or drug concentration is planned per protocol but is not available.
ADA Missing	ADA sample not drawn, QNS, not tested, etc., causing there to be no laboratory result reported or the result is reported as "no test".

Abbreviations: ADA = anti-drug antibody; TE = treatment-emergent; QNS = quantity not sufficient.

Parallel terminology applies for Neutralizing ADA (NAb) Detected, NAb Not Detected, NAb Present, NAb Not Present, NAb Inconclusive, NAb Missing. Anti-drug antibodies and Neutralizing ADA (NAb) are distinct assays and have different assay operating characteristics.

A post-baseline immunogenicity sample with ADA Present is said to have TE ADA titer if the titer meets the criteria to classify the participant as TE ADA+.

7.13.6.1.1. Definitions of Immunogenicity Assessment Periods

- **Immunogenicity Baseline Observations:** Baseline period for immunogenicity assessment for each participant includes all observations on or prior to the date of the first administration of study drug. In instances where multiple baseline observations are collected, to determine participant ADA status the lowest titer/not detected is used to determine treatment-emergent status (see below).
- **Immunogenicity Post-baseline Period Observations:** Post-baseline period observations for each participant includes all observations after the first administration of study drug.

7.13.6.1.2. Definitions of Participant ADA Status

- **Participant evaluable for treatment-emergent ADA:** A participant is evaluable for TE ADA if the participant has a non-missing baseline ADA result, and at least one non-missing post-baseline ADA result.
- **Treatment-emergent ADA positive (TE ADA+) participant:** A participant who is evaluable for TE ADA is treatment-emergent ADA positive (TE ADA+) if either of the following holds:
    a. The participant has baseline status of ADA Not Present and at least one post-baseline status of ADA Present with titer ≥2 × MRD, where the MRD is the minimum required dilution of the ADA assay.
    b. The participant has baseline and post-baseline status of ADA Present, with the post-baseline titer being 2 dilutions (4-fold) greater than the baseline titer. That is, the participant has baseline status of ADA Present, with titer 1:B, and at least one post-baseline status of ADA Present, with titer 1:P, with P/B ≥4.
- **Treatment-emergent ADA Inconclusive participant:** A participant who is evaluable for TE ADA is TE ADA Inconclusive if ≥20% of the participant's post-baseline samples, drawn pre-dose, are ADA Inconclusive and all remaining post-baseline samples are ADA Not Present.
- **Treatment-emergent ADA negative (TE ADA-) participant:** A participant who is evaluable for TE ADA is TE ADA negative (TE ADA-) when the participant is not TE ADA+ and the participant is not TE ADA Inconclusive.

7.13.6.2. Immunogenicity Statistical Analyses

A listing will be provided of all immunogenicity assessments for those participants who at any time had ADA Present. This includes the laboratory ADA assay result (Detected or Not Detected), solanezumab concentration from a simultaneous pharmacokinetic sample, and the clinical interpretation result that combines these (ADA Present, ADA Not Present, ADA Inconclusive, Missing). When Detected, a titer will be included, and TE ADA+ observations will be flagged. Also included, when the NAb assay was performed, will be the laboratory NAb assay result (Detected or Not Detected) and the NAb clinical interpretation result (NAb Present, NAb Not Present, NAb Inconclusive, Missing).

For the remainder of this section, mention of ADA result and NAb result will refer to the respective clinical interpretation result.

The number and proportion of participants who are TE ADA+ will be tabulated by treatment group, where proportions are relative to the number of participants who are TE ADA evaluable, as defined above. This analysis will include all post-baseline observations and will examine solanezumab ADA and solanezumab NAb. The tabulation will include the number and proportion of participants with ADA Present at baseline, and also the number and proportion of TE ADA+ participants exhibiting NAb+. For analysis sets involving both solanezumab and placebo, results between solanezumab and placebo-treated groups will be compared using a Cochran-Mantel-Haenszel test stratified by study, for TE ADA+ and for TE ADA+ with NAb+.

A summary will be provided of the number and percentage of solanezumab-treated participants experiencing TEAE (overall and by PT) by participant TE ADA status (TE ADA+, TE ADA-, TE ADA Inconclusive). The PT will be ordered by decreasing incidence in TE ADA+ status group.

In order to assess the clinical relevance of TE immunogenicity, specific acute and non-acute AEs will be evaluated. A listing will be provided of these events for all participants who were TE ADA+. This listing includes a time course of ADA (clinical interpretation result, plus flags for samples with TE ADA titer or NAb+ samples) along with the AE. Listings with the same structure will be provided of (i) these events of interest for participants who had ADA Present at any time (including baseline) but who were not TE ADA+, and (ii) all TEAE for TE ADA+ participants.

The time from first dose to the first TE ADA titer will be summarized cumulatively for representative time intervals over the course of the study.

Sample clinical interpretation result, represented as ADA Not Present, ADA Inconclusive, or as a titer value for ADA Present samples, will be displayed in a shift table from baseline to maximum post-baseline. For participants with a post-baseline ADA Present sample, the maximum post-baseline value is the maximum titer observed at any time post-baseline. For participants without a post-baseline ADA Present sample, the maximum value is ADA Inconclusive, if such a result was observed. If no participants had maximum value of ADA Inconclusive, then ADA Inconclusive will not be displayed in the table.

7.13.7. Analyses of Additional Safety Data

7.13.7.1. Columbia-Suicide Severity Rating Scale

Suicidal ideation, suicidal behavior, and self-injurious behavior without suicidal intent, based on the Columbia-Suicide Severity Rating Scale (C-SSRS), will be listed by participant and visit. Only participants that show suicidal ideation/behavior or self-injurious behavior without suicidal intent will be displayed (i.e., if a participant answers are all ‘no’ for the C-SSRS, then that participant will not be displayed). However, if a participant reported any suicidal ideation/behavior or self-injurious behavior without suicidal intent at any time point then all their ideation and behavior will be displayed, even if not positive.

Suicidal ideation, suicidal behavior, and self-injurious behavior without suicidal intent occurring during treatment, based on the Columbia-Suicide Severity Rating Scale (C-SSRS), will be summarized by treatment. In particular, for each of the following events, the number and percent of participants with the event will be enumerated by treatment: completed suicide, nonfatal 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, nonspecific active suicidal thoughts, wish to be dead, and self-injurious behavior without suicidal intent. Although not suicide-related, the number and percent of participants with non-suicidal

self-injurious behavior occurring during the treatment period will also be summarized by treatment.

In addition, the number and percent 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 number and percent 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.

- 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.

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.

7.13.7.1.1. Columbia Suicide-Related Severity by Dose

For participants treated with solanezumab, summaries of suicidal ideation, suicidal behavior, and self-injurious behavior without suicidal intent occurring during treatment, based on the Columbia-Suicide Severity Rating Scale (C-SSRS), will be displayed with events assigned to the dose (400 mg versus 1600 mg for participants assigned to solanezumab, with events occurring at the 800-mg level assigned to the 1600-mg group, and low dose placebo and high dose placebo for participants assigned to placebo based on the blinded dose level) that the participant was taking at the time of the event. At the first occurrence of a high dose for a participant, all subsequent observations will also be considered as high dose.

7.13.7.2. Assessment of Psychological Well-Being

Change from baseline to each post-baseline visit at which Psychological Well-Being is assessed will be analyzed using an ANCOVA model. The model for the fixed effects will include terms for the following independent effects: baseline Psychological Well-Being score, treatment.

The Geriatric Depression Scale and the State Trait Anxiety Inventory will also be looked at separately to compare changes between treatments.

7.14. Analyses of Study Drug Exposure and Infusion

Duration of exposure to treatment will be summarized and compared between treatment groups using an ANOVA with treatment as the independent factor.

Duration of exposure (in days) will be derived as follows:

Duration of exposure = Date of study completion/termination – Date of first infusion + 1 day.

Because dosing occurs at study visits, participants who attend all visits and successfully receive complete infusions are automatically compliant with treatment. Any infusion at which 75% or more of the infusate is given will be considered a complete infusion.

Summary statistics will be provided for

- the proportion of participants who received complete infusions at each visit
- duration of complete infusion at each visit, and
- volume of complete infusion at each visit.

The proportion of participants who received complete infusion between treatments will be compared using Fisher's exact test. ANOVA will be used to compare treatment groups for duration of complete infusion, and volume of complete infusion. Treatment will be the independent factor included in the ANOVA.

7.15. Subgroup Analyses

To assess the effects of various demographic and baseline characteristics on treatment outcome, subgroup analyses for the primary endpoint, PACC, will be performed based on the following variables:

- Gender: Male/Female
- Age: ≥ 65 years and < 75 years, ≥ 75 years and < 85 years
- Race: Dichotomized based on distribution of race at baseline.
- APOE4 Carrier Status: Carrier defined as ε2/ε4, ε3/ε4, or ε4/ε4 genotype; Non-Carrier defined as all other genotypes.
- APOE4 Genotype: ε2/ε4, ε3/ε4, ε4/ε4, No ε4
- Family History of AD: Yes/No
- Reason for Discontinuation: Categories as defined by the subject disposition case report form. Due to sparseness within certain categories, categories may be collapsed after seeing the actual distributions.

The primary outcome measures will be modeled using an NCS approach. This general model will include effects for (i) NCS basis expansion terms, (ii) subgroup (iii) NCS basis expansion terms-by-treatment-by-subgroup interaction, (iv) PACC test version administered, (v) age, (vi) education, (vii) APOE4 Carrier Status (yes/no), and (viii) baseline florbetapir cortical SUVr. This model will constrain the baseline treatment group means to be the same within subgroup levels, but allow different baseline means for the subgroups. Redundant terms will be dropped from the model in those cases in which the subgroup of interest is overlapping with this general model. The interaction will be tested by likelihood ratio test comparing the full model to the reduced model without any subgroup interaction effects.

To evaluate dose effect on the primary outcome measure, marginal structural models, inverse probability of treatment weighting (MSM, IPTW) methodology will be used (Lipkovich et al, 2008). Dose profiles for mean changes from baseline will be evaluated using weighted MSM with a repeated measures model. To adjust for selection bias due to non-random dose assignment and dropouts, participant-specific time-dependent weights will be determined as products of (i) stable weights based on inverse probability of receiving the sequence of dose assignments that was actually received by a participant up to given time multiplied by (ii) stable weights based on inverse probability of participant remaining on treatment by that time.

7.16. Protocol Violations

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

- Informed consent violation.

- Did not have an assessment of PACC at any of the visits at which the scale was scheduled to be assessed.
- 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.)
- Had unqualified raters or raters with substantial scoring errors for the PACC.
- Greater than or equal to 20% of infusions are incomplete.
- One or more missed infusions.
- Violated the protocol in any other way.

7.17. Interim Analyses and Data Monitoring

The ATRI-Coordinating Center (ATRI-CC) currently has a Data Safety Monitoring Board (DSMB) that reviews the safety of all subjects enrolled in trials on an ongoing basis. The initial task of the DSMB will be to review the protocol and consent forms to identify any necessary modifications. If modifications are necessary, revisions will be reviewed by the DSMB prior to its recommendation on initiation of the project. The DSMB, based on its review of the protocol, will identify the data parameters and format of the information to be regularly reported. The DSMB will be informed of the occurrence of any SAEs and immediately notified of fatal or life-threatening events. The DSMB may at any time request additional information from the ATRI-CC. The DSMB will initially be provided with data blinded to treatment status, but they may request unblinded data if there is a safety concern. The DSMB and National Institute on Aging (NIA) representative will meet in person or by conference call on a quarterly basis. Based on the review of safety data, the DSMB will make recommendations regarding the conduct of the study. These may include amending safety monitoring procedures, modifying the protocol or consent, terminating the study or continuing the study as designed. The DSMB will also be informed in a real-time basis of all immediately reportable AEs (FDA-defined serious AE). All reports are stripped of identifying information. Using the ATRI safety review process and the DSMB, there is substantial oversight and case review to alert the investigators, in a timely manner, to any safety issues that may arise. For further details, please refer to the DSMB charter.

7.18. 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 AEs, provided as a dataset which will be converted to an XML file. Both SAEs and ‘Other’ AEs are summarized: by treatment group, by MedDRA preferred term.

- An AE is considered 'serious' whether or not it is a treatment emergent adverse event (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, and
    - the number of events experienced.
- Consistent with www.ClinicalTrials.gov requirements, 'Other' AEs that occur in fewer than 5% of participants or participants in every treatment group may not be included if a 5% threshold is chosen (5% is the minimum threshold).

8. References

Andersen SW, Millen BA. On the practical application of mixed effects models for repeated measures to clinical trial data. *Pharm Stat*. 2013;12(1):7-16. https://doi.org/10.1002/pst.1548

Carpenter JR, Kenward MG. *Multiple Imputation and Its Application*. 1st ed. John Wiley & Sons Ltd; 2013.

Hastie TJ. Generalized additive models. In: Chambers JM, Hastie TJ, eds. Statistical Models in S. Chapman & Hall/CRC. 1992.

Liang K, Zeger S. Longitudinal data analysis of continuous and discrete responses for pre-post designs. Sankhyā: The Indian Journal of Statistics, Series B (1960-2002). 2000;62(1):134-148.

Lipkovich I, Adams DH, Mallinckrodt C, et al. Evaluating dose response from flexible dose clinical trials. *BMC Psychiatry*. 2008;8:3. https://doi.org/10.1186/1471-244x-8-3

Sparks J, Lo AC, Evans CD, et al. Reproducibility of applying multiple statistical analyses to data from the TRAILBLAZER-ALZ clinical trial of donanemab in early symptomatic Alzheimer's disease. *Alzheimers Dement*. 2021;17(S9):e057700. https://doi.org/10.1002/alz.057700

9. Appendices

Appendix 1. List of Florbetapir and Flortaucipir Parameters

The following is the list of brain regions for which florbetapir standardized uptake value ratios (SUVr) will be provided.

1	xlaal_frontal_med_orb[a]
2	new_temporal_2[a]
3	lprecuneus_gm[a]
4	lnew_parietal[a]
5	llposterior_cingulate_2[a]
6	lanterior_cingulate_2[a]
7	blcere_all
8	Composite_Summary

[a] Regions used in calculation of the composite summary SUVr.

The following is the list of 8 composite brain regions for which flortaucipir standardized uptake value ratios (SUVr) will be provided.

Basis	Region	Components	N Voxels Left / Right
Staging	Early	MGH-GTM amygdala MGH-GTM entorhinal MGH-GTM parahippocampal	258 / 272 256 / 241 328 / 316
Staging	Middle	MGH-GTM fusiform MGH-GTM inferiortemporal MGH-GTM middletemporal MGH-GTM inferiorparietal	1495 / 1487 1771 / 1715 1737 / 1816 1927 / 2214
Staging	Late	MGH-GTM lateraloccipital MGH-GTM posteriorcingulate MGH-GTM superiorparietal MGH-GTM supramarginal	1622 / 1403 411 / 500 1703 / 1586 1717 / 1614
Staging	Early Extension	MGH-GTM amygdala MGH-GTM entorhinal MGH-GTM parahippocampal MGH-GTM fusiform MGH-GTM inferiortemporal MGH-GTM middletemporal MGH-GTM inferiorparietal	258 / 272 256 / 241 328 / 316 1495 / 1487 1771 / 1715 1737 / 1816 1927 / 2214

Basis	Region	Components	N Voxels Left / Right
Staging	Late Extension	MGH-GTM fusiform	1495 / 1487
		MGH-GTM inferiortemporal	1771 / 1715
		MGH-GTM middletemporal	1737 / 1816
		MGH-GTM inferiorparietal	1927 / 2214
		MGH-GTM lateraloccipital	1622 / 1403
		MGH-GTM posteriorcingulate	411 / 500
		MGH-GTM superiorparietal	1703 / 1586
		MGH-GTM supramarginal	1717 / 1614
Staging	Combo	MGH-GTM amygdala	258 / 272
		MGH-GTM entorhinal	256 / 241
		MGH-GTM parahippocampal	328 / 316
		MGH-GTM fusiform	1495 / 1487
		MGH-GTM inferiortemporal	1771 / 1715
		MGH-GTM middletemporal	1737 / 1816
		MGH-GTM inferiorparietal	1927 / 2214
		MGH-GTM lateraloccipital	1622 / 1403
		MGH-GTM posteriorcingulate	411 / 500
		MGH-GTM superiorparietal	1703 / 1586
		MGH-GTM supramarginal	1717 / 1614
Anatomy	Lateral Temporal	MGH-GTM inferiortemporal	1771 / 1715
		MGH-GTM middletemporal	1737 / 1816
		MGH-GTM superiortemporal	1774 / 1601
		MGH-GTM transversetemporal	155 / 117
		MGH-GTM bankssts	379 / 355
Anatomy	Parietal	MGH-GTM inferiorparietal	1927 / 2214
		MGH-GTM superiorparietal	1703 / 1586
		MGH-GTM supramarginal	1717 / 1614
		MGH-GTM postcentral	1268 / 1124
		MGH-GTM precuneus	1400 / 1445

Appendix 2. List of Preferred Terms for Specific Clusters of TEAEs

Specific TEAE cluster	Preferred terms
Infusion-related reactions including anaphylaxis and urticiaria	infusion related reaction procedural dizziness procedural headache procedural hypotension procedural nausea procedural hypertension anaphylactic reaction anaphylactic shock anaphylactic transfusion reaction anaphylactoid reaction anaphylactoid shock first use syndrome Kounis syndrome shock Type I hypersensitivity urticaria
Suicidal Ideation or Behaviors	completed suicide depression suicidal intentional overdose intentional self-injury multiple drug overdose intentional poisoning deliberate self-injurious behavior self-injurious ideation suicidal behavior suicidal ideation suicide attempt

Specific TEAE cluster	Preferred terms
Hemorrhagic Stroke	basal ganglia haemorrhage basal ganglia stroke brain stem haemorrhage brain stem stroke cerebellar haematoma cerebellar haemorrhage cerebral haematoma cerebral haemorrhage cerebrovascular accident cerebrovascular disorder haemorrhage intracranial haemorrhagic cerebral infarction haemorrhagic stroke haemorrhagic transformation stroke intracerebral haematoma evacuation intracranial haematoma intraventricular haemorrhage meningorrhagia putamen haemorrhage spinal cord haemorrhage spinal cord haematoma stroke in evolution subarachnoid haemorrhage thalamus haemorrhage

Specific TEAE cluster	Preferred terms
Cardiac Ischaemic Events	acute coronary syndrome acute myocardial infarction coronary artery embolism coronary artery occlusion coronary artery reocclusion coronary artery thrombosis coronary bypass thrombosis Kounis syndrome myocardial infarction myocardial reperfusion injury papillary muscle infarction post procedural myocardial infarction postinfarction angina silent myocardial infarction angina pectoris angina unstable arteriosclerosis coronary artery arteriospasm coronary coronary angioplasty coronary arterial stent insertion coronary artery bypass coronary artery disease coronary artery dissection coronary artery insufficiency coronary artery restenosis coronary artery stenosis coronary endarterectomy coronary no-reflow phenomenon coronary ostial stenosis coronary revascularisation dissecting coronary artery aneurysm ECG signs of myocardial ischaemia external counterpulsation haemorrhage coronary artery in-stent coronary artery restenosis ischaemic cardiomyopathy microvascular angina myocardial ischaemia percutaneous coronary intervention Prinzmetal angina stress cardiomyopathy subclavian coronary steal syndrome subendocardial ischaemia

Specific TEAE cluster	Preferred terms
Cardiac Arrhythmias	Arrhythmia related investigations, signs and symptoms (Standardized MedDRA Query [SMQ]) Chronotropic incompetence Electrocardiogram repolarisation abnormality Electrocardiogram RR interval prolonged Electrocardiogram U-wave abnormality Electrocardiogram U-wave biphasic Gallop rhythm present Sudden cardiac death Bradycardia Cardiac arrest Cardiac death Cardiac telemetry abnormal Cardio-respiratory arrest Electrocardiogram abnormal Electrocardiogram ambulatory abnormal Electrocardiogram change Heart rate abnormal Heart rate decreased Heart rate increased Loss of consciousness Palpitations Rebound tachycardia Sudden death Syncope Tachycardia Tachycardia paroxysmal Bradyarrhythmias (including conduction defects and disorders of sinus node function) (SMQ) Accessory cardiac pathway Adams-Stokes syndrome Agonal rhythm Atrial conduction time prolongation Atrioventricular block Atrioventricular block complete Atrioventricular block first degree Atrioventricular block second degree Atrioventricular conduction time shortened Atrioventricular dissociation Bifascicular block Bradyarrhythmia Brugada syndrome Bundle branch block Bundle branch block bilateral Bundle branch block left Bundle branch block right Conduction disorder

Specific TEAE cluster	Preferred terms
	Electrocardiogram delta waves abnormal Electrocardiogram PQ interval prolonged Electrocardiogram PR prolongation Electrocardiogram PR shortened Electrocardiogram QRS complex prolonged Electrocardiogram QT prolonged Electrocardiogram repolarisation abnormality Lenegre's disease Long QT syndrome Nodal arrhythmia Nodal rhythm Sick sinus syndrome Sinoatrial block Sinus arrest Sinus arrhythmia Sinus bradycardia Trifascicular block Ventricular asystole Ventricular dyssynchrony Wandering pacemaker Wolff-Parkinson-White syndrome <u>Cardiac arrhythmia terms, nonspecific (SMQ)</u> Arrhythmia Heart alternation Heart rate irregular Pacemaker generated arrhythmia Pacemaker syndrome Paroxysmal arrhythmia Pulseless electrical activity Reperfusion arrhythmia Withdrawal arrhythmia <u>Supraventricular tachyarrhythmias (SMQ)</u> Arrhythmia supraventricular Atrial fibrillation Atrial flutter Atrial parasystole Atrial tachycardia Supraventricular extrasystoles Supraventricular tachyarrhythmia ECG P wave inverted Electrocardiogram P wave abnormal Retrograde p-waves Sinus tachycardia Supraventricular tachycardia

Specific TEAE cluster	Preferred terms
	Tachyarrhythmia terms, nonspecific (SMQ) Anomalous atrioventricular excitation Atrioventricular extrasystoles Cardiac flutter Extrasystoles Tachyarrhythmia Ventricular tachyarrhythmias (SMQ) Accelerated idioventricular rhythm Cardiac fibrillation Parasystole Rhythm idioventricular Torsade de pointes Ventricular arrhythmia Ventricular extrasystoles Ventricular fibrillation Ventricular flutter Ventricular parasystole Ventricular pre-excitation Ventricular tachyarrhythmia Ventricular tachycardia
Amyloid-related imaging abnormality – hemorrhage (ARIA-H, also known as microhemorrhage)	Amyloid related imaging abnormality-microhaemorrhages and haemosiderin deposits Brain stem microhaemorrhage Cerebral microhemorrhage Cerebellar microhemorrage Cerebral haemorrhage Superficial siderosis of central nervous system Cerebral haemosiderin deposition

Abbreviations: ECG = electrocardiogram; TEAE = treatment-emergent adverse event.

Appendix 3. List of Clinical Laboratory Parameters

Blood Chemistry	Hematology	Urinalysis
Albumin	Basophils	Blood
Alkaline phosphatase	Eosinophils	Color
ALT/SGPT	Erythrocyte count	Glucose
AST/SGOT	Hematocrit	Ketones
Bilirubin, Direct	Hemoglobin	Leukocyte esterase
Bilirubin, Total	Leukocyte Count	pH
C-reactive protein	Lymphocytes	Protein
Calcium	Mean Cell Hemoglobin	Specific Gravity
Chloride	Mean Cell Volume	
Cholesterol	Monocytes	
Creatine Phosphokinase	Neutrophils, segmented	
Creatinine	Platelet Count	
Gamma Glutamyltransferase		
Glucose		
Phosphorus		
Potassium		
Sodium		
Total protein		
Urea Nitrogen		
Uric acid		

Abbreviations: ALT = alanine aminotransferase; AST = aspartate aminotransferase; SGOT = serum glutamic oxalacetic transaminase; SGPT = serum glutamic-pyruvic transaminase.

Appendix 4. List of Preferred Terms for Clinical Relevance of Treatment-Emergent Immunogenicity

Preferred Terms	
Allergic oedema	Myalgia
Anaphylactic reaction	Nephritis
Anaphylactic shock	Oedema mouth
Anaphylactic transfusion reaction	Oropharyngeal swelling
Anaphylactoid reaction	Palatal oedema
Anaphylactoid shock	Pharyngeal oedema
Angioedema	Pruritus
Arthralgia*	Pruritus allergic
Bronchospasm	Pruritus generalised
Circumoral oedema	Pyrexia
Drug hypersensitivity	Rash
Epiglottic oedema	Rash erythematous
Eyelid oedema	Rash generalised
Face oedema	Rash pruritic
Gingival oedema	Renal failure
Gingival swelling	Renal failure acute
Gleich's syndrome	Serum sickness
Haematuria	Swollen tongue
Hypotension	Tongue oedema
Kounis syndrome	Type I hypersensitivity
Laryngeal oedema	Urticaria
Lip oedema	Urticaria papular
Lip swelling	

Note:* There is no MedDRA term for "arthralgia with fever". Participants must have both arthralgia **and** pyrexia to meet criteria

Signature Page for VV-CLIN-074356 v1.0

Approval	PPD Statistician 18-Jan-2023 13:23:32 GMT+0000

Signature Page for VV-CLIN-074356 v1.0