STATISTICAL ANALYSIS PLAN (SAP) FOR PART 1

Protocol Number:	INFRONT-3 (AL001-3)
Protocol Title:	A Phase 3, Multicenter, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of AL001 in Individuals at Risk For or With Frontotemporal Dementia Due to Heterozygous Mutations in the Progranulin Gene
Protocol Version and Date:	Version 6.6, 30 Jul 2025
Compound Name:	Latozinemab (AL001)
Indication:	Frontotemporal dementia (FTD)
Dosage Form/Strength:	IV 60 mg/kg
Study Phase:	3
Regulatory Agency Identifier(s):	IND Number: [redacted] EU Clinical Trial Number: 2023-506873-36

Registry Identifier(s):	Registry Name	Registry ID
	clinicaltrials.gov	NCT04374136

Sponsor: Alector Inc.
131 Oyster Point Boulevard, Suite 600
South San Francisco, CA 94080, US
Telephone: +1 650-837-0384

SAP Version and Date: Version 2.1 – United States, 31 Jul 2025

CONFIDENTIAL

This document contains confidential information. This document must not be disclosed to anyone other than the site study staff and members of the institutional review board, independent ethics committee, institutional scientific review board, or equivalent. The information in this document cannot be used for any purpose other than the evaluation or conduct of the clinical investigation without the prior written consent of Alector, Inc.

SAP APPROVAL

A Phase 3, Multicenter, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of AL001 in Individuals at Risk For or With Frontotemporal Dementia Due to Heterozygous Mutations in the Progranulin Gene

I approve the Statistical Analysis Plan for INFRONT-3 Part 1, Version 2.1 - United States, 31 Jul 2025.

Authored by: Leonard Guizzetti, PhD [redacted]

Refer to the signature manifest at the end of the document.

Signature | Date (dd-Mmm-yyyy)

Approved by: Arthur Mayorga, PhD [redacted]

Refer to the signature manifest at the end of the document.

Signature | Date (dd-Mmm-yyyy)

Approved by: Jinglin Zhong, PhD [redacted]

Refer to the signature manifest at the end of the document.

Signature | Date (dd-Mmm-yyyy)

SAP VERSION HISTORY

SAP Version	Approval Date	Based on Protocol (Version, Approval Date)	Brief Description of Change(s) and Rationale
2.1	31 Jul 2025	Version 6.6, 30 Jul 2025	• Plasma PGRN concentration was moved into a co-primary endpoint alongside the clinical primary endpoint following a recommendation from the FDA during the review of the SAP. Therefore, related changes to accommodate the co-primary endpoints are made to the statistical hypotheses, multiplicity adjustment and analysis set definitions. Additional analyses were also included as a result of the change to include co-primary endpoints. Please refer to Section 6 for complete details. • A *post hoc* power calculation was added to ensure that there was sufficient power for this study to support the recommended change of two co-primary endpoints. • A qualitative check that the proportionality assumption for the pMMRM is reasonable and fits the data approximately well, the estimated mean trajectories of both treatment groups from both the pMMRM and the MMRM model was added to the SAP. • Details for the time savings analysis based on the MMRM results was updated based on feedback from the FDA. • Minor editorial changes were made throughout to add detail or improve readability of the document.
1.0	03 Mar 2025	Version 6.0, 07 Sep 2023	Original document creation

TABLE OF CONTENTS

SAP APPROVAL 2
SAP VERSION HISTORY 3
LIST OF ABBREVIATIONS 8
1. INTRODUCTION 11
1.1. Study Objectives and Endpoints 11
1.2. Study Design 13
1.2.1. Part 1: Double-blind, Placebo-controlled Treatment Period 14
1.2.2. Part 2: Optional Open-label Extension 14
1.2.3. Sample Size Determination 14
2. STATISTICAL HYPOTHESES, MULTIPLICITY ADJUSTMENT AND ANALYSIS SETS 16
2.1. Multiplicity Adjustment 16
2.2. Additional Testing of Biomarkers 17
2.3. Definition of Analysis Set 18
3. STATISTICAL ANALYSIS 18
3.1. General Considerations 18
3.1.1. Definitions, Derivations and Classifications 19
3.1.2. Windowing, Unscheduled Visits and Repeat Test Results 19
3.1.3. Handling of Values Above or Below Limits of Quantification 20
3.1.4. Handling of Missing Data or Dropouts 20
3.1.4.1. Handling of Missing Data in Concomitant Medication and Adverse Event 20
3.2. Participant Disposition and Protocol Deviations 20
3.2.1. Participant Disposition 20
3.2.2. Protocol Deviations 21
3.3. Participant Characteristics 21
3.3.1. Demographic and Other Baseline Characteristics 21
3.3.2. Prior Medical and Surgical History 21
3.3.3. Concomitant Medical and Surgical Treatment Procedures 21
3.3.4. Prior and Concomitant Medications/Treatments 22
3.3.5. Extent of Exposure to Study Drug and Duration of Study 22
3.3.5.1. Extent of Exposure to Study Drug 23
3.3.5.2. Study Duration 23

3.3.5.3. Dosing Intensity 23
3.3.5.4. Treatment Compliance 24
3.4. Efficacy Analyses 24
3.4.1. Adjustment for Covariates 24
3.4.2. Clinical Co-primary Endpoint: Clinical Dementia Rating Dementia Staging Instrument Plus National Alzheimer's Disease Coordinating Center Frontotemporal Lobar Degeneration Behavior and Language Domains - Sum of Boxes 24
3.4.2.1. Definition of Estimand 25
3.4.2.2. Definition of Endpoint 25
3.4.2.3. Main Analytical Approach 26
3.4.2.4. Sensitivity Analyses for Missing Data 27
3.4.2.5. Supportive Analysis: Mixed Models for Repeated Measures 28
3.4.2.6. Exploratory Analyses 29
3.4.3. Biomarker Co-primary Endpoint: Plasma Progranulin 30
3.4.3.1. Definition of Estimand 30
3.4.3.2. Definition of Endpoint 30
3.4.3.3. Main Analytical Approach 30
3.4.3.4. Sensitivity Analysis for Missing Data 31
3.4.4. Key Secondary Endpoint: Clinical Global Impression-Severity 31
3.4.5. Additional Secondary Endpoints 31
3.4.5.1. Frontotemporal Dementia Rating Scale Score 31
3.4.5.2. Clinical Global Impression-improvement 32
3.4.5.3. Repeatable Battery for the Assessment of Neuropsychological Status 32
3.5. Exploratory Analyses 33
3.5.1. Time-to-event Analysis 33
3.5.1.1. Time-to-symptomatic Worsening 33
3.5.1.2. Time to Study Treatment Discontinuation 34
3.5.2. "Responder" Type of Analysis 34
3.5.3. Quality of Life: EuroQOL-5 Dimensions 34
3.5.4. Quality of Life: Zarit Burden Interview 35
3.5.5. Pharmacoeconomic: Resource Utilization in Dementia-Lite Version 35
3.5.6. Fluid Biomarker Analyses 36

3.5.7. Magnetic Resonance Imaging Biomarker Analyses 36
3.5.7.1. Treatment Effect on Magnetic Resonance Imaging Biomarkers by MMRM 37
3.5.8. Winterlight Labs Speech Analysis (Optional Assessment) 38
3.5.9. Association Analyses Between Biomarkers and Clinical Co-primary Endpoint 38
3.5.10. Causal Mediation Analyses of Selected Biomarkers 38
3.6. PK and PK/PD Analyses 38
3.7. Subgroup Analyses of Interest 39
3.7.1. Subgroup Analyses for Clinical Efficacy Endpoints 39
3.7.2. Subgroup Analyses for Fluid Biomarkers 40
3.8. Analysis Considerations for At-risk Participants 40
3.9. Safety Analyses 40
3.9.1. Adverse Events 41
3.9.1.1. Treatment-emergent Adverse Events 41
3.9.1.2. Serious Adverse Events 41
3.9.1.3. Relationship of Adverse Events to Study Drug 41
3.9.1.4. Summaries of Adverse Events 41
3.9.2. Sheehan-Suicidality Tracking Scale 42
3.9.3. Laboratory Data 43
3.9.4. Vital Signs 43
3.9.5. Electrocardiogram Parameters 43
3.9.6. Physical and Neurological Examination 44
3.9.7. Magnetic Resonance Imaging 44
3.9.8. Other Safety Assessments 44
3.9.8.1. Pregnancy Test 44
3.9.8.2. Impacts Due to COVID-19 44
3.10. Other Analyses 44
3.10.1. Immunogenicity Analyses 44
4. INTERIM ANALYSIS 45
5. CHANGES FROM AND ADDITIONS TO METHODS PLANNED IN THE PROTOCOL 45
5.1. Secondary Endpoints 46
5.2. Hypothesis Testing Strategy 46

5.3. Covariate Adjustment for Co-primary Clinical Outcome and Secondary Clinical Outcome Assessment Endpoints 46
6. CHANGES FROM OR ADDITIONS TO PREVIOUS VERSIONS OF THIS SAP 47
7. REFERENCES 49
8. SUPPORTING DOCUMENTATION 51
APPENDIX 1. MODEL DETAILS FOR THE PROPORTIONAL MIXED MODEL FOR REPEATED MEASURES 51
APPENDIX 2. DETAILS OF TIME-SAVINGS OF TREATMENT EFFECT ANALYSIS 53
APPENDIX 3. CAUSAL MEDIATION ANALYSIS 54
APPENDIX 4. MAPPING S-STS TO C-CASA 2010 CATEGORIES 56
APPENDIX 5. POTENTIALLY CLINICALLY SIGNIFICANT ABNORMALITIES 57
APPENDIX 6. SUMMARY OF PLANNED CLINICAL AND BIOMARKER ANALYSES 58

LIST OF TABLES

Table 1: Objectives and Endpoints 11
Table 2: Analysis Sets 18
Table 3: Summary of Treatment-emergent Adverse Event Tables by Analysis Set 42
Table 4: Mapping from the Sheehan-STS to the FDA 2010 C-CASA Categories 56
Table 5: Definitions of Potentially Clinically Significant Criteria for Vital Signs and Laboratory Values 57
Table 6: Definitions of Potentially Clinically Significant Criteria for Electrocardiogram Data 57
Table 7: Summary Table of Planned Clinical Efficacy Analyses 58
Table 8: Summary Table of Planned Fluid and Imaging Biomarker Analyses 59

LIST OF FIGURES

Figure 1: Diagram of Sequential Testing Strategy 17
Figure 2: Causal Mediation Directed Acyclic Diagram 54

LIST OF ABBREVIATIONS

Abbreviation	Description
ADA	antidrug antibody
AE	adverse event
ALLFTD	FTD research consortium consisting of Advancing Research and Treatment in Frontotemporal Lobar Degeneration (ARTFL) and Longitudinal Evaluation of Familial Frontotemporal Dementia Subjects (LEFFTDS)
ARS	all randomized set
ATC	Anatomical therapeutic chemical
BMI	body mass index
BPP	backwards projection to placebo
bvFTD	Behavioral variant frontotemporal dementia
C-CASA	Columbia Classification Algorithm of Suicide Assessment
CDR plus NACC FTLD	Clinical Dementia Rating Dementia Staging Instrument plus National Alzheimer's Disease Coordinating Center frontotemporal lobar degeneration Behavior and Language domains
CDR plus NACC FTLD-GS	Clinical Dementia Rating Dementia Staging Instrument plus National Alzheimer's Disease Coordinating Center frontotemporal lobar degeneration Behavior and Language domains - global score
CDR plus NACC FTLD-SB	Clinical Dementia Rating Dementia Staging Instrument plus National Alzheimer's Disease Coordinating Center frontotemporal lobar degeneration Behavior and Language domains - sum of boxes
CFB	change from baseline
CGI-I	clinical global impression - improvement
CGI-S	clinical global impression - severity
CI	confidence interval
COA	clinical outcome assessment
COVID-19	coronavirus disease 2019
CSF	cerebrospinal fluid
CSR	clinical study report
CV	Coefficient of variation
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
EMA	European Medicines Agency
ET	early termination

Abbreviation	Description
EQ-5D	EuroQOL-5 Dimensions, 5 levels
EuroQol	European Quality of Life
FAS	Full analysis set
FDA	Food and Drug Administration
FDR	false-discovery rate
FRS	Frontotemporal dementia rating scale
FTD	Frontotemporal dementia
FTLD	Frontotemporal lobar dementia
GENFI	Genetic FTD Initiative
GFAP	Glial fibrillary acidic protein
GRN	Progranulin gene
iDMC	independent data monitoring committee
IRT	interactive response technology
IV	intravenous(ly)
LLOQ	lower limit of quantitation
LP	lumbar puncture
LSMean(s)	least-squares mean(s)
MAR	Missing-at-random
Max	Maximum
MedDRA	medical dictionary for regulatory activities
MI	multiple imputation
Min	Minimum
ML	maximum likelihood
MCMC	Markov chain Monte Carlo
MMRM	mixed model for repeated measures
MNAR	Missing-not-at-random
MRI	magnetic resonance imaging
Msec	millisecond
NfL	neurofilament light chain
OLE	open-label extension
PCFB	percentage change from baseline
PD	Pharmacodynamic

Abbreviation	Description
PGRN	Progranulin protein
PK	Pharmacokinetic
PMM	pattern-mixture model
pMMRM	proportional MMRM
PPA	primary progressive aphasia
PT	preferred term
q4w	every 4 weeks
QTcF	QT interval corrected using the Fridericia formula ($QTcF = QT/(RR)^{1/3}$)
RBANS	repeatable battery for the assessment of neuropsychological status
RBPI	Reference-based pattern imputation
REML	Restricted maximum likelihood
ROI	region of interest
RUD-Lite	Resource Utilization in Dementia-Lite version
SAE	serious adverse event
SAF	Safety analysis set
SAP	statistical analysis plan
SAS®	Statistical Analysis Software, SAS Institute
SD	standard deviation
SE	standard error
SFV	safety follow-up visit
SOC	system organ class
Sheehan-STS	Sheehan-suicidality tracking scale
TEAE	treatment-emergent adverse event
TFL	table, figure, legend
ULOQ	upper limit of quantitation
VAS	visual analogue scale
WHO DD	World Health Organization Drug Dictionary
WLSA	Winterlight Labs speech assessment
ZBI	Zarit burden interview

1. INTRODUCTION

Study INFRONT-3 is a Phase 3 study designed to evaluate the efficacy and safety of AL001 (latozinemab) compared to placebo in individuals with frontotemporal dementia (FTD) due to a heterozygous granulin gene (*GRN*) mutation. The study is conducted in 2 parts. Part 1 is a 96-week randomized, double-blind placebo-controlled treatment period. Part 2 is an optional 96-week open-label extension (OLE) for eligible participants who have completed Part 1. All participants in the OLE, whether they received latozinemab or placebo in Part 1, will receive open-label latozinemab.

There are 2 separate statistical analysis plans (SAP) for Part 1 and Part 2 data, respectively. The clinical data cutoff date for the Part 1 SAP is when all participants have either completed Visit 25/Week 96 or have prematurely discontinued the planned study treatment period of Part 1 of this study.

At the time of database lock, all efficacy and safety data and PK/Biomarker data collected in Part 1 (including the Part 1 safety follow-up visit) are described and analyzed according to this Part 1 SAP, which constitutes the final clinical study report analysis for Part 1 of this study. Analysis of Part 2 data will be described and analyzed according to a separate Part 2 SAP. All data and analyses included in these 2 SAPs will be submitted for health authority review in the initial marketing application submissions.

The purpose of this SAP is to provide general guidelines for the analysis of data collected from Part 1 of the INFRONT-3 study. The SAP contains detailed technical elaborations of the principal features of the analysis described in the protocol and includes detailed procedures for executing the statistical analyses. Deviations from these guidelines must be substantiated by a sound statistical rationale and will be documented in the clinical study report (CSR). If there are differences between the protocol and the SAP, this SAP will take precedence. Detailed table, figure and listing (TFL) specifications will be provided in a separate document, and additional technical details will be outlined in a technical specifications document, and both will be finalized prior to execution of the analyses outlined in this SAP.

1.1. Study Objectives and Endpoints

Table 1: Objectives and Endpoints

Objectives	Endpoints
Co-primary	
To evaluate the efficacy of latozinemab compared with placebo in symptomatic participants as measured by the CDR plus NACC FTLD-SB	Change from baseline to Weeks 48, 72 and 96 in the CDR plus NACC FTLD-SB
To evaluate the treatment effect of latozinemab compared with placebo in symptomatic participants as measured by plasma PGRN concentration	Change from baseline of log-10 transformed plasma PGRN concentration at Week 96

Table 1: Objectives and Endpoints (Continued)

Objectives	Endpoints
Key Secondary	
To evaluate the efficacy of latozinemab compared with placebo in symptomatic participants as measured by the CGI-S	Change from baseline to Weeks 48, 72 and 96 in the CGI-S
Additional Secondary	
To evaluate the efficacy of latozinemab compared with placebo in symptomatic participants as measured by the FRS	Change from baseline to Weeks 48, 72 and 96 in the FRS
To evaluate the efficacy of latozinemab compared with placebo in symptomatic participants as measured by the CGI-I	Observed values at Weeks 48, 72 and 96 in the CGI-I
To evaluate the efficacy of latozinemab compared with placebo in symptomatic participants as measured by the RBANS	Change from baseline to Weeks 48, 72 and 96 in the RBANS
Safety	
To evaluate the safety and tolerability of latozinemab compared with placebo as measured by safety assessments and antidrug antibodies (ADAs) in the symptomatic population, the at-risk population, and all FTD participants (both symptomatic and at-risk).	Incidence and severity of treatment-emergent adverse events (AEs) and serious adverse events (SAEs) Physical examination abnormalities Neurological examination abnormalities Change in vital signs from baseline over time Changes in ECGs from baseline over time MRI abnormalities Changes in clinical laboratory tests from baseline over time Sheehan-suicidality Tracking Scale (Sheehan-STS) Incidence of ADAs to latozinemab
Exploratory Endpoints	
To evaluate the effects of latozinemab compared with placebo as measured by:	
Exploratory Quality of Life objective: Quality of Life COAs	*Exploratory Quality of Life endpoints:* Change from baseline to Weeks 48, 72 and 96 in EQ-5D visual analogue scale (VAS) score Change from baseline to Weeks 48, 72 and 96 in ZBI score
Exploratory Pharmacoeconomic objective: Pharmacoeconomic COA	*Exploratory Pharmacoeconomic endpoint:* Resource Utilization in Dementia-Lite (RUD-Lite) score

Table 1: Objectives and Endpoints (Continued)

Objectives	Endpoints
Exploratory PD and Disease Biomarker objectives: Exposure-response relationships Exploratory imaging Exploratory biomarkers	*Exploratory PD and Disease Biomarker endpoints:* Exposure-response relationships for relevant biomarkers (eg, progranulin [PGRN], neurofilament light chain [NfL]), efficacy responses in relation to latozinemab exposure Arterial spin labeling, diffusion-tensor imaging, and structural volumetric regional MRI measures PGRN concentrations in optional CSF NfL concentrations in serum and optional CSF Glial fibrillary acidic protein (GFAP) concentrations in plasma and optional CSF
Exploratory COA objective: Winterlight Lab Speech Assessment (WLSA) Time-savings of treatment effect Win probability of treatment effect	*Exploratory COA endpoints:* WLSA, for participants who agree to participate in these optional assessments only Time to decline to the value of the CDR plus NACC FTLD-SB at Week 96 in the latozinemab group compared to the same decline in the placebo group projected backwards to an earlier timepoint A ranked analysis based on a composite of the Week 96 change from baseline in the CDR plus NACC FTLD-SB or duration of study participation or time to death
Time-to-symptomatic worsening:	As measured by: CDR plus NACC FTLD-GS CDR plus NACC FTLD-SB
Responder-type analyses:	As measured by: CDR plus NACC FTLD-GS

CDR plus NACC FTLD=Clinical Dementia Rating Dementia Staging Instrument plus National Alzheimer's Disease Coordinating Center frontotemporal lobar degeneration Behavior and Language domains; CDR plus NACC FTLD-GS=CDR plus NACC FTLD-Global score, CDR plus NACC FTLD-SB=CDR plus NACC FTLD-Sum of boxes; CGI-I=clinical global impression-improvement; CGI-S=clinical global impression-severity; FRS=frontotemporal dementia rating scale; RBANS=repeatable battery for the assessment of neuropsychological status.

1.2. Study Design

The study has 2 parts: A double-blind, placebo-controlled treatment period (Part 1), followed by an optional OLE period (Part 2).

Each participant must be a known carrier of a heterozygous loss-of-function *GRN* mutation causative of FTD with a global CDR plus NACC FTLD score of 0 to 2, and:

- Participants with a CDR plus NACC FTLD-SB score ≤0.5 must have an elevated serum NfL (at-risk participants), or
- Participants with a CDR plus NACC FTLD-SB score >0.5 must have 1 or more of the 6 behavioral/cognitive symptoms required for a diagnosis of possible bvFTD, or a diagnosis of PPA (symptomatic participants).

1.2.1. Part 1: Double-blind, Placebo-controlled Treatment Period

Part 1 is a 96-week randomized, double-blind, placebo-controlled treatment period to evaluate the efficacy and safety of latozinemab 60 mg/kg, q4w compared to placebo in participants who are known carriers of a heterozygous loss-of-function *GRN* mutation causative of FTD. Participants are randomized to receive latozinemab or placebo in a 3:2 ratio, respectively.

Participants were stratified by:

- CDR plus NACC FTLD-SB score ≤0.5
- CDR plus NACC FTLD-Global score 0.5 (with a CDR plus NACC FTLD-SB >0.5)
- CDR plus NACC FTLD Global score 1
- CDR plus NACC FTLD Global score 2

A Schedule of Assessments can be referred to in the Protocol.

1.2.2. Part 2: Optional Open-label Extension

If a Part 1 participant completes 96 weeks of treatment, they are eligible to be rolled over into the optional Part 2 OLE portion of the study. Part 1 participants may be excluded from continuing in the Part 2 OLE if in the opinion of the investigator, continued treatment with latozinemab at the conclusion of Part 1 is not beneficial or safe for the participant. Part 1 participants who have been admitted to a nursing home/assisted living facility are not eligible for participation in the OLE. An individual participant may continue to receive latozinemab (60 mg/kg) in the OLE portion of the study for up to 25 doses over a 96-week dosing period.

Treatment assignment in Part 1 (latozinemab or placebo) will remain blinded to investigators, participants and raters until Part 2 is completed and the Part 2 data for all participants have been cleaned and locked.

1.2.3. Sample Size Determination

The study was initially designed to enroll both symptomatic and at-risk participants for the primary analysis. Data from the GENFI registry of *GRN* mutations with multi-year follow-up on the CDR plus NACC FTLD-SB was analyzed to obtain an estimate of the expected change over time in the clinical primary endpoint CDR plus NACC FLTD-SB, and its associated variability, in the at-risk and symptomatic participants for the purpose of sample size estimation. This analysis showed the 2-year coefficient of variation (CV) for the change in the CDR plus NACC FLTD-SB was 77% for the symptomatic cohort and 165% for the at-risk cohort. Based on these results, we assumed a 3:2 randomization ratio of participants treated with latozinemab:placebo, with a distribution of ~35% at-risk participants and ~65% symptomatic participants, and assuming ~8% dropout in the at-risk participants (2% per visit) and ~20% dropout in the symptomatic participants (5% per visit). There was approximately 85% power under a two-sided type 1 error of 5% to detect a 40% slowing in disease progression across Weeks 48, 72 and 96, in the treatment group relative to the placebo group, with an initial sample size estimate of N=180 combined at-risk and symptomatic participants. Power and sample size estimation were conducted by simulation using the pMMRM model.

Given the uncertainty in the variability of the clinical primary endpoint, a single, blinded sample size re-estimation was conducted. Sample size re-estimation was performed by simulation and focused on symptomatic participants. The sample size re-estimation procedure was conducted by pooling both treatment groups, treating the sample as an internal pilot study to provide a more accurate estimate of the variability of the clinical primary endpoint over time, which was treated as a nuisance parameter (Friede 2013). This re-estimation process has been shown to control the Type I error rate, with little effect on the overall power to detect a treatment effect if the new sample size is at least close to the originally planned sample size.

Updated variability estimates were informed by the observed 2-year variability (CV=68%) from the GENFI and ALLFTD registries as reported by Staffaroni (2022). A slightly smaller variability (65%) has also been assumed due to routine blinded monitoring of the accumulating INFRONT-3 data. The assumed 48- and 96-week dropout rates were 30% and 40%, respectively, while the allocation ratio and treatment effect estimate remain the same as the initial design. Based on these assumptions, a sample size of N=90 is required to achieve at least 85% power under a two-sided type 1 error of 5% to detect a 40% reduction in disease progression in the latozinemab group compared to placebo in the primary analysis population of symptomatic participants.

Following a recommendation from the FDA during the review of the AL001-3 SAP, the study objective was changed to include plasma PGRN as a biomarker co-primary endpoint. As a result of this change, a *post hoc* power calculation was performed to ensure that there was still sufficient power for this study to support the recommended change of two co-primary endpoints. Based on available data from AL001-2 study for symptomatic FTD-*GRN* participants treated with latozinemab, it is expected that after 96 weeks of treatment, participants would experience a mean (SD) change from baseline of log-10 transformed plasma PGRN concentration of 0.45 (0.0815). Based on longitudinal data from a large observational cohort of FTD patients (Sellami 2020), plasma PGRN levels are low and stable up to 10 years after disease onset. As a result, plasma PGRN levels are not expected to change over 96 weeks while on placebo. Therefore, the change from baseline of log-10 transformed plasma PGRN concentration for the placebo arm is assumed to be 0 with a common SD of 0.0815 to match the latozinemab treated individuals. Anticipating that there may be as many as 63 participants with a baseline and Week 96 plasma PGRN concentration, in a 2:3 allocation ratio of placebo:latozinemab, a two-sample t-test shows that there is >99% power at a 1-sided 2.5% significant level to detect the treatment difference. Therefore, there is adequate power in this study to detect a treatment effect on both the clinical co-primary endpoint and the biomarker co-primary endpoint.

Based on the small number of at-risk participants who have enrolled in the trial and external registry data showing more than twice the variability in CDR plus NACC FTLD-SB in at-risk participants compared to symptomatic participants (165% vs. 77%), discussions with the FDA and the Committee for Medicinal Products for Human Use led to an agreement to change the primary analysis population to focus on symptomatic participants.

2. STATISTICAL HYPOTHESES, MULTIPLICITY ADJUSTMENT AND ANALYSIS SETS

The primary analysis for the clinical co-primary and all secondary endpoints will use a proportional mixed model for repeated measures (pMMRM), in which the percent slowing of disease progression due to study treatment is modeled as a single proportional (θ) reduction across Weeks 48 to 96.

The primary analysis for the biomarker co-primary (plasma PGRN) and all other exploratory biomarker endpoints will use a mixed model for repeated measures (MMRM) to compare the difference between treatment and placebo group in change from baseline (Δ) of the log-10 transformed values at Week 96.

For clinical endpoints, hypotheses will be tested using a 1-sided 2.5% significance level for the primary estimand as follows:

- H_0 (null): $\theta = 0$, indicating that latozinemab group disease progression rate is the same as that of the placebo group.
- H_1 (alternative): $\theta > 0$, indicating that latozinemab is slowing the disease progression relative to placebo.

The hypothesis to be tested for the biomarker co-primary and exploratory fluid biomarker endpoints (modeling the log-10 transformed values) will be tested using a 1-sided 2.5% significance level as follows:

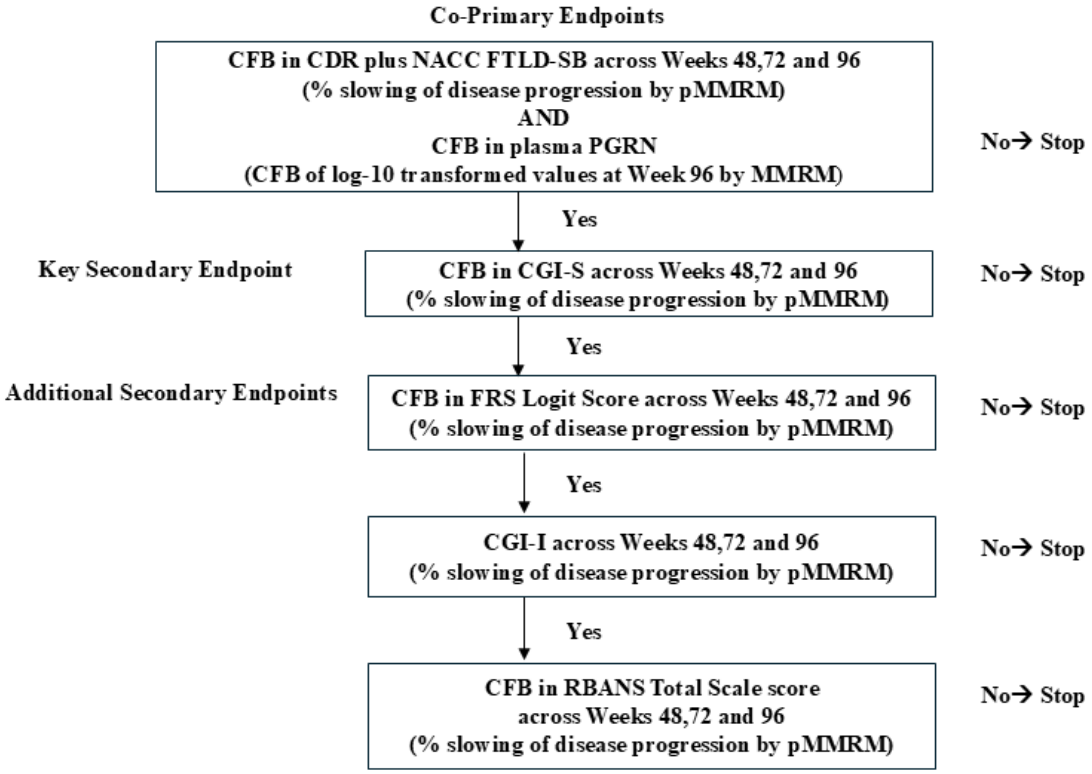
- H_0 (null): $\Delta = 0$, indicating that the change from baseline at Week 96 in treatment group is the same as that of the placebo group.
- H_1 (alternative): $\Delta > 0$ for plasma PGRN, or $\Delta < 0$ for serum NfL and plasma GFAP, indicating a treatment benefit favoring the latozinemab group.

2.1. Multiplicity Adjustment

A sequential hierarchical testing strategy will be applied, with the testing order specified below. All endpoints will be tested at a 1-sided 0.025 alpha level. Each of the co-primary endpoints will be tested first at the 0.025 level. If the primary hypothesis for both primary endpoints is rejected at the 0.025 level, the remaining endpoints will be tested in sequential order as presented below in Figure 1. During the testing, if any endpoint fails to be rejected at the 0.025 level, the testing procedure will stop, and for all subsequent endpoints only nominal p-values will be reported.

The sequential testing strategy is outlined in Figure 1.

Figure 1: Diagram of Sequential Testing Strategy



CDR plus NACC FTLD=Clinical Dementia Rating Dementia Staging Instrument plus National Alzheimer's Disease Coordinating Center frontotemporal lobar degeneration Behavior and Language domains; CDR plus NACC FTLD-GS=CDR plus NACC FTLD-Global score; CDR plus NACC FTLD-SB=CDR plus NACC FTLD-Sum of Boxes; CFB=change from baseline; CGI-I=clinical global impression - improvement; CGI-S=clinical global impression - severity; FRS=frontotemporal dementia rating scale; MMRM=mixed model for repeated measures; PGRN=progranulin; pMMRM=proportional mixed model for repeated measures; RBANS=repeatable battery for the assessment of neuropsychological status.

2.2. Additional Testing of Biomarkers

The biomarker endpoints (serum NfL and plasma GFAP) will be analyzed separately by MMRM, examining the treatment effect of AL001 on changes in these biomarkers. Nominal p-values will be reported for treatment group LSMean differences (vs. placebo) to provide supportive evidence to the treatment effects of latozinemab.

The volumetric MRI endpoints (from the following regions of interest: frontal, temporal, parietal; and bilateral ventricular volumes) will be analyzed separately by MMRM, examining the treatment effect of AL001 on changes in these regions of interest (ROI) volumes. Nominal p-values will be reported for treatment group LSMean differences (vs. placebo) to provide supportive evidence to the treatment effects of latozinemab.

All other p-values will be reported at nominal levels and no adjustment for multiplicity will be conducted.

2.3. Definition of Analysis Set

The following analysis sets are defined in Table 2.

Table 2: Analysis Sets

Analysis Set	Definition
All enrolled participants (ENR)	All participants who signed an informed consent form.
All randomized set (ARS)	All participants randomized to study treatment arms.
Full analysis set (FAS)	The subset of the ARS who have received any amount of study drug. Participants in this group will be analyzed according to their randomized treatment group.
FAS-symptomatic	The subset of the FAS who are symptomatic at baseline. This set will be used as focal inferential analyses for co-primary, secondary and biomarker endpoints (including causal mediation analysis).
FAS-at Risk	The subset of the FAS who are at-risk (pre-symptomatic) at baseline.
Per Protocol Set (PPS)-symptomatic	The subset of the FAS-symptomatic excluding participants who had important protocol deviation that would affect the interpretation of the efficacy data. Participants in this set will be analyzed according to the randomized treatment group. The inclusion/exclusion of participants in this analysis set will be determined prior to database lock.
Safety analysis set (SAF)	The subset of participants in the ARS who have been exposed to any amount of study drug. Participants in this set will be analyzed according to the actual treatment received, rather than as assigned, during Part 1. If a participant received any amount of latozinemab, they will be included in the latozinemab treatment group, regardless of randomization assignment. This set is the focal set for safety analyses.
SAF-symptomatic	This is the subset of the SAF set which includes participants who were symptomatic at baseline.
SAF-at Risk	This is the subset of the SAF set which includes participants who were at risk (pre-symptomatic) at baseline.
Pharmacokinetic set (PK)	The subset of participants in the SAF who had at least one post-baseline measurable assessment for determination of PK concentration.

3. STATISTICAL ANALYSIS

3.1. General Considerations

Unless specified otherwise, TFLs will be presented by treatment group, visit and timepoint (ie, pre-dose and post-dose), as applicable.

All assessments will be provided in respective by-participant data listings, except data collected only for confirmation of study entry criteria and for study administrative purposes.

Specifications of each table, listing and figure will be documented in a separate programming document, which may include technical definitions or sample code snippets.

All analyses will be conducted using SAS Version 9.4 or later. If necessary, additional analysis conducted in other software will be documented in the CSR.

3.1.1. Definitions, Derivations and Classifications

Definition of Day 1

Study Day 1 is defined as the first day of treatment. This will also be considered the reference start date.

Definition of Baseline

The baseline value will be the last available non-missing value prior to treatment on Day 1 (Visit 1/Week 0). Unless specified otherwise, assessments taken on the same day as trial treatment will be assumed to be before trial treatment (pre-dose) unless indicated to be obtained after trial treatment (post-dose).

The baseline value for any measures planned to be taken with replicates at the same timepoint (eg, triplicate electrocardiogram parameters) will be the average of those replicate measurements using the same definition of baseline above.

Definition of Study Day

Study day will be computed as follows. There is no study day 0 for the purposes of analysis.

- Study day = (Date of event - Date of first dose of study treatment) + 1 if the date of the event is on or after the date of the first study treatment.
- Study day = (Date of event - Date of first dose of study treatment) if the date of the event is prior to the date of the first study treatment.

Derivations Involving Dates

The number of days between 2 dates, the date of event and the reference date, will be calculated as (Date of event - Reference date) + 1 when the event date is on or after the reference date; or, as (Reference date - Date of event) when the event date is before the reference date. The resulting value is then converted to years, months, or weeks, as appropriate, using standard conversion factors. Partial dates will be imputed before calculating the number of days. Details are found in a separate technical document.

3.1.2. Windowing, Unscheduled Visits and Repeat Test Results

The general principle used for efficacy and biomarker analyses is that visits will be allocated into analysis visit windows for purposes of summarization and modeling. These window definitions and analysis value selection rules (ie, in the case of multiple visits within the same analysis window) will be described in a separate document.

3.1.3. Handling of Values Above or Below Limits of Quantification

Unless otherwise specified, clinical laboratory values that are continuous in nature (eg, CRP and safety lab tests) that are collected as "<xx" or ">xx" will be imputed and analyzed as the numerical value of those respective limits without the greater than (>) or less than (<) signs.

For fluid biomarkers, handling of values below the LLOQ and above the ULOQ if reported in the form of "<x" or ">x" (where x is considered the LLOQ or ULOQ) will use the following logic:

- A value that is reduced by the reported precision of the least significant (rightmost) digit of the LLOQ will be used to calculate descriptive statistics if the datum is reported in the form of "<x" (where x is considered the LLOQ). For example, if the values are reported as <50, <5.0 and <5.00, then the precision of the least significant digits are 1, 0.1 and 0.01, and the values of 49, 4.9 and 4.99 respectively, will be used to calculate summary statistics. An exception to this rule is any value reported as <1 or <0.1. For values reported as <1 or <0.1, a value of 0.9 or 0.09, respectively, will be used for calculating summary statistics.
- A value that is 1 unit above the ULOQ will be used to calculate descriptive statistics if the datum is reported in the form of ">x" (where x is considered the ULOQ). Values with decimal points will follow the same logic as above.

However, the LLOQ or ULOQ will be used directly to calculate descriptive statistics if the data is reported in the form of "≤x" or "≥x" (where x is considered the LLOQ or ULOQ).

In general, imputed values will be used for summary statistics and analyses in tables and graphs. For data listings, values will be reported as collected with the relevant signs.

3.1.4. Handling of Missing Data or Dropouts

For efficacy and biomarker analyses, handling of missing data will be specified in each analysis section, as applicable. For safety analyses, missing data will not be imputed, in general (eg, laboratory values).

3.1.4.1. Handling of Missing Data in Concomitant Medication and Adverse Event

The conventions for handling missing or partial dates for the purposes of concomitant medications and adverse events are outlined in a separate technical document. If the relationship of an AE is missing, it will be considered as related to study treatment. If the severity of an AE is missing, it will be considered as severe.

3.2. Participant Disposition and Protocol Deviations

3.2.1. Participant Disposition

Participant disposition will be described for all enrolled participants, including the number of participants enrolled (and primary reason for screen failure), the number of participants re-screened, the number and percentage of participants who completed the study, or discontinued early (and reasons for study discontinuation).

The number and percentage of participants included in the analysis sets as well as reasons for their exclusion, will be summarized by treatment group and overall using the All Randomized Set (ARS).

3.2.2. Protocol Deviations

Protocol deviations will be classified as significant or non-significant protocol deviations. Protocol deviations will be recorded separately from the clinical database. The classification of protocol deviations will be done in accordance with the Protocol Deviation Plan. Deviations leading to exclusion from analysis populations will be finalized prior to database lock, identified and summarized.

The number and percentage of participants with significant protocol deviations for each deviation category will be tabulated by treatment group for the ARS.

3.3. Participant Characteristics

3.3.1. Demographic and Other Baseline Characteristics

Demographic information, including year of birth, age, sex, race, ethnicity and handedness were collected at the Screening visit unless prohibited by local regulatory authorities.

Age will be categorized using the following categories: ≥18 to <65, ≥65 years; and ≥18 to <65, ≥65 to <85, ≥85 years.

Baseline clinical outcome assessments will include the baseline scores on the CDR plus NACC FTLD-GS, CDR plus NACC FTLD-SB, CGI-S, FRS and RBANS. Baseline fluid biomarkers will be summarized for serum and CSF NfL, plasma and CSF GFAP and plasma and CSF PGRN concentrations. Baseline imaging biomarkers will be summarized for each vMRI ROI.

Baseline diagnostic characteristics will include symptomatic status, clinical phenotype, approximate age at diagnosis of first FTD diagnosis and approximate age at onset of symptoms of first diagnosis.

Descriptive summaries of demographic, baseline clinical characteristics and baseline biomarker values by treatment group and overall, will be provided for the FAS and by baseline symptomatic status.

Baseline diagnostic characteristics will be summarized by treatment group and overall, using the FAS-symptomatic.

3.3.2. Prior Medical and Surgical History

Medical and surgical history will be coded using the latest version of the Medical Dictionary for Regulatory Activities (MedDRA, Version 26 or newer).

Medical history will be summarized using numbers and percentages by primary system organ class, preferred term, treatment group, using the FAS.

3.3.3. Concomitant Medical and Surgical Treatment Procedures

Medical and surgical treatment procedures will be coded using MedDRA, Version 26 or newer.

Medical or surgical treatment procedure is any therapeutic and diagnostic intervention or diagnosis assessment performed between the dates of the first dose of study drug and Part 1 End of Study.

The number and percentages of all concomitant medical and surgical treatment procedures will be summarized by treatment group using the FAS.

3.3.4. Prior and Concomitant Medications/Treatments

All prior and concomitant medications will be coded using the World Health Organization Drug Dictionary (WHO DD), Version March 2023 or newer.

A medication will be considered as "prior" if the end date of administration is before the start of the first study treatment administration.

A medication will be considered "concomitant" if the end date of administration is after the start of study treatment administration or if the start date is at or after the start of study treatment administration. If the date and time of administration contains partial information such that the attribution of concomitant administration cannot be ruled out, then it will be considered as “concomitant”. Medications that begin on or after the start of the first dose of Part 2 will not be considered “concomitant” in Part 1. More details on handling of partial dates are specified in a separate technical document. Concomitant medications will be collected at all visits. Medications which are on-going (with a missing end date) are considered concomitant.

All prior and concomitant medications data will be summarized by ATC and preferred drug name. The number and percentage of all concomitant medications will be summarized by ATC level 3 and PT by study treatment group and overall. The total number of concomitant medications and the number and percentages of participants with at least 1 concomitant medication will be summarized.

The number and percentage of all prior medications will be summarized by treatment group using the FAS.

3.3.5. Extent of Exposure to Study Drug and Duration of Study

Study treatment administration data will be summarized for the FAS by duration of exposure to study drug, number of infusions received, number of infusions planned, total cumulative received volume (mL), total cumulative planned infusions, and treatment compliance as defined below.

Infusion interruption, duration of exposure to study treatment and duration of study will be summarized by treatment group and overall. Duration of exposure to study treatment, duration of study follow-up and treatment compliance will be listed.

3.3.5.1. Extent of Exposure to Study Drug

Exposure to study drug will be quantitatively summarized in weeks, by treatment group.

Duration of exposure (days) will be calculated as the minimum of:

- If the participant did not die:
  - the date of last dose of study treatment during Part 1 - date of first dose + 28 days dosing interval
  - the date of first dose in Part 2 - date of first dose, if the participant continued to Part 2
- If the participant did die: date of death - date of first dose + 1 day

In addition, a categorical summary of duration of exposure will be provided using the following intervals:

- <24 weeks
- ≥24 to <48 weeks
- ≥48 to <72 weeks
- ≥72 to <96 weeks
- ≥96 weeks (Part 1 dosing visits only)

3.3.5.2. Study Duration

Duration of study follow-up during Part 1 of the Study will be quantitatively summarized by treatment group. Duration will be calculated as:

Duration (days) = study completion date - first date of study drug + 1 day

In addition, a categorical summary of duration of study follow-up in weeks will be provided using the following intervals:

- <24 weeks
- ≥24 to <48 weeks
- ≥48 to <72 weeks
- ≥72 to <96 weeks
- ≥96 weeks (Part 1 data only, including Part 1 safety follow-up)

3.3.5.3. Dosing Intensity

The planned and actual IV volumes (mL) are used to calculate dosing intensity to study drug and are recorded on the Exposure - IV eCRF.

The total cumulative received IV volume will be the sum of the received IV volume of study treatment while they are on study treatment across all planned doses during Part 1, up to the date of treatment withdrawal, if participants discontinued from study drug early.

The total cumulative expected IV volume will be the sum of the expected IV volume of study treatment while they are on study treatment across all planned doses during Part 1, up to the date of withdrawal, if participants discontinued from study drug early.

Dosing intensity will be calculated as:

$$\text{Dose Intensity}(\%) = 100 \times \frac{\text{total cumulative received IV volume (mL)}}{\text{total cumulative expected IV volume (mL)}}$$

3.3.5.4. Treatment Compliance

The number of infusions of study drug is defined as the total number of infusions a participant received, as recorded in the eCRF. The total number of infusions received and expected of study drug will be summarized as a continuous measure.

Treatment compliance is defined as the number of infusions received divided by the number of infusions expected, summarized by treatment group. Treatment compliance is calculated across all planned doses during Part 1, up to the date of withdrawal, if participants discontinued from study drug early.

Treatment compliance (%) will be summarized in the following categories: "<80%", "≥80-<120%", "≥120%".

3.4. Efficacy Analyses

The following sections describe all the efficacy and PK/PD analyses of clinical and biomarker endpoints. Unless otherwise specified, all efficacy analyses are conducted using the FAS-symptomatic.

A summary table of the analyses planned for COA endpoints and biomarkers can be found in Appendix 6.

3.4.1. Adjustment for Covariates

For the analysis of change from baseline using pMMRM or MMRM models, the baseline of the endpoint will be included as a continuous covariate in the model. For the analysis of CGI-I observed values using pMMRM and MMRM models, the baseline of CGI-S score will be included as a continuous covariate. For these COA endpoints, the log-10 transformed baseline serum NfL concentration will be included in both pMMRM and MMRM models as a covariate.

Unless stated otherwise, visits will be considered as categorical.

3.4.2. Clinical Co-primary Endpoint: Clinical Dementia Rating Dementia Staging Instrument Plus National Alzheimer's Disease Coordinating Center Frontotemporal Lobar Degeneration Behavior and Language Domains - Sum of Boxes

This clinical co-primary endpoint is the CFB in the CDR plus NACC FTLD-SB score across Weeks 48, 72 and 96 as estimated by the percentage slowing of disease progression. All CFB values up to Week 96 will be included in the model.

3.4.2.1. Definition of Estimand

The primary clinical question of interest is: What is the percent slowing of disease progression across Weeks 48, 72 and 96 for the CDR plus NACC FTLD-SB score in symptomatic participants with FTD who are carriers of a *GRN* mutation causative of FTD treated with latozinemab versus placebo, regardless of study treatment discontinuation/adherence or use of prohibited medications?

The estimand is described by the following attributes:

Treatment condition: the investigational intervention regardless of study treatment discontinuation/adherence or use of prohibited medication.

Population: symptomatic participants with FTD who are carriers of a *GRN* mutation causative of FTD (hereafter, FTD-*GRN*). Further details of the included population can be found in the Protocol, including inclusion and exclusion criteria.

Endpoint: change from baseline in the CDR plus NACC FTLD-SB score at Weeks 48, 72 and 96.

Intercurrent events: The 2 intercurrent events 'study treatment discontinuation/adherence' and 'use of prohibited medications' are both addressed in the treatment condition and are handled with an intention-to-treat principle using the treatment policy strategy. The treatment policy will include observed values in the analysis whether or not these intercurrent events occurred.

Population-level summary: Percent slowing of disease progression in the treatment group relative to the placebo group as measured from the change from baseline across Weeks 48, 72, and 96.

Rationale for this estimand: This estimand aims to reflect how symptomatic participants with FTD-*GRN* will be treated in clinical practice while taking into account all available data for efficacy, including those who discontinue study treatment early.

Hypotheses: Refer to Section 2 for the hypothesis definitions.

3.4.2.2. Definition of Endpoint

Washington University's CDR is a global assessment instrument that yields 2 scores: an overall global score (GS) and the sum of boxes (SB). The sum of boxes score is a quantitative general index that provides more detailed information than the global score in participants with mild dementia (O'Bryant 2010). The CDR "boxes" are 6 domains of cognitive and functional performance applicable to Alzheimer's disease and related dementias: Memory, Orientation, Judgment and Problem Solving, Community Affairs, Home and Hobbies, and Personal Care. The CDR's later expansion to include 2 additional domains (Behavior/Comportment/Personality and Language) more specific to the assessment of frontotemporal lobar degeneration (FTLD) patients (Miyagawa 2020a; Miyagawa 2020b) raises the total sum of boxes score without changing the potential range of GS. The FTD-relevant version of the CDR is called the CDR plus NACC FTLD. The necessary information to rate each domain is scored using standardized methodology and obtained by a trained interviewer through a semi-structured interview of the participant and a reliable informant or collateral source (eg, a caregiver). Participants were excluded from the study if they had a CDR plus NACC FTLD-GS of 3. For the purposes of this trial, the FDA

recommended to modify the anchor for the Behavior/Comportment/Personality and to recode its values to enable clarity of scoring and ensure that outcomes are clinically meaningful.

The modified domains and their possible scores are listed below.

- Memory - 0, 0.5, 1, 2, 3
- Orientation - 0, 0.5, 1, 2, 3
- Judgment and Problem Solving - 0, 0.5, 1, 2, 3
- Community Affairs - 0, 0.5, 1, 2, 3
- Home and Hobbies - 0, 0.5, 1, 2, 3
- Personal Care - 0, 1, 2, 3 (Note: this domain does not include a 0.5 level)
- Addition for FTLD: Behavior/Comportment/Personality - 0, 1, 2, 3 (Note, the FDA recommended to combine the 0 and 0.5 levels of the Behavior box original values and to score it as 0.)
- Addition for FTLD: Language - 0, 0.5, 1, 2, 3

Higher scores indicate greater impairment. The CDR plus NACC FTLD-SB score is obtained by a simple sum over all 8 domain scores and thus ranges from 0 to 24. Partially scored domains or missing CDR plus NACC FTLD-SB will be considered as missing.

3.4.2.3. Main Analytical Approach

The primary estimator will use the pMMRM approach (Wang 2022), with the CFB in CDR plus NACC FTLD-SB to Weeks 24 through 96 as the outcome variable. The model will include fixed-effects of baseline CDR plus NACC FTLD-SB, visit (categorical), treatment group indicator and log-10 transformed baseline serum NfL concentration as a covariate. The model utilizes all available post-baseline data; however, as greater noise than signal is expected at the early stages of the trial, the pMMRM model will only assume the treatment effect is proportional across Week 48 to Week 96, and use the time points of Weeks 48, 72 and 96, to estimate the common percentage slowing of disease progression (θ). At Week 24, the treatment effect of percentage slowing will be modeled separately from Weeks 48, 72 and 96. The model is likelihood-based, in which missing data are assumed to be missing-at-random (MAR) (Mallinckrodt 2003). An unstructured covariance matrix will be used to model within-participant variance and covariance, and will be estimated using REML with robust "sandwich" variance estimation. In the event that the model fails to converge using the unstructured covariance matrix, the Fisher scoring algorithm will be used (using the option SCORING of PROC MIXED) to obtain initial estimates of the unstructured variance-covariance matrix. If this fails to converge, then a prespecified covariance structure will be used with robust (empirical) variance estimates until convergence is reached. Additional model details are described in Appendix 1.

The primary inference is the percentage slowing of disease progression in the treatment group compared to the placebo group across Weeks 48, 72 and 96 as calculated by the θ parameter in the pMMRM model. The point estimate, corresponding SE, 2-sided 95% confidence interval and p-value from the model will be presented. Additionally, least squares mean (LSMeans) from pMMRM and corresponding SEs by visit and treatment group will be reported and plotted.

Results from the main and supportive and sensitivity analyses (excluding the tipping point analysis) described below will be presented graphically in a forest plot. Effect estimates on the change from baseline scale at Week 96 will be plotted and reported alongside the corresponding Week 96 percentage slowing of disease progression.

3.4.2.4. Sensitivity Analyses for Missing Data

The following sensitivity analyses will evaluate multiple missing data imputation strategy to evaluate the robustness of the treatment effect under different missing-not-at-random (MNAR) mechanisms. A minimum of 100 imputed datasets will be used, and results will be synthesized using Rubin's rules (Rubin 1987).

Each multiple imputation (MI) analysis will assume a monotone missing data pattern. If it is necessary to create a monotone missing pattern, the data will first be imputed to create the monotone pattern, assuming intermittent missing data are MAR, using a multivariate Gaussian imputation model using the Markov chain Monte Carlo (MCMC) method with Jeffrey's prior for each treatment group, imputing the missing analysis values at each visit.

3.4.2.4.1. Sensitivity Analysis 1: Reference-based Pattern Imputation

Using an MNAR assumption, participants treated with latozinemab who discontinue early from treatment would be expected to resemble participants in the placebo group after discontinuation. The placebo group will be used as reference and imputation will follow a reference-based pattern imputation (RBPI) method under a pattern-mixture model framework (Ratitch 2013). Missing data are imputed as if the participant was in the placebo group throughout the study, but the imputation model parameters come only from participants in the placebo group. The observed values from the participants before early discontinuation will be used together with the imputation model to impute the missing data. Missing data will be imputed sequentially by time. Monotone linear regression models will be the basis of imputation models. Participants in the latozinemab group with missing data will contribute any observed values for calculation of their imputed values.

Analysis will follow the pMMRM main analytic model as outlined in Section 3.4.2.3. The statistical inference will be the percentage slowing in disease progression across Weeks 48 to 96, which will be pooled by Rubin's rules. The pooled estimate of the percentage slowing in disease progression and associated 95% CI and p-value will be reported.

3.4.2.4.2. Sensitivity Analysis 2: Tipping Point Analysis

A tipping point analysis assumes that when data are missing post-deviation, there is an underlying MNAR mechanism responsible, and a fixed amount (called a "shift") is added to correct for the implied bias (Ratitch 2013). In this way, the effect of treatment is tested under a range of clinically plausible shifts which deviate from the predicted values of the time points from participants with missing data.

A single set of multiply imputed data will be created under an MAR assumption. In each imputed dataset, the shift values will be applied to all participants at visits with missing data following early discontinuation, and they will be independently applied to the placebo group (δ_P) and the treatment group (δ_T). In this way, the tipping point for each pair of shifts can be graphically depicted. The initial values of the shift to be considered will be 0, 0.5, 1, 1.5 and

2 point worsening in the CDR plus NACC FTLD-SB. Additional values may be explored if deemed appropriate.

Imputed datasets will be analyzed using the same pMMRM main analysis model. Pooled estimates of the percentage slowing of disease will be computed for each shift combination. A graphical representation will be presented to aid interpretation and indicate the tipping point(s) where conclusions are overturned.

3.4.2.5. Supportive Analysis: Mixed Models for Repeated Measures

The estimator will be an MMRM model with the CFB in CDR plus NACC FTLD-SB as the outcome variable, using all post-baseline data up to Week 96. The model will include fixed-effects of treatment group, baseline CDR plus NACC FTLD-SB score, log-10 transformed baseline serum NfL concentration, time point and time point-by-treatment interaction. The model will assume a treatment policy estimand and will treat missing data as MAR. An unstructured covariance structure will be used for within-participant residuals, and the model will be estimated using REML with robust "sandwich" variance estimation. In case of model convergence issues, the Fisher scoring algorithm will be used (using the option SCORING of PROC MIXED) to obtain initial estimates of the unstructured variance-covariance matrix. If this fails to converge, then additional residual covariance structures will be explored with robust (empirical) variance estimates following the testing order as specified in Appendix 1.

The primary time point for inference of treatment comparison will be Week 96. The treatment group contrast in LSMeans of change from baseline for the treatment comparison of latozinemab versus placebo will be reported with associated 95% CIs and p-value for each post-baseline visit. In addition, the within-treatment group LSMean and corresponding SE and 95% CI for each post-baseline visit will be reported. Treatment group-specific LSMeans (±SE) over time will be plotted.

The LSMeans of CFB will also be used to convert the observed effect to a percent slowing of disease progression scale, expressed as the ratio of negative difference in LSMean of CFB of latozinemab over the LSMean of CFB of placebo, $\left[\frac{(-\text{difference in LSMeans of CFB})}{\text{LSMean of CFB placebo}}\right] \times = 100\%$, for each post-baseline visit.

Analogous to the pMMRM estimation of a common percent reduction in disease progression across visits, the LSMean difference from Weeks 48 to 96 will also be averaged. The average LSMean difference across these visits with associated 95% CI and p-value will be reported and will also be transformed to the percent slowing of disease progression scale.

As a qualitative check that the proportionality assumption for the pMMRM is reasonable and fits the data approximately well, the estimated mean trajectories of both treatment groups from both the pMMRM and the MMRM model will be overlaid in the same plot. Close alignment of the two sets of curves suggests the proportionality assumption holds reasonably well. However, in certain situations where the proportionality assumption is obviously violated, such as crossover of treatment groups between Weeks 48 to 96, the pMMRM model and results derived from this model (such as time savings of treatment effect) will be deemed uninterpretable. As a result, all statistical inferences will be based on results from the MMRM including the time savings of treatment effect analysis.

3.4.2.6. Exploratory Analyses

The following exploratory analyses are planned for this co-primary endpoint. These analyses will use the FAS-symptomatic under the treatment policy estimand.

Exploratory Analysis 1: Time-savings of Treatment Effect Analysis

This analysis will use the same pMMRM primary efficacy model and will complement the primary efficacy analysis (Wang 2024). The treatment effect will be estimated as time-savings in disease progression relative to the decline observed in the placebo group, using the co-primary endpoint CFB in CDR plus NACC FTLD-SB. The time-savings over the 96-week trial duration is calculated by $\theta * 96$, directly corresponds to the percentage slowing in disease progression. Additional details are provided in Appendix 2.

Time savings analysis based on the MMRM results will be calculated using a backward projection from Week 96. This is complementary to the pMMRM derived time savings and does not depend on a proportionality assumption. Additional details are provided in Appendix 2.

Exploratory Analysis 2: Win Probability of Treatment Effect

The win probability offers a complementary analysis in the case of a positive treatment effect of latozinemab. The win probability is the probability that a randomly selected participant from the latozinemab group does at least as well as a randomly selected participant from the placebo group. As such, the potential treatment effect is expressed as a probability-based statement of benefit. It is advantageous because the composite makes use of data from all participants, even if they discontinue from the study or die. It does not rely on distributional assumptions regarding the distribution or trajectory of scores, and only assumes the data are ordinal in nature.

The outcome is a composite of the CFB in the CDR plus NACC FTLD-SB, and survival time. A similar hierarchical endpoint has been used elsewhere (eg, the Combined Assessment of Function and Survival score for amyotrophic lateral sclerosis; Berry 2013). Participants will be assigned a single score to reflect symptomatic/functional change using the Week 96 CFB or the duration of their time in the study, and survival status. The following hierarchy will be used to determine the composite endpoint:

1. The CFB score at Week 96 in CDR plus NACC FTLD-SB (where smaller values are better) will be used first; or if not available then
2. Survivors without a CFB score at Week 96 will use their study duration in Part 1 (where longer survival times are better); or if not available then
3. Time to death due to any cause (where longer survival times are better).

The composite score will be ranked within each level of the hierarchy, where each level of the hierarchy are ranked higher (better) than all lower levels. That is, the smallest Week 96 CFB score in the CDR plus NACC FTLD-SB is given a rank of 1 (best); the largest Week 96 CFB is ranked higher than the longest survival time in the study; the shortest survival time in the study is ranked higher than the longest time to death, and the shortest time to death is ranked last (worst).

The ranked outcome will be transformed to win probabilities and analyzed by linear regression using “HC2” robust standard errors with a fixed effect of treatment group (Zou 2022; Zou 2023).

Treatment effect will be quantified by the win probability, and associated 95% CI and p-value will be reported based on the logit transformation.

3.4.3. Biomarker Co-primary Endpoint: Plasma Progranulin

This biomarker co-primary endpoint is the change from baseline of the log-10 transformed plasma PGRN concentration at Week 96. All CFB values up to Week 96 will be included in the model.

3.4.3.1. Definition of Estimand

The estimand for this endpoint is defined with similar attributes as for the clinical co-primary endpoint, with a similar clinical question, and the same treatment condition, population and intercurrent event strategy. The endpoint and population-level summary are presented below which differ from the clinical co-primary endpoint.

Endpoint: Change from baseline of the log-10 transformed plasma PGRN concentration at Week 96.

Population-level summary: The mean change from baseline of log-10 transformed concentrations in the treatment group relative to the placebo group at Week 96.

Hypotheses: Refer to Section 2 for the hypothesis definitions.

3.4.3.2. Definition of Endpoint

Analysis values for plasma PGRN concentrations will be used. If necessary, values below or above quantitation limits will be imputed based on the description in Section 3.1.3. The change from baseline of the log-10 transformed values will be used for analysis. Missing concentration values will not be imputed.

3.4.3.3. Main Analytical Approach

This analysis will use the FAS-symptomatic and will be conducted using an MMRM model in similar fashion as described for the supportive analysis of the clinical co-primary endpoint in Section 3.4.2.5.

The main analytical approach will use an MMRM model with the change from baseline of the log-10 transformed values as the outcome. The log-10 transformed biomarker value at baseline will be included as a covariate in the model, as well as the baseline CDR plus NACC FTLD-SB score.

The LSMeans (based on log-10 transformation) for each treatment group and associated 95% CIs will be reported for each assessed time point. Treatment group-specific LSMeans (±SE) over time will be plotted.

For ease of interpretation, LSMeans and LSMean differences (vs. placebo) will be further converted to original scale as a percentage, within-group percent difference in geometric mean from baseline and placebo-adjusted percent difference in geometric mean ratio from baseline, respectively.

3.4.3.4. Sensitivity Analysis for Missing Data

A reference-based pattern imputation will be conducted with the similar attributes as described for the clinical co-primary endpoint (refer to Section 3.4.2.4.1). The analysis of each imputed dataset will use the same MMRM model as specified in the main analysis method.

3.4.4. Key Secondary Endpoint: Clinical Global Impression-Severity

The key secondary efficacy endpoint is the CFB in CGI-S score. The estimand for the key secondary endpoint has the same characteristics and attributes as for the clinical co-primary endpoint estimand, with the only change being the outcome of interest.

The CGI-S is used by a clinician to rate the severity of a participant's disease relative to the clinician's past experience with participants who have the same disease using a 7-point ordinal scale (ranging from 1=normal, not at all ill; 2=borderline mentally ill; 3=mildly ill; 4=moderately ill; 5=markedly ill; 6=severely ill; and 7=among the most extremely ill participants).The CGI-S values and CFB values will be summarized by descriptive statistics treating values as continuous, by visit and by treatment group.

The CGI-S will be summarized as a categorical variable using frequencies and percentages, by visit and by treatment group.

The CGI-S will also be summarized by visit and by treatment group as the percentage with a score of at least 5 ("markedly ill"). In addition, CGI-S CFB scores showing a worsening of ≥2 points will be summarized by visit and by treatment group.

The CGI-S will be analyzed using a similar pMMRM analysis approach as in the clinical co-primary efficacy endpoint main analytical approach (refer to Section 3.4.2.3).

In addition, a supportive analysis using MMRM will also be conducted in a similar fashion as in the clinical co-primary efficacy endpoint.

For sensitivity analyses, the CGI-S will be analyzed using the methods described as in the clinical co-primary efficacy endpoint using the pMMRM with the RBPI method (see Section 3.4.2.4.1).

3.4.5. Additional Secondary Endpoints

The estimand for each additional secondary endpoint has the same characteristics and attributes as for the clinical co-primary endpoint estimand, with the only change being the outcome of interest. Note that for the CGI-I, as there is no baseline measure by definition, the baseline CGI-S will be used as a covariate in the model. Unless specified otherwise, the analyses will use the FAS-Symptomatic, using a treatment policy estimand.

3.4.5.1. Frontotemporal Dementia Rating Scale Score

The FRS is a 30-item scale with item responses of "All the time", "Sometimes", "Never" or "Not applicable" for each item. The total percentage score will be calculated as the number of items with a response of "Never", divided by the number of applicable items, excluding those that marked as "not applicable." The percentage score ranges from 0 to 100 inclusive. The percentage score will be converted to a logit score ranging from 5.39 to -6.66 using a look-up table that is defined as part of the FRS instrument.

Qualitative interpretation of the FRS can also be classified as:

- Very mild: if the percentage score is 96 (non-inclusive) to 100 (inclusive)
- Mild: if the percentage score is 79 (non-inclusive) to 96 (inclusive)
- Moderate: if the percentage score is 40 (non-inclusive) to 79 (inclusive)
- Severe: if the percentage score is 12 (non-inclusive) to 40 (inclusive)
- Very Severe: if the percentage score is 2 (non-inclusive) to 12 (inclusive)
- Profound: if the percentage score is 0 to 2 (inclusive)

A descriptive summary of the FRS percentage score and logit score (actual values and CFB) will be presented at each visit by treatment group. In addition, the FRS severity categorization will be summarized by frequency and percentages at each visit by treatment group.

The FRS logit score will be analyzed using the same pMMRM analysis approach as in the clinical co-primary efficacy endpoint main analytical approach (refer to Section 3.4.2.3).

In addition, a supportive analysis using MMRM will also be conducted in a similar fashion as in the clinical co-primary efficacy endpoint using the FRS logit score (refer to Section 3.4.2.5).

3.4.5.2. Clinical Global Impression-improvement

The CGI-I is used by a clinician to rate how much a participant's disease has improved or worsened relative to baseline using a 7-point ordinal scale (ranging from 1=very much improved; 2=much improved; 3=minimally improved; 4=no change from baseline; 5=minimally worse; 6=much worse; and 7=very much worse). Therefore, the CGI-I is not assessed at the baseline visit.

The CGI-I values will be summarized by descriptive statistics treating values as continuous, by visit and by treatment group. The CGI-I will also be summarized as a categorical variable, treating each point value as its own category, as well as dichotomized it into 1-5 vs. 6-7. The summary table will report frequency and percentages, by visit and by treatment group.

The observed CGI-I values will be analyzed using the same pMMRM analysis approach as in the clinical co-primary efficacy endpoint main analytical approach (refer to Section 3.4.2.3) with the difference being that the baseline CGI-S as covariate in the model.

In addition, a supportive analysis using MMRM will also be conducted in a similar fashion as in the clinical co-primary efficacy endpoint (refer to Section 3.4.2.5) with the difference being that the baseline CGI-S as covariate in the model.

3.4.5.3. Repeatable Battery for the Assessment of Neuropsychological Status

The RBANS is an assessment consisting of twelve subtests yielding index scores for 5 domains (immediate memory, visuospatial/constructional, language, attention, and delayed memory) and a total scale score. The raw scores from each subtest within a domain are converted to a summary score, or index score, for the domain by consulting normative data tables. Then the sum of index scores across the 5 domains is converted to a total score that summarizes the participant’s overall level of performance on this measure. The total score can range from 40 to

160 with a lower score indicating a greater cognitive deficit. The total scale score will be used for analysis.

If any participant has a missing sub-scale score, the RBANS total scale score will be considered as missing.

The RBANS values and CFB values will be summarized by descriptive statistics treating values as continuous, by visit and by treatment group.

The RBANS total scale score will be analyzed using the same pMMRM analysis approach as in the clinical co-primary efficacy endpoint main analytical approach (refer to Section 3.4.2.3).

In addition, a supportive analysis using MMRM will also be conducted in a similar fashion as in the clinical co-primary efficacy endpoint using the RBANS total scale score (refer to Section 3.4.2.5).

3.5. Exploratory Analyses

3.5.1. Time-to-event Analysis

3.5.1.1. Time-to-symptomatic Worsening

This analysis will use the FAS-Symptomatic.

The following definitions are based on overall scores to assess changes in disease progression or worsening. An increase in the global CDR score is a well-accepted definition of disease progression and clinical staging in neurodegenerative conditions such as Alzheimer's Disease (Lanctôt 2024). The definitions based on changes in the sum of boxes are intended to assess worsening by at least 2 points (Lansdall 2024) or by requiring confirmation at a subsequent visit to account for potential variation in the sum of boxes score when assessed repeatedly over time.

Time to symptomatic worsening will be defined as follows:

- Definition 1: time from first dose to first progression (worsening) in CDR plus NACC FTLD-GS by at least 1 level (eg, 0.5 to 1, 1 to 2) or death.
- Definition 2: time from first dose to first progression (worsening) in CDR plus NACC FTLD-SB by ≥2 points (O'Bryant 2010) or death.
- Definition 3: time from first dose to the first visit with any progression (worsening) in CDR plus NACC FTLD-SB which is confirmed by at the next consecutive visit, or death. If the participant's visit meeting the increase is the last available visit, then the time to this visit will be used instead.

Event times will be determined by the actual time, or for participants who die, the time to death will be used. If participants have no post-baseline scores to determine worsening based on the above definitions, or do not experience an event during follow-up, they will be censored at the earliest of the following: study day 1, study discontinuation or study completion date in Part 1, and will not be considered to have worsened.

Time-to-event analyses will present overall Kaplan-Meier curves for each treatment group. Quantiles of the time-to-event and associated 95% CIs will be estimated for each treatment group, if possible. Additional percentiles may be considered if the 25th percentile is not

estimable. Comparisons of time-to-event will use a stratified log-rank test. The stratification factor will be the as-randomized disease severity stratum.

In addition, the Kaplan-Meier estimator of worsening at Weeks 48, 72 and 96 and associated 95% CI will be estimated for both treatment groups. Confidence intervals will be estimated using the complementary log-log transformation. In addition, the risk difference at Weeks 48, 72 and 96 and associated 95% CIs will also be calculated using a pseudo-value approach based on generalized estimating equations with identity link (Klein 2007).

3.5.1.2. Time to Study Treatment Discontinuation

The following analysis will use the FAS.

Time to study treatment discontinuation is defined as the time from the date of first dose to any cause study treatment or study discontinuation, or death, whichever occurs first.

Time-to-event analyses will present overall Kaplan-Meier curves for each treatment group. The median time-to-event and 95% CI will be estimated for each treatment group. Comparisons of time-to-event will use a stratified log-rank test. The stratification factor will be the as-randomized disease severity stratum.

3.5.2. "Responder" Type of Analysis

The following analysis will use the FAS-symptomatic.

A binary outcome of worsening (nonresponse) status by Week 96 will be defined as the first worsening in CDR plus NACC FTLD-GS by at least 1 level (eg, 0.5 to 1, 1 to 2) or death. If this criterion is met, the participant is considered to have worsened. Otherwise, the participant is considered to as not worsened. Event status will use all available post-baseline scores. If the participant does not have any post-baseline scores, they will be considered as not worsened.

The frequency and proportion of worsening by Week 96 will be summarized by treatment group.

An alternative definition will be considered in which the event will be a composite of either the first worsening in CDR plus NACC FTLD-SB by ≥2 points or death.

Logistic regression will be used to analyze worsening status. The model will include baseline CDR plus NACC FTLD-SB score, baseline log-10 transformed serum NfL concentration, and treatment group.

The probability of worsening for each treatment group, and the treatment group LSMean difference (vs. placebo) and its associated 95% CI and p-value will be reported.

3.5.3. Quality of Life: EuroQOL-5 Dimensions

The EQ-5D is a standardized instrument developed by the EuroQol Group and consists of 2 parts: the EQ-5D descriptive system and the EQ VAS. EQ-5D descriptive system is comprised of 5 dimensions: Mobility, Self-care, Usual Activities, Pain/Discomfort, and Anxiety/Depression. Each dimension is rated on a 5-point scale for severity:

- Level 1: no problems
- Level 2: slight problem

- Level 3: moderate problem
- Level 4: severe problem
- Level 5: unable to function/extreme problem

This study uses the EQ-5D proxy version 1 in which the caregiver (proxy) provides their impression of the participant's health. The EQ-5D VAS records the study partner's responses to a self-administered participant quality of life assessment on that day, using a vertical VAS from 0 (best health imaginable) to 100 (worst health imaginable). As the VAS is a single number summary, no imputation will be performed.

The EQ-5D VAS (actual values and CFB) will be summarized at each visit, by treatment group.

For categorical ratings of each of the EQ-5D domains, the frequency and proportion of reported problems will be reported by visit and treatment group.

An MMRM analysis will be conducted with similar attributes as the clinical co-primary efficacy endpoint for the CFB in EQ-5D VAS score (refer to Section 3.4.2.5) using the FAS-symptomatic.

3.5.4. Quality of Life: Zarit Burden Interview

The ZBI is a 22-item caregiver self-reported questionnaire. Each item in the interview is a statement on which the caregiver provides a rating on a 5-point scale with response options ranging from 0 (Never or Not at All) to 4 (Nearly Always or Extremely). The total score is the simple sum of the item scores and ranges from 0 to 88. If any item is missing, the total score will also be considered as missing.

The scale further defined a categorization of qualitative interpretation of items will be:

- 0 to 20: little or no burden
- 21 to 40: mild-to-moderate burden
- 41 to 60: moderate-to-severe burden
- 61 to 88: severe burden

A descriptive summary of the ZBI, change from baseline ZBI, and the qualitative interpretation (defined above) will be reported at each visit by treatment group.

An MMRM analysis will be conducted with similar attributes as the clinical co-primary efficacy endpoint for the CFB in ZBI score (refer to Section 3.4.2.5) using the FAS-symptomatic.

3.5.5. Pharmacoeconomic: Resource Utilization in Dementia-Lite Version

The RUD-Lite is a two-section questionnaire including questions about the caregiver and time spent caring for the participant, and about the participant, including questions about the participant's healthcare resource usage.

The RUD-Lite is a questionnaire rather than an assessment scale, and as such, it is not "scored." Rather, individual items are summarized. A subset of items focused on objective measures of health care utilization by the participant (eg, hospital admissions, ER visits) and time spent by

the caregiver (eg, assisting the participant with activities of daily living, supervision) will be summarized using descriptive statistics by visit by treatment group.

Data will be summarized as observed, by visit and by treatment group using the FAS-symptomatic.

3.5.6. Fluid Biomarker Analyses

PD fluid biomarker analyses will include blood-based biomarkers listed below. Optional CSF biomarkers will be summarized descriptively only. Missing biomarker data will not be imputed.

Blood-based and optional CSF biomarkers include:

- PGRN protein concentrations in optional CSF
- NfL protein concentrations in blood serum and optional CSF
- GFAP protein concentrations in blood plasma and optional CSF

Descriptive statistics will be used to summarize fluid biomarker parameters by analysis visit and treatment group, in terms of:

- observed (untransformed) values
- CFB
- PCFB
- log-10 transformed values (for geometric means)

Analysis of Blood-based Biomarkers

This analysis will use the FAS-symptomatic and will be conducted in similar fashion as described for the biomarker co-primary endpoint in Section 3.4.3.3.

These analyses will be conducted on the blood-based biomarkers (NfL and GFAP). An MMRM analysis will be used to model the change from baseline of the log-10 transformed concentration for each blood-based biomarker. The log-10 transformed biomarker value at baseline, and baseline CDR plus NACC FTLD-SB score will be included as covariates in the model.

The LSMeans (based on log-10) for each treatment group and associated 95% CIs will be reported for each assessed time point. Treatment group-specific LSMeans (±SE) over time will be plotted.

For ease of interpretation, LSMeans and LSMean differences (vs. placebo) will be further converted to original scale as a percentage, within-group percent difference in geometric mean from baseline and placebo-adjusted percent difference in geometric mean ratio from baseline, respectively.

3.5.7. Magnetic Resonance Imaging Biomarker Analyses

Structural data are derived from the 3D T1 sequence, acquired as per the parameters outlined in the AL001-3 Clario MRI Charter.

The FAS-symptomatic will be used for MRI imaging parameters summaries and analyses.

The minimal set of MRI parameters to summarize will be:

- Grey matter volumes from the following ROIs will decrease if atrophy occurs:
  - Bilateral frontal
  - Bilateral temporal
  - Bilateral parietal
- Bilateral ventricular volume will increase if atrophy occurs

PD imaging parameters (listed above) will be summarized by baseline volume and post-baseline atrophy measurements, by visit and treatment group. Volumes will be analyzed in units of cm^3. MRI scans at baseline will be measured for volumes of each ROI, whereas all post-baseline scans will be post-processed to provide atrophy (ie, change in volume since baseline) based on the MRI Imaging Charter.

Annualized rate of atrophy for the 4 vMRI bilateral ROIs (frontal, temporal, parietal, ventricular) will be summarized descriptively. Annualized rates are defined as follows:

$$\frac{\text{atrophy } [\text{cm}^3]}{(\text{post-baseline scan date - baseline scan date} + 1 \text{ [days]})/365.25 \text{ [days/year]}}$$

where atrophy is the change in volume (cm^3) in the brain region from the baseline to the last post-baseline scan, divided by the interval between these scans. This yields the average annualized atrophy rate over the study.

3.5.7.1. Treatment Effect on Magnetic Resonance Imaging Biomarkers by MMRM

These analyses will be restricted for 4 vMRI bilateral ROIs (frontal, temporal, parietal and ventricular) and will use the FAS-symptomatic. An MMRM analysis will be conducted with similar attributes as the clinical co-primary efficacy endpoint for the atrophy in each of the vMRI imaging parameters (refer to Section 3.4.2.5). The model will include baseline regional bilateral volume, log-10 transformed serum NfL at baseline, baseline CDR plus NACC FTLD-SB score, and age at baseline as covariates.

The LSMeans for each treatment group will be reported by visit and by treatment group. A line plot will also be presented showing the LSMean (±SE) for each treatment group by visit.

3.5.7.1.1. Supportive Analysis: Treatment effect on Magnetic Resonance Imaging Biomarkers by pMMRM

These analyses will be restricted to same bilateral ROIs and the FAS-symptomatic as Section 3.5.7.1. A pMMRM analysis will be conducted using the atrophy for each bilateral ROI as the outcome variable with similar attributes as the clinical co-primary endpoint (refer to Section 3.4.2.3). The model will include the baseline bilateral volume for each respective ROI, log-10 transformed serum NfL at baseline, baseline CDR plus NACC FTLD-SB score, and age at baseline as covariates. The results from pMMRM model can only be used for interpretation when the proportionality assumption is not violated after the qualitative check on proportionality assumption.

The percentage slowing of disease progression in the treatment group compared to the placebo group as calculated by the θ parameter will be reported for Weeks 48 and 96. The point estimate, corresponding SE, 2-sided 95% confidence interval and p-value from the model will be presented. Additionally, least squares mean (LSMeans) from pMMRM and corresponding SEs by visit and treatment group will be reported and plotted.

3.5.8. Winterlight Labs Speech Analysis (Optional Assessment)

Analyses of Winterlight Lab Speech data may be conducted as data permits and will be covered in a separate exploratory analysis plan, if sufficient data are available to conduct an analysis.

3.5.9. Association Analyses Between Biomarkers and Clinical Co-primary Endpoint

Treatment group-level correlations will be used to assess the association between changes in the clinical co-primary endpoint and biomarker changes across visits. To assess the group-level association between the treatment effect in the clinical co-primary endpoint (CDR plus NACC FTLD-SB) and the log10 transformed biomarkers (serum NfL and plasma GFAP), a correlation analysis will be conducted on the LSMeans of the change from baseline from each treatment group across visits in which both endpoints are measured (ie, Weeks 24, 48, 72 and 96). LSMeans and associated SEs will be derived from the respective MMRMs for the clinical co-primary endpoint and biomarker. LSMeans from corresponding visits for each treatment group will be plotted with SE error bars.

Spearman's correlation coefficient, weighted by sample size at each time point will be displayed on the plot for each treatment group.

3.5.10. Causal Mediation Analyses of Selected Biomarkers

The biomarker and imaging endpoints considered are the fluid biomarkers (log-10 transformed serum NfL and plasma GFAP) and 4 vMRI bilateral ROIs (frontal, temporal, parietal and ventricular). The objective of these analyses is to determine whether fluid or imaging biomarkers mediate putative treatment effects of latozinemab on the CFB in CDR plus NACC FTLD-SB score at Week 96. These analyses will be applied to participants from the FAS-Symptomatic with available data for the relevant biomarkers and endpoints.

A causal mediation analysis will be conducted based on the counterfactual framework using flexible regression methods (VanderWeele 2005; VanderWeele 2009). Details of the method can be found in Appendix 3.

The total treatment effect (vs. placebo) of CDR plus NACC FTLD-SB at Week 96 will be decomposed into direct and indirect effects, as well as the percentage of the treatment effect that is mediated by changes in biomarker levels at Week 48. Associated standard errors and 95% confidence intervals will be estimated using bias-corrected bootstrap methods with 2,000 replicates.

3.6. PK and PK/PD Analyses

The following planned analyses will only apply to listing and summary of serum and CSF PK concentration data. A separate analysis plan will be developed to describe population PK and exposure-response analyses.

Serum and CSF concentration at each study visit and timepoint will be presented in data listings. Serum and CSF concentration data will be summarized by study visit and timepoint for the PK analysis population using descriptive statistics: number of participants with non-missing data (n), mean, SD, CV, median, minimum, maximum, geometric mean, and geometric CV. For summary statistics and mean plots, all concentrations that are below LLOQ (BLQ) before the first dose will be set to zero and after first dose will be set to missing.

3.7. Subgroup Analyses of Interest

To permit modeling convergence, if a subgroup is too small, categories may be pooled prior to analysis, and footnoted appropriately.

This study is not statistically powered to detect a prespecified treatment effect for any predefined subgroups. Nominal p-values will be presented without adjustment for multiple comparisons.

When appropriate, forest plots will be used to show the treatment effect and 95% CI at Week 96 to visually assess the consistency of the treatment effect across subgroups.

3.7.1. Subgroup Analyses for Clinical Efficacy Endpoints

The clinical co-primary endpoint (CDR plus NACC FTLD-SB) and key secondary endpoint CGI-S will be summarized and analyzed using pMMRM (refer to Section 3.4.2.3) and the MMRM (refer to Section 3.4.2.5) by the following subgroups, based on the FAS-symptomatic. Each subgroup will be modeled separately.

- Age group (<65 or ≥65 years)
- Sex (Male, Female)
- Clinical phenotype based on first diagnosis at baseline
  - Defined as either PPA or bvFTD if that is the only diagnosis; or
  - If both bvFTD and PPA were diagnosed, use the first identified diagnosis. If the first diagnosis is unknown, then the diagnosis will be determined by sequential comparison of: ages at diagnosis, ages at symptom onset and finally comparison of CDR plus NACC FTLD language and behavior scores at baseline.
- Geographic region
  - Definition 1: North America versus outside of North America
  - Definition 2: Europe versus outside of Europe
- CDR plus NACC FTLD Global score at baseline (≤1; versus 2).
  - The CDR plus NACC FTLD Global score will be rederived following the recoding of the Behaviour box (refer to Section 3.4.2.2). The algorithm proposed by the Advancing Research and Treatment in Frontotemporal Lobar Degeneration (ARTFL)/Longitudinal Evaluation of Familial Frontotemporal Dementia Subjects (LEFFTDS) Consortium (Miyagawa 2020a) is used to rederive the CDR plus NACC FTLD-Global score after performing the recoding of Behaviour box.

In addition, subgroup analyses of the FRS logit score will be conducted for clinical phenotype and CDR plus NACC FTLD Global score at baseline.

Race will not be included as a subgroup as there will be insufficient data to perform subgroup analyses. For geographical regions, the Rest of World category was too small to perform subgroup analyses, so it will be included with "outside North America" and "outside Europe" subgroups.

3.7.2. Subgroup Analyses for Fluid Biomarkers

The fluid biomarkers (plasma PGRN, serum NfL and plasma GFAP) will be summarized and analyzed using the MMRM analysis model (refer to Section 3.4.2.5) by the following subgroups based on the FAS-symptomatic:

- Age group (<65 or ≥65 years)
- Sex (Male, Female)
- CDR plus NACC FTLD Global score at baseline (≤1; versus 2) (as noted above)

3.8. Analysis Considerations for At-risk Participants

As described above, the focal analysis population for efficacy and biomarker endpoints are based on symptomatic participants. Given the small sample size of at-risk participants, we plan for the following summary and selected modeling analysis to characterize the treatment effect in the combined symptomatic and at-risk participants.

Summary statistics will be provided for all efficacy, biomarker endpoints, as well as for some selected imaging endpoints for the FAS (Symptomatic and At Risk) and FAS-at Risk.

For the clinical co-primary endpoint CDR plus NACC FTLD-SB, the pMMRM model will be conducted in a similar fashion as described in the primary analysis for the FAS (Symptomatic and At-risk).

For the fluid biomarkers (plasma PGRN, serum NfL and plasma GFAP), the MMRM model will be conducted in a similar fashion as described in the supportive analysis for CDR plus NACC FTLD-SB for the FAS (Symptomatic and At-risk).

For the time-to-event analyses, Kaplan-Meier plots and a log-rank test will be conducted using the FAS (At-risk).

3.9. Safety Analyses

Safety assessments include AEs; physical and neurological examinations; vital signs and body weight; ECGs; MRIs; clinical laboratory analyses in blood and urine; suicidality assessments; and ADAs.

The default analysis set for safety is the SAF (ie, all dosed participants regardless of symptomatic or at-risk status), unless otherwise specified.

3.9.1. Adverse Events

AEs will be collected and coded by system organ class (SOC) and PT using MedDRA, Version 26 or newer.

3.9.1.1. Treatment-emergent Adverse Events

Treatment-emergent AE (TEAE) will be defined as those events that first occurred after treatment initiation in Part 1, and on or before the earliest of +28 days after the last treatment received in Part 1, or the date of death, or the clinical data cutoff date (as agreed with the regulatory authority), or if participants continue into study Part 2 (OLE) then 1 day prior to the first dose received in Part 2.

For the purpose of inclusion in TEAE summaries, incomplete AE onset and end dates will be imputed using the rules outlined in a separate technical document.

3.9.1.2. Serious Adverse Events

SAEs will be identified and captured as SAEs if AEs meet the definitions of SAEs that were specified in the Protocol (Section 7).

3.9.1.3. Relationship of Adverse Events to Study Drug

The relationship of AEs to study drug is in the investigator's opinion in the eCRF.

3.9.1.4. Summaries of Adverse Events

An overall summary table will present the number and percentage of participants who experienced at least 1 TEAE for the following events for SAF and SAF-symptomatic:

- TEAE
- Treatment-related TEAE
- TEAE leading to treatment discontinuation
- TEAE leading to study drug interruption
- TESAE
- Treatment-related TESAE
- TEAE with severity grade of Severe or higher
- Treatment-related TEAE with severity grade of Severe or higher
- TEAE leading to death (ie, outcome is fatal)
- TEAE related to COVID-19 (as indicated on the Adverse Event eCRF)

In addition, the number of total TEAEs, TESAEs and deaths will also be summarized in the overall summary table. Table 3 describes which summaries of TEAEs will be produced by analysis set.

If more than one TEAE occurs with the same preferred term for the same subject, then the subject will be counted only once for that preferred term using the most severe for the summary by severity.

All AEs and TEAEs will be listed. The details will be included in a separate technical document and shell specifications.

Table 3: Summary of Treatment-emergent Adverse Event Tables by Analysis Set

TEAE Summary Table	Analysis Set(s)
All TEAEs:	
• TEAEs by SOC and PT	1,2
• TEAEs by PT in decreasing frequency	1
• TEAEs by SOC, PT and maximum severity	1
All treatment-related TEAEs:	
• Treatment-related TEAEs by SOC and PT	1,2
• Treatment-related TEAEs by PT in decreasing frequency	1
• Treatment-related TEAEs by SOC, PT and maximum severity	1
All TEAEs leading to study treatment discontinuation or interruption	
• TEAE leading to treatment discontinuation by SOC and PT	1,2
• TEAE leading to treatment discontinuation by PT in decreasing frequency	1
All TESAEs:	
• TESAEs by SOC and PT	1,2
• TESAEs by PT in decreasing frequency	1
All treatment-related TESAEs:	
• TESAEs by SOC and PT	1,2
• TESAEs by PT in decreasing frequency	1
All TEAE leading to death:	
• TEAE leading to death by SOC and PT	1

Note: 1=SAF; 2=SAF-symptomatic.

3.9.2. Sheehan-Suicidality Tracking Scale

The Sheehan-STS is designed to assess and monitor the phenomenon of suicidality. The Sheehan-STS is a 16-item scale that is administered by a clinician or participant through self-report. Each item is scored on a 5-point Likert scale (0=not at all, 1=a little, 2=moderately, 3=very, and 4=extremely). The standard Sheehan-STS will be used to generate a total score as the simple sum of item scores.

Sheehan-STS total score will be summarized descriptively at each visit by treatment group.

In addition, the Sheehan-STS will be mapped to the 2010 Columbia Classification Algorithm of Suicide Assessment (C-CASA) (see details in Appendix 4). The C-CASA scale is intended to aid in clinical interpretation of suicidal ideation and behavior. Categorical results of the C-CASA will be summarized by visit and by worst post-baseline visit for each treatment group.

3.9.3. Laboratory Data

Clinical laboratory data to be summarized includes hematology, clinical chemistry, coagulation, and urinalysis. Handling of non-numeric qualifiers for below or above limits of quantitation are outlined in Section 3.1.3.

Central laboratory test results (actual values and CFB) will be summarized with descriptive statistics by parameter for each post-baseline visit by treatment group.

Shift tables will be used to summarize the change from baseline to worst post-baseline value with counts and percentages for hematology and serum chemistry.

For selected laboratory parameters, potentially clinically significant (PCS) categories are defined in Appendix 5. The number and percentage of subjects with post-baseline PCS for each PCS category will be summarized by visit and overall post-baseline visit and by treatment group.

Central laboratory data along with the normal ranges when applicable will be listed. Local laboratory data will not be summarized but will be listed.

3.9.4. Vital Signs

Vital sign parameters to be summarized include: systolic blood pressure (mmHg), diastolic blood pressure (mmHg), respiration rate (breaths/min), temperature (C), and pulse rate (beats/min) and will be assessed at all visits. Height will be collected at Screening. Weight (kg) will be collected at scheduled visits (refer to Protocol for Schedule of Assessments).

Vital signs (actual values and CFB for continuous parameters) will be descriptively summarized by visit.

Selected vital signs have PCS categories defined in Appendix 5, and will be tabulated by visit and overall post-baseline visit by treatment group.

3.9.5. Electrocardiogram Parameters

Local ECGs will be analyzed from a clinical safety basis (without intensive QT analysis). Local ECG parameters to be summarized include QT, QTcF, PR, and QRS intervals in msec and heart rate (beats /min). Means of triplicate measurements will be used for this summary, excluding zero values. If there is at least 1 but less than 3 available measurements with non-zero values at a given timepoint, then mean will be calculated over these available values. If there are no available measurements with non-zero values at a given timepoint, the mean will be considered missing.

Average values and CFB for ECG parameters will be summarized separately by pre- and post-infusion and by treatment group.

Clinical interpretation of ECG tracings will be summarized by frequency and percentage of participants in each category.

Shift tables from baseline to post-baseline visit, and worst post-baseline visit will be provided separately by measurement time point (pre- and post-infusion) by visit and by treatment group.

PCS for ECG parameters will be summarized in a similar manner as the clinical judgment described previously. PCS criteria are listed in Appendix 5.

Triplicate ECG tracings for approximately 60 consecutive participants will be read by a core ECG laboratory. The purpose is to provide analysis of standard ECG waveforms and QTcF intervals. Results from centrally read ECGs will not be captured in the eCRF but will be provided by a vendor and will be analyzed according to a separate SAP.

3.9.6. Physical and Neurological Examination

Physical and Neurological Examination abnormalities are captured as part of Medical History or AEs, therefore no additional summaries or listings will be provided.

3.9.7. Magnetic Resonance Imaging

MRI Abnormalities are captured as part of Medical History or AEs.

Baseline and all post-baseline MRI findings will be provided in a listing for any participants that have had any documentation of lesions (eg, microhemorrhages, white matter disease involvement).

3.9.8. Other Safety Assessments

3.9.8.1. Pregnancy Test

Screening reproductive status and urine and/or serum pregnancy tests will be presented in data listings for all women of child-bearing potential.

3.9.8.2. Impacts Due to COVID-19

A listing will be provided for the impact on participant visits due to COVID-19 as documented in the eCRF.

3.10. Other Analyses

3.10.1. Immunogenicity Analyses

The exploratory analyses for immunogenicity endpoints will be performed on the SAF who received at least one dose of study drug and had at least one post-baseline sample evaluable for immunogenicity (with reportable result). Only evaluable subjects are considered for computing treatment-induced ADA incidence.

For each scheduled timepoint for ADA evaluation, blood samples are analyzed for anti-latozinemab antibodies. First, a screening assay is performed, and the results of which can be either negative (no further confirmation required) or potentially positive. If the screening result is potentially positive, then the data is validated by a confirmatory assay, which can be negative (no further assay required) or positive. If positive, additional quantification of the titer value will be performed. In addition, for positive ADA confirmatory assays, a neutralizing antibody (NAb) assay is performed.

The ADA status of a subject is determined based on the longitudinal results from the screening/confirmation assay for anti-latozinemab binding antibody and the titer in the confirmed samples as follows:

- ADA Negative: All samples (baseline and postbaseline) are negative.
- Positive Pre-existing ADA:
  - Baseline sample is positive and all postbaseline samples were negative, or
  - Both baseline and postbaseline samples are positive, but all positive postbaseline sample have a titer <4-fold of the baseline titer.
- Positive Treatment Emergent ADA (TEADA):
  - At least 1 postbaseline sample is positive if the baseline sample was negative, or
  - At least 1 postbaseline sample is positive with a titer ≥4-fold of the baseline titer if the baseline sample was positive.

Samples tested positive for ADA to latozinemab, NAb status is determined as follows.

- NAb Negative: All ADA positive samples of a subject are tested NAb negative
- NAb Positive: At least one postbaseline ADA positive sample is tested NAb positive

ADA incidence data from only the confirmatory assay will be presented in a frequency table by different subgroups. Treatment-induced overall ADA incidence will be computed as a percentage of the total number of evaluable subjects that were ADA-negative at baseline. Treatment-induced ADA incidence at each visit will be computed as a percentage of the total number of evaluable subjects at that particular visit. Continuous variable like ADA titer will be presented with summary statistics such as n, mean, SD, median and IQR.

NAb incidence will be presented as a frequency table by different subgroups. Only overall incidence will be reported.

Individual results from ADA screening assay, confirmatory assay, titer, and NAB assay will be reported in a listing for each subject by visit.

4. INTERIM ANALYSIS

No interim analyses are planned for Part 1 of this study. At the time of database lock, all participants will have either early discontinued or completed the Week 96 visit for the Part 1 double-blind phase of the study. The analysis will be considered as the final analysis for Part 1 of the study.

Data collected during Part 2 OLE will be analyzed according to a separate SAP.

5. CHANGES FROM AND ADDITIONS TO METHODS PLANNED IN THE PROTOCOL

The following changes from the Protocol are noted with rationale.

5.1. Secondary Endpoints

The Protocol specified that the secondary endpoints (CGI-S, CGI-I and RBANS) would be tested under multiplicity adjustment. In this SAP, we propose the following changes. We propose that the CGI-S is the key secondary endpoint and retain the FRS, CGI-S and RBANS as additional secondary endpoints.

The sequential hypothesis testing orders are modified accordingly to reflect the above changes as specified in Section 2.1. The rationale for the proposed changes is that since the CGI-I and CGI-S are both clinician assessments of disease severity, which are highly correlated measures of a similar concept, it is deemed to be reasonable and sufficient to choose one of them as a key secondary endpoint to be included in the higher order of testing strategy. The CGI-S was chosen as the key secondary endpoint instead of the CGI-I because the CGI-S asks the rater to consider their ratings (eg, of "severely ill" or "among the most extremely ill patients") in the context of the rater's "total clinical experience with this particular population" at the current visit and dose not rely on rater recall. In contrast, the CGI-I relies on rater recall relative to baseline, which could be subject to recall bias over such a lengthy study, could be misinterpreted as change since the previous visit, and could be confounded by changes in raters over time. The RBANS was moved below the FRS in the testing order because of literature showing that the RBANS is not likely to be sensitive to changes in FTD (Fremont 2020) and due to the potential for a higher amount of missing data on the RBANS due to the functional components that participants are less likely to be able to complete in an advanced disease state. Thus, the testing strategy will prioritize the CGI-S followed by the secondary COA endpoints of the FRS, CGI-I, and RBANS.

5.2. Hypothesis Testing Strategy

The protocol stated the strategy for multiple testing using the Benjamini-Hochberg approach. In this SAP, we propose a sequential hypothesis testing order with predefined order specified in Section 2.1, where each step is gated on the success of the previous test in the hierarchy. This sequential testing procedure provides strong control of the overall family-wise type I error rate at 1-sided 2.5% level.

5.3. Covariate Adjustment for Co-primary Clinical Outcome and Secondary Clinical Outcome Assessment Endpoints

The Protocol did not specify in detail all covariates that would be included in the pMMRM model, which is used in the main analysis for the co-primary efficacy endpoint (CDR plus NACC FTLD-SB) or secondary efficacy COA endpoints (CGI-S, CGI-I, FRS and RBANS). For each of these efficacy endpoints, the pMMRM and supportive MMRM models will include the log-10 transformed baseline serum NfL concentration as a covariate. Inclusion of potentially prognostic covariates is supported by the 2023 FDA Guidance on covariate adjustment, in which the inclusion of potentially prognostic covariates may realize a gain in efficiency, or if not, there may be no gain at all.

6. CHANGES FROM OR ADDITIONS TO PREVIOUS VERSIONS OF THIS SAP

Below is a summary of changes made to version 1.0 of the SAP, dated 03 Mar 2025, and reflected in version 2.1-United States, dated 31 Jul 2025.

The major change since the previous version of this SAP is based on the recommendation from the FDA requesting to add the plasma PGRN as a co-primary endpoint alongside the existing clinical primary endpoint. This has resulted in numerous updates throughout the document, with the important updates noted below:

- Section 2.1: Update the hypothesis testing strategy to account for the co-primary endpoints;
- Section 2.3: Revise the definition of the FAS to remove the condition that participants in the set need to have a baseline score for the clinical co-primary endpoint;
- Section 3.4.3: Elevate plasma PGRN as a co-primary endpoint;
- Section 1.2.3: Provide a *post hoc* power calculation for the biomarker co-primary endpoint to ensure adequate power to accommodate this change;
- Section 3.4.3.4: Add a sensitivity analysis for missing data for the biomarker co-primary endpoint;
- Section 3.7.2: Add relevant subgroup analyses for the biomarker co-primary endpoint;
- Section 3.8: Add analysis of the biomarker co-primary endpoint that includes the presymptomatic (at risk) participants.
- Section 3.4.2.5: To address the concerns that the pMMRM model proportionality assumption may not hold, we have added a graphical check comparing results to the MMRM and clarified that all statistical inferences including the time saving analysis will be based on results from the MMRM model, if the proportionality assumption does not hold.
- Section 3.5.2: The responder analysis definitions were modified to align with the definitions used for the time-to-event analysis for symptomatic worsening.
- Section 3.5.7: Some exploratory MRI parameters were removed in order to focus on key brain regions (arterial spin labelling, fractional inosotropy, laterality indices, cortical thickness, additional brain regions). In addition, a support analysis using the pMMRM as the analytical model was added to the planned primary analysis model, MMRM, for volumetric MRI parameters to assess whether atrophy rates demonstrate a proportional change in response to treatment.
- Section 3.7: The definitions for some subgroups variables have been clarified.
- Section 3.7.1: An additional set of subgroup analyses was added for the FRS secondary endpoint.
- Section 3.10: Removed exploratory analysis of *TMEM106b* gene variants.

- Appendix 1: Empirical “sandwich” error estimates will be used for both pMMRM and MMRM model estimation following the 2023 FDA Guidance on Covariate Adjustment. This change reflects that empirical estimates tend to produce consistent estimates of variance even if distributional assumptions or the covariance structure are mis-specified.

Remaining changes in the SAP are editorial in nature to add additional detail or to improve readability of this SAP.

7. REFERENCES

Berry JD, Miller R, Moore DH, et al. The Combined Assessment of Function and Survival (CAFS): a new endpoint for ALS clinical trials. Amyotrophic lateral sclerosis & frontotemporal degeneration, 2013; 14(3), 162-168.

Center for Drug Evaluation and Research (CDER). Guidance for Industry: Adjusting for Covariates in Randomized Clinical Trials for Drugs and Biological Products. Federal Drug Administration (FDA). May 2023.

Fremont R, Manoochehri M, Armstrong NM, et al. Tolcapone Treatment for Cognitive and Behavioral Symptoms in Behavioral Variant Frontotemporal Dementia: A Placebo-controlled Crossover Study. J Alzheimers Dis. 2020;75(4):1391-1403.

Friede T, Kieser M. Blinded sample size re-estimation in superiority and non-inferiority trials: bias versus variance in variance estimation. Pharmaceutical Statistics. 2013; 12:141-146.

Klein JP, Logan B, Harhoff M, et al. Analyzing survival curves at a fixed point in time. Statistics in Medicine. 2007; 26: 4505-4519.

Lanctôt KL, Boada M, Tariot PN, et al. Association between clinical dementia rating and clinical outcomes in Alzheimer's disease. Alzheimer's Dement. 2024; 16:e12522.

Lansdall CJ, Teng E, Chague J, et al. Care partner-informed meaningful change thresholds for the Clinical Dementia Rating-sum of Boxes for trials of early Alzheimer's disease. Alzheimer's Dement. 2024; 20: 5889-5900.

Lu K, Mehrotra DV. Specification of covariance structure in longitudinal data analysis for randomized clinical trials. Statistics in Medicine. 2010; 29: 474-488.

Mallinckrodt CH, Sanger TM, Dubé S, et al. Assessing and interpreting treatment effects in longitudinal clinical trials with missing data. Biol Psych. 2003; 53(8):754-760.

Miyagawa T, Brushaber D, Syrjanen J, et al. Use of the CDR(R) plus NACC FTLD in mild FTLD: Data from the ARTFL/LEFFTDS consortium. Alzheimers Dement. 2020a;16(1):79-90.

Miyagawa T, Brushaber D, Syrjanen J, et al. Utility of the global CDR((R)) plus NACC FTLD rating and development of scoring rules: Data from the ARTFL/LEFFTDS Consortium. Alzheimers Dement. 2020b;16(1):106-17.

O'Bryant SE, Lacritz LH, Hall J, et al. Validation of the new interpretive guidelines for the clinical dementia rating scale sum of boxes score in the national Alzheimer's coordinating center database. Arch Neurol. 2010;67(6):746-9.

Posner, K, MA Oquendo, M Gould, et al. Columbia Classification Algorithm of Suicide Assessment (C-CASA): Classification of Suicidal Events in the FDA's Pediatric Suicidal Risk Analysis of Antidepressants. Am J Psychiatry. 2007;164:1035-1043.

Ratitch B, O'Kelly M, Tosiello R. Missing data in clinical trials: from clinical assumptions to statistical analysis using pattern mixture models. Pharmaceutical Statistics 2013. 12:337-347.

Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York: John Wiley and Sons; 1987.

Sellami L, Rucheton B, Ben Younes I, et al. Plasma progranulin levels for frontotemporal dementia in clinical practice: a 10-year French experience. Neurobiol Aging. 2020;91:167.e1-167.e9.

Sheehan DV, Giddens JM, Sheehan IS. Status update on the Sheehan-suicidality tracking scale (S-STS). 2014. Innovations in clinical neuroscience. 2014 Sep;11(9-10):93.

Staffaroni AM, Quintana M, Wendelberger B, et al. Temporal order of clinical and biomarker changes in familial frontotemporal dementia. Nat Med. 2022;28(10):2194-206.

Valeri L, VanderWeele TJ. Mediation Analysis Allowing for Exposure-mediator Interactions and Causal Interpretation: Theoretical Assumptions and Implementation with SAS and SPSS Macros. Psychological Methods. 2013; 18:137-150.

VanderWeele TJ. Explanation in Causal Inference: Methods for Mediation and Interaction. 2005. New York: Oxford University Press.

VanderWeele TJ, Vansteelandt S. Conceptual Issues Concerning Mediation, Interventions and Compositions. Statistics and Its Interface 2009;2:457-468.

VanderWeele TJ. A Unification of Mediation and Interaction: A 4-Way Decomposition. Epidemiology 2014; 25:749–761.

Wang G, Liu L, Li Y, et al. Proportional constrained longitudinal data analysis models for clinical trials in sporadic Alzheimer's disease. Alzheimers Dement (N Y). 2022 Apr 5;8(1):e12286.

Wang G, Cutter G, Oxtoby NP, et al. Statistical considerations when estimating time-saving treatment effects in Alzheimer's disease clinical trials. Alzheimer' Dement. 2024 Aug;20(8):5421-5433.

Zou G, Zou L, Choi YH. Distribution-free Approach to the Design and Analysis of Randomized Stroke Trials With the Modified Rankin Scale. Stroke. 2022 Oct;53(10):3025-3031.

Zou G, Zou L, Qiu SF. Parametric and nonparametric methods for confidence intervals and sample size planning for win probability in parallel-group randomized trials with Likert item and Likert scale data. Pharm Stat. 2023 May-Jun;22(3):418-439.

8. SUPPORTING DOCUMENTATION

APPENDIX 1. MODEL DETAILS FOR THE PROPORTIONAL MIXED MODEL FOR REPEATED MEASURES

The pMMRM was developed as an extension of the traditional MMRM. The proportionality constraints introduced in the pMMRM are conceptually similar to the Cox proportional hazards model in as much as it models the treatment effect as a single proportional difference at each post-baseline visit. A pMMRM with freely estimated proportions at each visit is equivalent to the MMRM. By using time as a categorical variable in the pMMRM (same as in the MMRM), it thus releases the linearity assumption of a linear mixed effect (LME) model. On the other hand, pMMRM takes on the assumption of a proportional treatment effect of an LME model and thus uses assessments more efficiently than a MMRM.

Let y_{ijk} denote the longitudinal assessments for subject i at time j for treatment group k, $i = 1, 2 \ldots, \mathrm{n}_k, j = 0, 1, \ldots, m$, and $k = 1, 2$ represents the placebo group and the treatment group. The marginal pMMRM can be written as:

$$\Delta y_{ijk} = \beta_0 (y_{i0k} - \bar{y}_{\cdot 0 \cdot}) + \Delta_{j1}(1 - \theta_k) + \varepsilon_{ijk}$$

where Δy_{ijk} is the individual change from the baseline y_{i0k} and is calculated as $y_{ijk} - y_{i0k}$ for $j = 1, \ldots, m_i$. β_0 is an unknown parameter which represents the association between y_{i0k} and Δy_{ijk}. $\bar{y}_{\cdot 0 \cdot}$ is the marginal mean of the baseline values pooled over both treatment groups, and the inclusion of $\bar{y}_{\cdot 0 \cdot}$ enables the model to provide a treatment effect adjusted for the baseline value. $\bar{y}_{\cdot 0 \cdot}$ is the arithmetic mean of the baseline values of the treatment groups and the placebo group when the design is balanced (eg, equal sample size for both groups); and is the marginal mean when the design is unbalanced (eg, missing data lead to different sample size in the 2 groups). For an unbalanced design, $\bar{y}_{\cdot 0 \cdot}$ is the same as the one used in the MMRM least square mean estimation and can be obtained through the MMRM model. Parameters (to be estimated) include Δ_{j1}, the mean change from baseline for the placebo group at time j for $j \geq 1$, θ_k, the proportion treatment effect for group k with $\theta_k = 0$ for the placebo group; and ε_{ijk} is the within-subject error and is assumed to follow a multivariate normal distribution: $\mathrm{N}(\vec{0}, \boldsymbol{\Sigma}_{m \times \mathrm{m}})$, where $\boldsymbol{\Sigma}$ is an unstructured variance-covariance matrix. Like MMRM, the parameters in the marginal pMMRM model can be estimated using the ML or REML methods.

To facilitate model convergence, the estimates of Δ_{j1}, β_0 and $\boldsymbol{\Sigma}$ from MMRM will be used as the initial values of pMMRM. If the unstructured covariance matrix results in a lack of convergence, then the unstructured variance-covariance matrix will be estimated using the Fisher scoring algorithm (using the SCORING option of PROC MIXED) (Lu 2010). If this fails to converge, then the following structures will be used in sequence:

- Heterogeneous Toeplitz (TOEPH) covariance structure
- Heterogeneous first-order autoregressive (ARH(1)) covariance structure
- Heterogeneous compound symmetry (CSH) covariance structure
- Toeplitz (TOEP) covariance structure

- First-order autoregressive (AR(1)) covariance structure
- Compound symmetry (CS) covariance structure.

In the case that a structured variance-covariance matrix is used to enable the model to converge, the "sandwich" estimator of the variance-covariance matrix will be employed by using the EMPIRICAL option in the PROC MIXED procedure with REML estimation and default degree of freedom adjustment. The empirical (robust) covariance estimator is used following the 2023 FDA Guidance on Covariate Adjustment.

APPENDIX 2. DETAILS OF TIME-SAVINGS OF TREATMENT EFFECT ANALYSIS

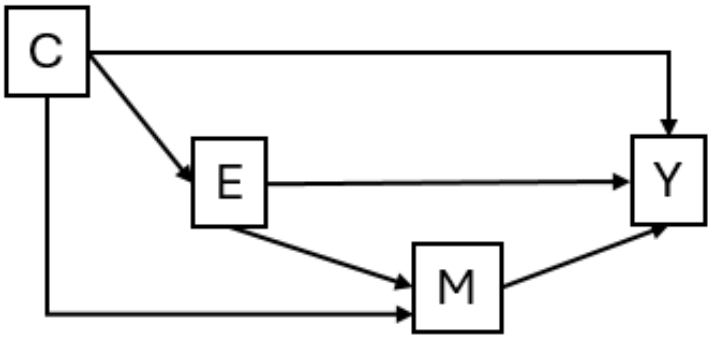
A time-savings of treatment analysis will be estimated using the pMMRM model of the primary efficacy analysis as $\theta * 96$. It will also be estimated based on MMRM outputs using a Backwards Projection to Placebo (BPP) method, the estimated mean change of the treatment arm (Δ_T) at a visit is projected backward in time until it aligns with the same mean value observed in the placebo arm at an earlier visit (Wang 2024). The time difference between these 2 points represents the treatment effect measured as time-savings in disease progression.

The underlying idea of the BPP method is that, within the trial duration, if both the placebo declined (Δ_P) and treatment declined (Δ_T), then the duration it took for the placebo arm to decline from Δ_T to Δ_P represents the time-savings throughout the trial. This method is directly related to the pMMRM estimation in that, conceptually, the proportion of time-savings over the trial duration (θ_T on the x-axis) is similar to the proportion of reduction in the disease progression (θ_{pMMRM} on the y-axis). When calculating each proportion based on the same visits, these 2 proportions are essentially the same. In such cases, the proportion of time-savings over 96 weeks directly corresponds to the reduction proportion in placebo group disease progression, as calculated by $\theta * 96$.

APPENDIX 3. CAUSAL MEDIATION ANALYSIS

Causal mediation is conceptually depicted as a causal diagram, showing the hypothesized relationships among the variables (Figure 2). There are 4 key variables of interest: the outcome variable (Y); treatment with latozinemab as the exposure variable (E); a putative mediator variable (M) and any confounding pre-exposure covariates (C).

Figure 2: Causal Mediation Directed Acyclic Diagram



Causal mediation is conducted using the regression approach under a counterfactual framework (Valeri 2013; VanderWeele 2014). The average total, direct and indirect effects will be calculated using linear regression by PROC CAUSALMED. In this context, the total effect of exposure to latozinemab is the difference in Week 96 CFB in CDR plus NACC FTLD-SB score, between those treated with latozinemab versus placebo. The total effect will be decomposed into 2 mutually exclusive components: the indirect (or mediated) effect, and the direct effect. The indirect effect is the effect of interest, and it indicates how much of the effect of latozinemab treatment on Week 96 CFB in CDR plus NACC FTLD-SB is mediated through changes in a specific biomarker (eg, Week 48 CFB in plasma NfL). Covariates will include the baseline (log-10 transformed) biomarker value, the baseline CDR plus NACC FTLD-SB score, and age. The direct effect is the effect that is through other pathways not involving the biomarker of interest as the putative mediator. The total effect will be decomposed into 4 distinct components: the controlled direct effect (ie, the effect due to neither mediation nor interaction), the pure indirect effect (ie, the effect due to mediation alone), the reference interaction (ie, the effect due to interaction alone), and the mediated interaction (ie, the effect due to both mediation and interaction). The percentage mediated is an estimate of the degree of mediation, obtained by dividing the indirect effect by the total effect. This approach gives valid inferences under the condition that the baseline covariates are sufficient to control for the exposure, the outcome, the mediator-outcome and the exposure-mediator confounding.

The outcome and mediator causal models will both be modeled by linear regression. Separate causal mediation models will be conducted using a common outcome (Week 96 CFB in CDR plus NACC FTLD-SB) and common set of covariates (baseline CDR plus NACC FTLD-SB and

an appropriate [log-10 transformed] baseline biomarker value), for each of the following mediators:

- log-10 transformed serum NfL
- log-10 transformed plasma GFAP
- Volumes of selected MRI ROIs (Bilateral frontal; bilateral temporal; bilateral parietal; bilateral ventricular)

The bootstrap method with bias-correction adjustment using 2,000 replicates will be used to construct 95% CI for all estimates.

APPENDIX 4. MAPPING S-STS TO C-CASA 2010 CATEGORIES

The C-CASA is a 9-point ordinal scale of severity of suicidality (Posner 2007). The mapping is provided to the 2010 C-CASA categories by Sheehan (2014) and is adapted below in Table 4.

To score a participant with a C-CASA code, each participant is evaluated according to the most severe category endorsed (ie, lowest code number) on the C-CASA.

Table 4: Mapping from the Sheehan-STS to the FDA 2010 C-CASA Categories

C-CASA Code Number	FDA 2010 C-CASA Coding Category	Sheehan-STS
1	Completed suicide	A "Yes" response to Question 17 OR Serious Adverse Event with Preferred Term of "Suicide completed" and the outcome is "Fatal"
2	Suicide attempt	(A positive response (>0) to Question 14 or Question 20) OR A "Yes" response to Question 1b OR Serious Adverse Event with Preferred Term of "Attempted suicide"
3	Preparatory acts toward imminent suicide behavior	A positive response (>0) to Question 12
4	Suicidal ideation	A positive response (>0) to any of: Questions 2, 3, 4, 5, 6, 7, 8, 9, 10 or 11 OR AE/SAE with Preferred Term of "Suicidal ideation"
5	Self-injurious behavior, intent unknown	[A positive response (>0) to Question 1a, AND (All of Questions 1b, 9, 10, 12, 14, 15, 16, 17 and 20 left unanswered/blank)] OR AE/SAE with Preferred Term of "Suspected suicide attempt" with non-fatal outcome
6	Not enough information (fatal)	A "Yes" response to Question 18
7	Self-injurious behavior, no suicide intent	A positive response (>0) to Question 13 OR (A positive response (>0) to Question 1a AND A negative response ("No") to Question 1b)
8	Other (accidental, psychiatric, medical), no deliberate self-harm	[(A "Yes" response or unanswered/blank response to Question 1) AND (A negative response (=0) to Questions 1a, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 17 and 18) AND (A "No" response to Question 1b OR A "No" response to Question 1 OR A negative response (=0) to Question 1a)] OR A positive response (>0) to Question 19
9	Not enough information (non-fatal)	A "Yes" response to Question 21 or if there is missing or incomplete information on the S-STS beyond the explicit rules above to allow mapping to codes 1-8 of the C-CASA.
10	No suicidal ideation or intent detected	No endorsement of codes 1 through 9, inclusive.

Adapted from Sheehan 2007.
The 10th category was added to facilitate counting participants with a complete Sheehan-STS assessment with otherwise no apparent suicidal ideation or intent.

APPENDIX 5. POTENTIALLY CLINICALLY SIGNIFICANT ABNORMALITIES

The following are the PCS criteria to determine potentially clinically significant abnormalities for vital signs and laboratory values (Table 5) and for ECG data (Table 6).

Table 5: Definitions of Potentially Clinically Significant Criteria for Vital Signs and Laboratory Values

Parameter or Analyte	PCS Flag	PCS Criteria
Vital Signs		
Pulse rate	Decreased	≤50 beats/min and decrease from baseline ≥15 beats/min
Pulse rate	Increased	≥120 beats/min and increase from baseline ≥15 beats/min
Systolic blood pressure	Decreased	≤90 mm Hg and decrease from baseline ≥20 mm Hg
Systolic blood pressure	Increased	≥180 mm Hg and decrease from baseline ≥20 mm Hg
Diastolic blood pressure	Decreased	≤50 mm Hg and decrease from baseline ≥20 mm Hg
Diastolic blood pressure	Increased	≥105 mm Hg and decrease from baseline ≥15 mm Hg
Weight	Decreased	Decrease from baseline ≥10%
Weight	Increased	Increase from baseline ≤10%

Table 6: Definitions of Potentially Clinically Significant Criteria for Electrocardiogram Data

Parameter	Label	Grade or Flag	PCS Criteria
Absolute QTcF interval	Absolute QTcF interval prolongation	Normal	≤450 msec
		Mild	>450 msec Note: if the QRS interval is ≥120 msec then a QTcF ≥450 msec and <480 msec is not considered abnormal
		Moderate	>480 msec and ≤500 msec
		Severe	>500 msec
Change from Baseline QTcF interval	QTcF interval prolongation	Normal	Increase from baseline ≤30 msec
		Mild	Increase from baseline >30 msec and ≤60 msec
		Moderate	Increase from baseline >60 msec
Absolute PR interval	PR interval	Low	<120 msec
	PR interval	High	≥200 msec
QRS duration	QRS duration	Low	<60 msec
	QRS duration	High	≥100 msec
Heart rate	Bradycardia	Low	<50 beats/min
	Tachycardia	High	>120 beats/min

APPENDIX 6. SUMMARY OF PLANNED CLINICAL AND BIOMARKER ANALYSES

Table 7: Summary Table of Planned Clinical Efficacy Analyses

		Co-primary Endpoint	Key Secondary Endpoint	Additional Secondary Endpoints		
Analysis	**Model**	**CDR plus NACC FTLD-SB**	**CGI-S**	**CGI-I**	**FRS**	**RBANS**
Main Analysis	pMMRM	X[a]	X[a]	X	X	X
Sensitivity Analysis to evaluate the impact of missing data[b]	Placebo-based Pattern Imputation	X[b]	X[b]			
	Tipping point analysis	X[a]				
Supportive Analysis	MMRM	X	X	X	X	X
Exploratory Analysis	Time-savings of treatment effect analysis	X				
	Win probability of treatment effect	X				
	Responder-type analysis	X[c]				
	Time-to-event analysis	X[c]				

[a] Analyses based on the FAS are also planned.

[b] Sensitivity analysis models make an MNAR assumption.

[c] Based on changes in the CDR plus NACC FTLD-SB or CDR plus NACC FTLD-GS.

Note: All analyses are based on FAS-symptomatic. Unless indicated otherwise, all models assume a treatment policy estimand and missing data will be considered to be missing-at-random.

Table 8: Summary Table of Planned Fluid and Imaging Biomarker Analyses

		Co-primary Endpoint	Exploratory Endpoints		
Analysis	**Model**	**Plasma PGRN**	**Serum NfL**	**Plasma GFAP**	**Imaging vMRI ROIs**
Main analysis	MMRM	X[a]	X[a]	X[a]	X
Sensitivity Analysis to evaluate the impact of missing data		X[b]			
Exploratory association analysis between biomarkers and CDR plus NACC FTLD-SB	Group-level correlation analysis		X	X	
	Causal mediation analysis		X	X	X

[a] The FAS will also be used for these analyses.
[b] The sensitivity analysis model makes an MNAR assumption.
Note: All analysis are based on FAS-symptomatic, unless specified otherwise. All models assume a treatment policy estimand and missing data will be considered to be missing-at-random.

Document Approvals

Approval Task Verdict: Approve	Arthur Mayorga, (arthur.mayorga@alector.com) Author Approval 31-Jul-2025 17:09:05 GMT+0000

Approval Task Verdict: Approve	Leonard Guizzetti, (leonard.guizzetti@alector.com) Author Approval 31-Jul-2025 17:15:31 GMT+0000

Approval Task Verdict: Approve	Jinglin Zhong, (jinglin.zhong@alector.com) Management Approval 31-Jul-2025 17:17:17 GMT+0000