STATISTICAL ANALYSIS PLAN

STUDY TITLE:	**A DOUBLE-BLIND, PLACEBO-CONTROLLED PARALLEL-GROUP STUDY IN PRECLINICAL PSEN1 E280A MUTATION CARRIERS RANDOMIZED TO CRENEZUMAB OR PLACEBO, AND IN NON-RANDOMIZED, PLACEBO-TREATED NON-CARRIERS FROM THE SAME KINDRED, TO EVALUATE THE EFFICACY AND SAFETY OF CRENEZUMAB IN THE TREATMENT OF AUTOSOMAL-DOMINANT ALZHEIMER'S DISEASE**
STUDY NUMBER:	GN28352
VERSION NUMBER:	*2*
TEST COMPOUND:	Crenezumab (RO5490245)
EUDRACT NUMBER:	TBD
IND NUMBER:	100839 (Non-IND study)
NCT NUMBER:	NCT01998841
SPONSOR'S:	F. Hoffmann-La Roche Ltd
LEGAL REGISTERED ADDRESS:	Grenzacherstrasse 124 4070 Basel, Switzerland
DATE FINAL:	See electronic date stamp above

PLAN PREPARED BY:

Date and Time(UTC)	Rea[redacted]ning	Name
09-Mar-2022 08:09:39	Company Signatory	[redacted] Ph.D.

STATISTICAL ANALYSIS PLAN AMENDMENT APPROVAL

STATISTICAL ANALYSIS PLAN VERSION HISTORY

This Statistical Analysis Plan (SAP) was developed based on Roche SAP model document Version 2.0, 26 October 2020.

SAP Version	Approval Date	Based on Protocol (Version, Approval Date)
2	See electronic date stamp on title page	Version 10, 23 August 2021
1	2 September 2021	Version 9, 19 March 2021

STATISTICAL ANALYSIS PLAN AND AMENDMENT RATIONALE

The SAP has been amended to update the safety analysis to include 16-week follow-up data and to update the type I error proof of the blinded power analysis. In particular, the protocol indicates that adverse events reported during the 16-week follow-up period of Study Period A will be summarized separately but according to this SAP, those events will be included in the analysis of Study Period A. The rationale of this departure from the protocol is explained in Section 5.6.

Key changes to the SAP, along with the rationale(s) for each change, are summarized below.

Section	Description of Change	Rationale for Change
1, 5.6	Adverse events reported during the 16-week follow-up period in Study Period A will not be summarized separately.	Study participants are considered treatment-exposed during the 16-week follow-up period (approximately 5 half-lives).
1.2.4	Details clarifying the process by which control cases are identified for adjudication are added.	
5	Types of analyses from Study Period A are clarified.	
5.3.2, 5.3.4	The main analysis and sensitivity analysis have been modified to include medications that may impact cognition.	
5.3.3, 5.4	The order of removing stratification factors is clarified.	
5.3.4	Sensitivity analysis changed to "first occurrence of discontinuing from crenezumab" instead of "first occurrence of discontinuing from study from the crenezumab arm".	If a participant discontinues from the study, then no subsequent data is expected to be collected.
5.6.2	The list of adverse events with special interested is augmented.	
5.6.8	Rationale of only analyzing partial anti-drug antibody data is added.	
Appendix 2	Mathematical proof of the blinded power analysis type I error control is generalized to both strong and weak sense.	

Additional minor changes have been made throughout to improve clarity and consistency.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation or Term	Description
AD	Alzheimer’s disease
ADAD	autosomal-dominant Alzheimer's disease
ADA	anti-drug antibody
API	Alzheimer’s Prevention Initiative
APOE	apolipoprotein E
ARIA-E	amyloid-related imaging abnormalities–edema/effusion
ARIA-H	amyloid-related imaging abnormalities–hemosiderin deposition
BP	blinded power
CDR	Clinical Dementia Rating
CERAD	Consortium to Establish a Registry for Alzheimer’s Disease
COVID-19	coronavirus disease 2019
CSF	cerebrospinal fluid
CSR	Clinical Study Report
ET-A	end of treatment visit A
ET-B	end of treatment visit B
FCSRT	Free and Cued Selective Reminding Task
FDA	U.S. Food and Drug Administration
FDG	fluorodeoxyglucose
GDS	Geriatric Depression Scale
HR	hazard ratio
iDMC	independent Data Monitoring Committee
iPAC	independent progression adjudication committee
ITT	intent-to-treat
IxRS	interactive voice or Web-based response system
LS	least squares
MCI	mild cognitive impairment
mITT	modified intent-to-treat
MMRM	mixed model with repeated measure
MMSE	Mini-Mental State Examination
MRI	magnetic resonance imaging
NCI CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
NfL	neurofilament light
OLE	open-label extension

Abbreviation or Term	Description
PACC	Preclinical Alzheimer Cognitive Composite
PET	positron emission tomography
PD	pharmacodynamic
PK	pharmacokinetic
RBANS	Repeatable Battery for the Assessment of Neuropsychological Status
RCRM	random coefficient regression model
ROI	region of interest
RPSFT	rank-preserving structural failure time
RR	relative reduction
SAP	Statistical Analysis Plan
SE	standard error
SUVr	standardized uptake value ratio
TTE	time-to-event

1. INTRODUCTION

This Statistical Analysis Plan (SAP) provides details of the planned analyses and statistical method for the clinical efficacy and clinical safety for Period A of Study GN28352. *The analyses specified in this document supersede the analysis plan described in the study protocol and the previous version of the SAP.* After *the primary analysis results* from Period A of Study GN28352 are available, the Sponsor may *extend Study Period B, or* initiate an open-label extension (OLE) study for *PSEN1* E280A autosomal-dominant mutation carriers. Key efficacy analyses and general descriptions of safety and other analyses in Period B of Study GN28352 and in *the* potential OLE *study* are also included in *Section 5.10 of* this SAP.

1.1 OBJECTIVES AND ENDPOINTS

Table 1 Objectives and Endpoints

Primary Objectives	Corresponding Endpoints
• To evaluate the efficacy of crenezumab treatment compared with placebo for at least 260 weeks in change in cognitive function as measured by the API ADAD Cognitive Composite Test total score in preclinical *PSEN1* E280A autosomal-dominant mutation carriers • To evaluate the efficacy of crenezumab treatment compared with placebo for at least 260 weeks on change in episodic memory function as measured in the FCSRT Cueing Index in preclinical *PSEN1* E280A autosomal-dominant mutation carriers	• Annualized rate of change in the API ADAD Composite Cognitive Test total score, which is computed from the following 5 cognitive test scores: – CERAD Word List Recall (Rosen et al. 1984; Morris et al. 1989; Mohs et al. 1997) – Multilingual Naming Test (Gollan et al. 2012) – MMSE for orientation to time (Folstein et al. 1975) – CERAD Constructional Praxis (a measure of visuospatial ability; (Morris et al. 1989) – Raven's Progressive Matrices, Set A (a measure of nonverbal fluid reasoning and visuospatial abilities; Raven 1976) • Annualized rate of change on the FCSRT Cueing Index
Secondary Objectives	**Corresponding Endpoints**
• Assess the effect of crenezumab on clinical progression • Assess the effect of crenezumab on overall cerebral fibrillar amyloid burden using predefined ROI from florbetapir PET	• Time to progression from preclinical AD to MCI due to AD (Albert et al. 2011) or from preclinical AD to dementia due to AD (McKhann et al. 2011) • Time to progression to non-zero in the CDR Scale global score (Morris 1993)

Table 1 Objectives and Endpoints (cont.)

Secondary Objectives	Corresponding Endpoints
• Assess the effect of crenezumab on regional CMRgl using FDG-PET measurements in predefined ROIs • Assess the effect of crenezumab on brain atrophy as measured by volumetric measurements using MRI • Assess the effect of crenezumab on tau-based CSF biomarkers	• Annualized rate of change in the CDR Scale-Sum of Boxes (Morris 1993) • Annualized rate of change in a measure of overall neurocognitive functioning: RBANS (Randolph 1998) • Annualized rate of change in mean *cerebral* fibrillar amyloid accumulation using florbetapir PET *from a predefined ROI* • Annualized rate of change in *regional* CMRgl using FDG-PET *in a predefined ROI* (Chen et al. 2010; Van Dyck et al. 2019) • Annualized rate of change in *volumetric measurements* using MRI • Annualized rate of change in CSF tTau and pTau
Safety Objective	**Corresponding Endpoints**
• To assess the safety and tolerability of crenezumab in preclinical *PSEN1* E280A mutation carriers (comparing crenezumab with placebo)	• Frequency and severity of treatment-emergent adverse events and serious adverse events • Withdrawals of study drug due to adverse events • Incidence of treatment-emergent adverse events of special interest, including: – ARIA-E – ARIA-H – Cerebral macrohemorrhages – Pneumonia • Incidence of injection reactions and IRRs • Incidence of treatment-emergent anti-crenezumab antibodies
Pharmacokinetic/Pharmacodynamic Objective	**Corresponding Endpoints**
• Collect sparse PK samples to support confirmation of exposure to crenezumab and to explore the PD response (as measured by plasma total Aβ levels comparing crenezumab with placebo) in preclinical *PSEN1* E280A mutation carriers	• CSF and serum crenezumab concentrations at protocol-specified timepoints • Trough serum crenezumab $C_{trough, ss}$ in serum • Plasma and CSF Aβ1-40 and Aβ1-42 concentrations

Table 1 Objectives and Endpoints (cont.)

Exploratory Objectives	Corresponding Endpoints
• Assess further the effect of crenezumab in preclinical *PSEN1* E280A mutation carriers on additional clinical measures of efficacy and biological markers of disease that have not been prespecified as primary and secondary endpoints in the SAP • Explore pharmacogenetic effects, including but not limited to, a person's APOE ε4 carrier status on the active treatment's cognitive, clinical, and adverse effects • Explore effects of genetic variation, including but not limited to, how genes affect the biology of AD and related disorders and how genes influence biomarker responses • Examine clinical and biomarker changes in *PSEN1* E280A non-carriers and to compare these changes with those seen in carriers treated with placebo • Relate the treatment's biomarker effects to clinical outcomes and to examine predictive and prognostic utility of baseline characteristics • Assess the impact of treatment on brain tau load over time, as measured by tau PET imaging in an optional substudy (GN28352-1/BN40199)	• Changes from baseline over time in the following cognitive measures: – Trail Making Test (Armitage 1946) – MMSE (Folstein et al. 1975) – RBANS Index Scores (Randolph 1998) – Scores of each of the components of the API ADAD Composite Cognitive Test Battery – PACC (Donohue et al. 2014) modified for the API ADAD trial – Clinical endpoints not examined in secondary outcome measures • Changes in the NPI (Cummings et al. 1994; 1997) total score, items, and factors • Changes in the GDS (Sheikh and Yesavage 1986) total score • FAST (Sclan and Reisberg 1992) total score • Changes in Subjective Memory Checklist (Acosta-Baena et al. 2011) • Changes in other blood and CSF measures such as NfL and plasma tau markers • Analysis of *image* ROIs not selected in secondary outcome measures • Other imaging outcome measures and analytic methods, alone, or in conjunction with other imaging modalities not explored in secondary outcome measures • Change in tau burden over time, as measured by GTP1 PET in an optional substudy (GN28352-1/ BN40199) • Changes in primary, secondary, and exploratory outcomes in mutation non-carriers treated with placebo • Comparisons of clinical and biomarker outcomes between carriers treated with placebo and non-carriers • Changes in primary, secondary, and exploratory outcomes in carriers and non-carriers as functions of APOE genotype and other genetic variations

Table 1 Objectives and Endpoints (cont.)

Exploratory Objectives	Corresponding Endpoints
	• Short-term changes in imaging measures as functions of initiation (e.g., baseline to 12 weeks) • Analyses of outcome measures in relation to one another and in relation to baseline characteristics

AD = Alzheimer's disease; API = Alzheimer's Prevention Initiative; APOE = apolipoprotein E; ADAD = autosomal-dominant Alzheimer's disease; ARIA-E = amyloid-related imaging abnormalities–edema/effusion; ARIA-H = amyloid-related imaging abnormalities–hemosiderin deposition; CDR = Clinical Dementia Rating; CERAD = Consortium to Establish a Registry for Alzheimer's disease; CMRgl = cerebral metabolic rate of glucose; CSF = cerebrospinal fluid; $C_{trough, ss}$ = concentration at steady state; FAST = Functional Assessment Staging of Alzheimer's disease; FCSRT = Free and Cued Selective Reminding Task; FDG = fluorodeoxyglucose; GDS = Geriatric Depression Scale; GTP1 = Genentech Tau Probe 1; IRR = infusion-related reaction; MCI = mild cognitive impairment; MMSE = Mini-Mental State Examination; MRI = magnetic resonance imaging; NfL = neurofilament light chain; NPI = Neuropsychiatric Inventory; PACC = Preclinical Alzheimer's Cognitive Composite; PET = positron emission tomography; PD = pharmacodynamic; PK = pharmacokinetic; RBANS = Repeatable Battery for the Assessment of Neuropsychological Status; ROI = region of interest; SAP = Statistical Analysis Plan.

1.2 STUDY DESIGN

1.2.1 Study Period A

This is a prospective, randomized, double-blind, placebo-controlled, parallel-group study of crenezumab versus placebo in individuals who carry the *PSEN1* E280A autosomal-dominant mutation causing early-onset Alzheimer's disease (EOAD), and do not meet criteria for mild cognitive impairment (MCI) due to Alzheimer's disease (AD) (Albert et al. 2011) or dementia due to AD (McKhann et al. 2011) and are, thus, in a preclinical phase of AD (Sperling et al. 2011). The study also incorporates administration of placebo to individuals who are not *PSEN1* E280A autosomal-dominant mutation carriers. This efficacy and safety study is being conducted at a single primary site and 3 satellite sites in Colombia.

In the Study GN28352, *PSEN1* E280A autosomal-dominant mutation carriers who met study eligibility criteria were randomized in a 1:1 ratio to one of two treatment groups: crenezumab or placebo. Crenezumab is being administered either SC (720 mg every 2 weeks) or IV (60 mg/kg every 4 weeks). Matching placebo is administered by the same routes at the same dosing regimen. The switch to the higher IV dose (approximately 4–fold higher exposure to crenezumab) was optional. Participants may decide whether to change from the SC dosing route to the IV dose route; however, once IV dosing in a given participant has occurred, it is not intended for a participant to switch back to the lower SC dose. In order to maintain genotype blind

and to have a genetic kindred control, a cohort of *PSEN1* E280A mutation non-carrier kindred family members were also enrolled in the study and dosed only with placebo.

The study includes three arms *and originally aimed to have* approximately 100 participants per arm; two arms with participants that have a *PSEN1* E280A mutation randomized in a 1:1 ratio to active treatment or placebo and a third arm of non-carriers randomized to placebo (see Figure 1). Participants will receive the randomized treatment until the last participants have reached their last treatment visits at 260 weeks (see Figure 2).

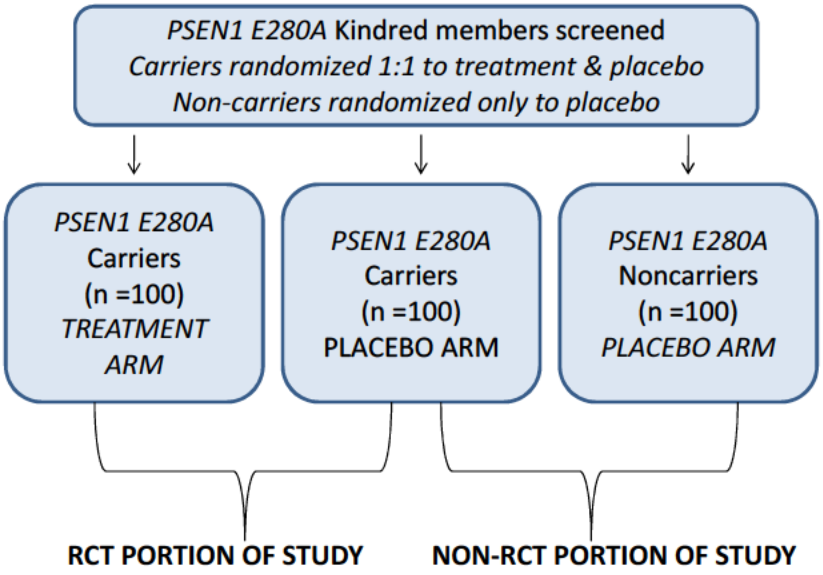
The study duration for individual participants will be at least approximately 284 weeks, including an 8-week screening period, a double-blind treatment period of at least 260 weeks in length, a 4-week final dose visit, and a final safety follow-up visit 16 weeks after the last dose of study drug (crenezumab or placebo) that will allow for clinical follow-up after treatment discontinuation for participants who do not continue with study drug beyond the end of Study Period A or who terminate study drug early. Per protocol, all participants are assessed with clinical cognitive measures, magnetic resonance imaging (MRI), fluorodeoxyglucose (FDG) positron emission tomography (PET), amyloid PET, and blood-based biomarkers per the schedule of activities. Cerebrospinal fluid (CSF) collection is optional and tau PET is collected in an optional substudy.

1.2.2 Study Period B

Following the completion of Study Period A, participants will be offered the opportunity to continue to receive study drug until the results of the study are known. This is termed Study Period B. All *PSEN1* mutation carriers will be provided crenezumab in Study Period B regardless of treatment allocation during Study Period A, and all non-carriers will continue on placebo. Treatment allocation for both Study Periods A and B will remain blinded. Study Period B may last up to approximately 12 months depending on when participants enter Study Period B. See Figure 2.

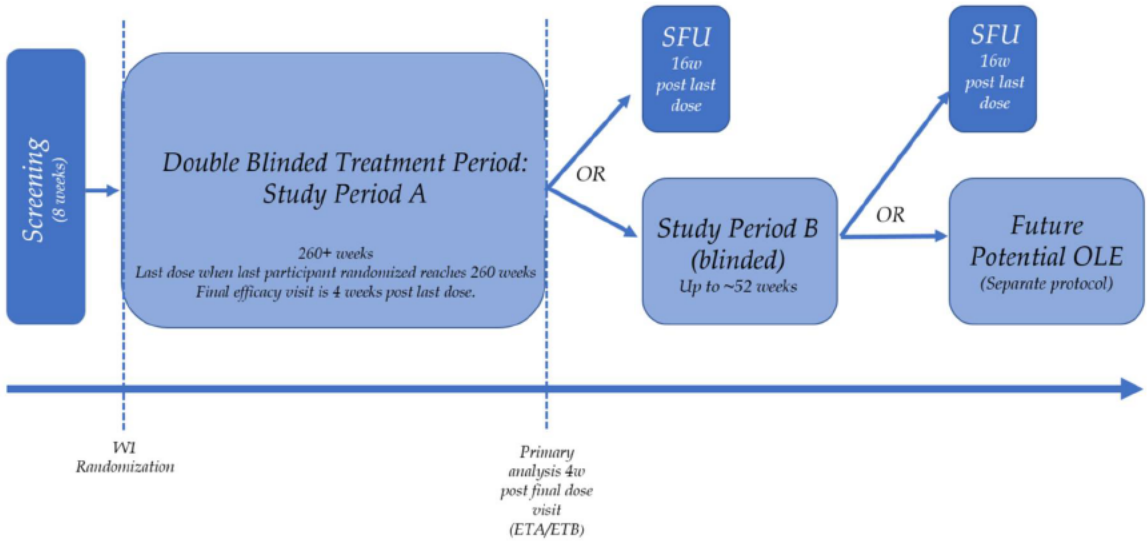
After the primary efficacy results from Study Period A are available, the Sponsor may extend Study Period B, or initiate an OLE study for PSEN1 E280A autosomal-dominant mutation carriers, or development of crenezumab may be discontinued.

Figure 1 Study Design



RCT = randomized, controlled trial.

Figure 2 Study Schematic



ETA/ETB = end of treatment visit A or end of treatment visit B; OLE = open-label extension; SFU = safety follow-up; W = week.

1.2.3 Treatment Assignment and Blinding

Both *PSEN1* E280A mutation carriers and non-carriers were enrolled. The majority of participants in the *PSEN1* E280A kindred cohort from Colombia do not wish to know their *PSEN1* genotype. In most cases, the participant will not know whether he or she is a mutation carrier or a non-carrier. Participants who had their mutation carrier status

disclosed by a physician in conjunction with genetic counseling outside of the auspices of this study and are mutation carriers were eligible for the study. Genetic disclosure performed independent of this study was not required for a mutation carrier to be eligible to participate. Non-carriers who received information about their genetic status prior to the study were not included.

Participants, blinded study personnel, and the Sponsor will not know if a participant has been assigned to the active treatment or one of the two placebo treatment groups in Study Period A. In Study Period B, participants and blinded study personnel will continue to be blinded to original and current treatment assignment. Following *the data snapshot* for *the* Study Period A *primary analysis*, *s*elected members of the Sponsor team may become unblinded for the purpose of analysis and interpretation of Study Period A data.

In Study Period A, a dynamic randomization design was used with treatment allocation assigned by the interactive voice or Web-based response system (IxRS) vendor. Mutation carriers were randomized to crenezumab and placebo arms in a 1:1 ratio. A smaller group of mutation non-carriers were assigned to placebo to conduct the study, such that autosomal-dominant AD (ADAD) family members were not required to receive information about their genetic risk and to help distinguish those changes related to the predisposition to AD from those associated with normal aging in the two placebo groups (in Study Period B, all carrier participants will be assigned crenezumab by the IxRS vendor and non-carrier participants will be assigned placebo).

While pharmacokinetic (PK) samples should be collected from participants assigned to the placebo arms to maintain the blinding of treatment assignment, PK assay results for these participants are generally not needed for the safe conduct or proper interpretation of this trial. A defined set of personnel responsible for performing PK assays and not otherwise involved in the conduct of the trial will be unblinded to participants' treatment assignments to identify appropriate PK samples to be analyzed. Samples from participants assigned to the placebo arms will not be analyzed except by request (i.e., to evaluate a possible error in dosing). In addition, PK and plasma pharmacodynamic (PD) assay results will not be released to blinded personnel until the study is unblinded.

If unblinding is necessary for participant management (e.g., in the case of a serious adverse event for which participant management might be affected by knowledge of treatment assignment), the investigator will be able to break the treatment code by contacting the IxRS. Treatment codes should not be broken except in emergency situations.

For regulatory reporting purposes, and if required by local health authorities, the Sponsor will break the treatment code for all serious, unexpected, and suspected

adverse reactions that are considered by the investigator or Sponsor to be related to study drug.

1.2.4 Independent Progression Adjudication Committee

Ascertainment of AD dementia or MCI progression will be made by the investigator. An independent progression adjudication committee (iPAC) was formed to review the information provided by the site investigators regarding participants enrolled in the trial. The iPAC will review data on:

- Participants who, in the opinion of the site investigator, appear to have progressed from not meeting criteria for MCI due to AD or dementia due to AD to meeting criteria for MCI due to AD or *to* dementia due to AD;
- Participants who, in the opinion of the investigator, appear to have progressed from meeting criteria for MCI due to other cause to meeting criteria for MCI due to AD or to dementia due to AD;
- Participants who, in the opinion of the investigator, appear to have progressed from meeting criteria for dementia due to other cause to meeting criteria for MCI due to AD or to dementia due to AD, and;
- A matching number of participants who, in the opinion of the investigator, do not appear to have progressed. *For each "case" the site submits, the data-coordinating center randomly selects a participant from the list of participants who have not yet been adjudicated as a progression using a random number generator. Once the participant has been identified, the data-coordinating center instructs the site to provide information needed for adjudication for that participant.*

The primary roles of iPAC members will be to:

- Provide guidance to the site and Sponsor regarding evaluation and as needed, collection of information necessary to adjudicate progression to MCI or dementia
- Review all participants' clinical information supplied in the supporting documents and make a determination as to whether the participant has progressed or not progressed from "normal cognition" (i.e., in this protocol, not meeting criteria for MCI or dementia) to MCI due to AD or dementia due to AD at the time of the most recent major clinical assessment at the site

If there is lack of consensus between the iPAC members as to progression of an individual participant, they will confer and arrive at a consensus. If the members' opinion differs with that of site investigators, one or both members will correspond with site investigators and attempt to arrive at a consensus. Failing that, the opinion of the iPAC will be used for purposes of statistical analysis; *the opinion* of the site investigators will be used for clinical management of the study participant. If after adjudication, both the iPAC members and site agree that the participant has not progressed, the participant will be re-adjudicated if deemed to progress by the site at a future visit.

1.2.5 Data Monitoring

An independent Data Monitoring Committee (iDMC) is being used during the conduct of this trial to periodically review unblinded safety and efficacy data. In addition to periodic review of safety and efficacy data, the iDMC may make recommendations regarding study conduct. Details of the iDMC membership and roles are detailed in the iDMC Charter.

2. STATISTICAL HYPOTHESES

The primary endpoint family consists of 2 endpoints: 1) annualized change on the Alzheimer's Prevention Initiative (API) ADAD Composite Cognitive Test total score and 2) annualized change on Free and Cued Selective Reminding Task (FCSRT) Cueing Index. The type I error is split between testing the treatment effect on the API ADAD Composite Cognitive Test total score ($\alpha = 0.04$) and on the FCSRT Cueing Index ($\alpha = 0.01$). If at least one of the 2 primary analyses is statistically significant, then the trial is considered positive. The primary endpoint family in the crenezumab arm and placebo arm will be compared among mutation carriers and tested using a random coefficient regression model (RCRM) with the null and alternative hypotheses as follows.

- API ADAD Composite Cognitive Test total score:
  - H_0: There is no difference in annualized change on the API ADAD Composite Cognitive Test total score between the crenezumab and placebo arms among mutation carriers
  - H_1: There is a difference in annualized change on the API ADAD Composite Cognitive Test total score between the crenezumab and placebo arms among mutation carriers
- FCSRT Cueing Index:
  - H_0: There is no difference in annualized change on FCSRT Cueing Index between the crenezumab and placebo arms among mutation carriers.
  - H_1: There is a difference in annualized change on FCSRT Cueing Index between the crenezumab and placebo arms among mutation carriers.

Details on the type I error control are described in Section 3.1.

3. SAMPLE SIZE DETERMINATION

This study will enroll up to 300 participants: up to 200 participants will be enrolled in the carrier cohort and up to 100 participants will be enrolled in the non-carrier cohort. Participants in the carrier cohort will be randomized in a 1:1 randomization ratio to active treatment or placebo. The study was initially powered to compare the mean change from baseline over 260 weeks in the API ADAD Composite Cognitive Test total score between the active treatment group and the placebo group. Assuming a 25% dropout rate with use of two-sided testing at the overall 0.05 level, a placebo group coefficient of variation of 65% for the Week 260 change scores ($= 100\% \times$ standard deviation of

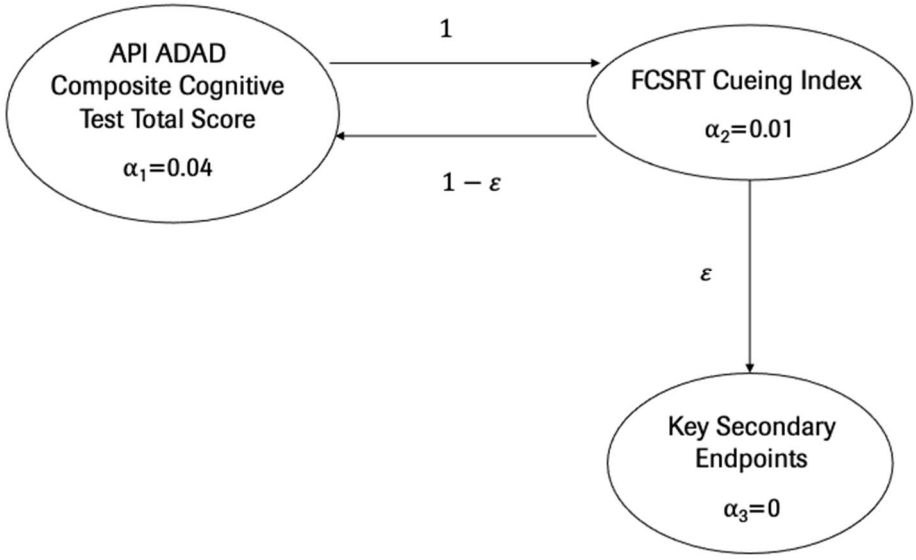
placebo participant change scores/mean of placebo participant change scores) and 100 participants per arm, the study will have at least 80% power to detect a true effect of 30% reduction of the mean decline in the API ADAD Composite Cognitive Test total score in the placebo group using a t-test. Participants in the non-carrier cohort will all receive placebo and will be included in the study in order to maintain the genotype blind.

3.1 TYPE I ERROR CONTROL

A graph-based (Bretz et al. 2009; Bretz et al. 2011; U.S. Food and Drug Administration [FDA 2017]) testing procedure will be used to ensure the family-wise type I error will be controlled at a two-sided $\alpha = 0.05$. In Figure 3, we illustrate how hypothesis testing of the primary endpoint family will be performed. The key secondary endpoints will be tested hierarchically at $\alpha = 0.05$ ordered by the following, only if both primary endpoints are significant; after the first occurrence of the p-value > 0.05 in the list below, subsequent hypothesis testing will stop.

1. Annualized rate of change in mean fibrillar amyloid accumulation in a composite region (including frontal, temporal, parietal, and cingulate cortices) with a subcortical white matter reference region using florbetapir PET
2. Time to progression from preclinical AD to MCI due to AD or from preclinical AD to dementia due to AD
3. Annualized rate of change in the Clinical Dementia Rating (CDR) Scale-Sum of Boxes
4. Time to progression to non-zero in CDR global score
5. Annualized rate of change in a measure of overall neurocognitive functioning: Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) total score

Figure 3 Graph-Based Testing Procedure for Type I Error Control



API = Alzheimer's Prevention Initiative; ADAD = autosomal-dominant Alzheimer's disease; FCSRT = Free and Cued Selective Reminding Task.

The arrows in Figure 3 denote the direction of α propagation; the notation ε indicates a positive real number close to zero, which in our case indicates the potential to pass α from the FCSRT Cueing Index to key secondary endpoints but only if it is not necessary to pass α from the FCSRT Cueing Index to the API ADAD Composite Cognitive Test total score. In the following, we illustrate how the procedure works in a few examples.

- **Example 1**: p-value from testing API ADAD Composite Cognitive Test total score hypothesis = 0.045; p-value from testing FCSRT Cueing Index hypothesis = 0.005. Our graph-based approach will declare statistical significance on both primary endpoints and $\alpha = 0.05$ is passed to key secondary endpoints being tested in our specified order.
- **Example 2**: p-value from testing API ADAD Composite Cognitive Test total score hypothesis = 0.06; p-value from testing FCSRT Cueing Index hypothesis = 0.005. Our graph-based approach will declare statistical significance only on the FCSRT Cueing Index and no α is passed to key secondary endpoints.
- **Example 3:** p-value from testing API ADAD Composite Cognitive Test total score hypothesis = 0.03; p-value from testing FCSRT Cueing Index hypothesis = 0.03. Our graph-based approach will declare statistical significance on both primary endpoints and $\alpha = 0.05$ is passed to key secondary endpoints being tested in our specified order.
- **Example 4:** p-value from testing API ADAD Composite Cognitive Test total score hypothesis = 0.03; p-value from testing FCSRT Cueing Index hypothesis = 0.06. Our graph-based approach will declare statistical significance only on the API ADAD

Composite Cognitive Test total score and no α is passed to key secondary endpoints.

Before study unblinding, the Sponsor performed the blinded power (BP) analysis on the original primary endpoint (API ADAD Composite Cognitive Test total score) and the secondary clinical endpoints using data from all participants regardless of carrier status. The BP analyses assumed the same crenezumab treatment effect across all endpoints considered. The Sponsor concluded that the FCSRT Cueing Index was expected to have a CV which meets the statistical power target specified in the protocol without speculation on the treatment effect based on the blinded data. Therefore, the Sponsor decided to add the FCSRT Cueing Index to the primary endpoint family due to its expected sensitivity to change over time given its clinical relevance as a sensitive measure of episodic memory decline, a hallmark of emerging symptomatic AD, and the most commonly observed, early, and consistent neuropsychological marker of AD (Tounsi et al. 1999; Bateman et al. 2012; Caselli et al. 2020). Appendix 2 describes details of the BP analysis along with *mathematical proof* of type I error control.

4. ANALYSIS SETS

The following analysis populations are defined for Study Period A:

Population	Definition
ITT	All randomized participants, whether or not the participant received the assigned treatment; participants are grouped according to the treatment assignment at randomization × mutation status.
mITT	All randomized participants who received at least 1 dose of study drug; participants are grouped according to the treatment assignment at randomization; all primary and secondary efficacy analyses are based on the mITT population among the mutation carriers.
Safety-evaluable	All randomized participants who received at least 1 dose of study drug; participants are grouped according to the actual treatment received × mutation status; mutation carriers randomized to placebo and mutation non-carriers randomized to placebo who received more than 2 doses of crenezumab will be grouped in the crenezumab arm; all participants who have received no more than 2 doses of crenezumab will be further grouped together in a placebo arm, regardless of the mutation status.
Pharmacokinetic-evaluable	All safety-evaluable participants with at least 1 plasma sample, provided sufficient dosing information (dose and dosing time) is available. Participants will be grouped according to the actual treatment received; participants randomized to placebo who received more than 2 doses of crenezumab will be grouped in the crenezumab arm.

Population	Definition
Tau-ITT	Tau substudy ITT population. All participants randomized in Study GN28352 and consented to the tau PET substudy, whether or not the participant received the assigned treatment; participants are grouped according to the treatment assignment at randomization × mutation status.
Tau-mITT	Tau substudy modified ITT population. All participants randomized in Study GN28352 and consented to the tau PET substudy who received at least 1 dose of study drug; participants are grouped according to the treatment assignment at randomization; analyses on change in tau burden are based on Tau-mITT population among the mutation carriers.
Tau-safety-evaluable	All participants randomized in Study GN28352 and consented to the tau PET substudy who received at least 1 dose of study drug; participants are grouped according to the actual treatment received × mutation status; mutation carriers randomized to placebo and mutation non-carriers randomized to placebo who received more than 2 doses of crenezumab will be grouped in the crenezumab arm; all participants who have received no more than 2 doses of crenezumab will be further grouped together in a placebo arm, regardless of the mutation status.

ITT = intent-to-treat; mITT = modified intent-to-treat; PET = positron emission tomography.

5. STATISTICAL ANALYSES

5.1 GENERAL CONSIDERATIONS

Analyses for Study Period A includes the following:

- A primary analysis *(initial readout)* will occur after all participants have completed the Week 260 assessments
- *During Study Period B, safety analyses from Study Period A will be refreshed (due to the additional 16-week safety follow-up data accrued in Period A after the primary analysis) and reported in the final Clinical Study Report (CSR); no other analyses (e.g., primary and secondary endpoints and PK/PD endpoints) will be refreshed*

The analyses for Study Period B or the OLE study are described in Section 5.10.

All efficacy analyses will be performed on the modified intent-to-treat (mITT) population, unless otherwise specified. Participants will be analyzed according to the treatment assigned at randomization by IxRS.

All safety analyses will be performed in the safety-evaluable population, unless otherwise specified.

All data, including data collected after Week 260, will be used for the primary analysis.

5.2 PARTICIPANT DISPOSITION

The analysis of participant disposition will be based on the intent-to-treat (ITT) population. Reasons for early discontinuation of treatment or early termination from the study will be listed and summarized by treatment group X mutation status.

5.3 PRIMARY ENDPOINTS ANALYSIS

The primary endpoint analysis from Study Period A will be based on the mITT population. The analysis will be adjusted by the factors used in randomization: age, education, apolipoprotein E (APOE) 4 carrier status (carrier vs. non-carrier), and baseline CDR zero versus CDR non-zero. The stratification factors as recorded in IxRS (see protocol for details) will be used.

5.3.1 Definition of Primary Endpoints

The primary endpoint family is 1) the annualized rate of change in the API ADAD Composite Cognitive Test total score and 2) the annualized rate of change in the FCSRT Cueing Index.

The API ADAD Composite Cognitive Test total score is computed from the following 5 cognitive test scores:

- Consortium to Establish a Registry for AD (CERAD) Word List: Recall (Rosen et al. 1984; Morris et al. 1989; Mohs et al. 1997)
- Multilingual Naming Test (Gollan et al. 2012)
- Mini-Mental State Examination (MMSE) for orientation to time (Folstein et al. 1975)
- CERAD Constructional Praxis (a measure of visuospatial ability; Morris et al. 1989)
- Raven's Progressive Matrices (a measure of nonverbal fluid reasoning and visuospatial abilities) Set A (Raven 1976)

Specifically, the ADAD-API Cognitive Composite Test total score is calculated as [(Multilingual Naming Test Score/15) + (MMSE Score/5) + (Raven's Progressive Matrices Score/12) + (CERAD Word List Recall Score/10) + (CERAD Constructional Praxis Score/11)] × 20.

If at least one of the primary analyses is statistically significant, then the trial will be deemed positive. Type I error will be controlled using a graph-based procedure (FDA 2017; see Section 3.1).

5.3.2 Estimands

Primary endpoint estimands are defined on the basis of the following attributes.

- Population: *PSEN1* E280A mutation carriers as defined by the inclusion/exclusion criteria specified in the protocol (mITT)
- Treatments: crenezumab versus placebo either by SC administration or by IV administration

- Variables: 1) annualized rate of change in the API ADAD Composite Cognitive Test total score and 2) annualized rate of change in the FCSRT Cueing Index
- Intercurrent events:
  - Events that lead to treatment withdrawal due to an adverse event. A treatment strategy will be applied where all observed values will be used regardless of the occurrence of the intercurrent event.
  - Death not due to AD. A hypothetical strategy will be applied where all values will be censored after the occurrence of the intercurrent event.
  - Death due to AD. All values on the primary endpoint after this event will be imputed as 0, the worst score possible
  - Pregnancy that leads to treatment withdrawal. A treatment strategy will be applied where all observed values will be used regardless of the occurrence of the intercurrent event.
  - Resumption of dosing after treatment discontinuation post-protocol amendment, version 8. A treatment policy strategy will be applied where all observed values will be used regardless of the occurrence of the intercurrent event.
  - Dose change from SC to IV administration. A treatment policy strategy will be applied where all observed values will be used regardless of the occurrence of the intercurrent event.
  - Events that lead to treatment interruption or withdrawal due to coronavirus disease 2019 (COVID-19) diagnosis. A hypothetical strategy will be applied where all primary efficacy data will be censored between COVID-19 diagnosis and recovery (if recovery date is missing, then all primary efficacy data after COVID-19 diagnosis date will be censored).
  - *Concurrent use of protocol-defined prohibited medications that can potentially impact cognition (e.g., anti-psychotics, anticonvulsants, benzodiazepines, etc.). A treatment policy strategy will be applied where all observed values will be used regardless of the occurrence of the intercurrent event.*
- Population-level summary: difference in annualized rate of change between the mutation carrier crenezumab arm and mutation carrier placebo arm

5.3.3 Main Analytical Approach for Primary Endpoint Family

The primary analysis on both primary endpoints will be performed using a RCRM (Richeldi et al. 2014; Flaherty et al. 2019; Hu et al. 2021) with a missing at random (MAR) assumption. This model will include all observations, including baseline scores as the response variable. The RCRM will include the randomization stratification factors described in Section 5.3 as fixed effects with an intercept, and will include the categorical treatment groups as interaction terms with time*; stratification factor(s) may be removed from the model if there is risk of non-convergence or sparse strata (in case*

of non-convergence, stratification factors will be removed in the order of the smallest number of observations in any cell caused by that stratification factor).
Participant-level random intercept and random slope are also included in the model, where an unstructured covariance between the random intercept and random slope is assumed. The RCRM is constructed as follows:

$$Y_{ij} = \alpha + (\text{Randomization Factors})_i + u_{0i} + (\beta + \beta_x I_{x,i} + u_{1i}) t_{ij} + \varepsilon_{ij}$$

where Y_{ij} is the score for participant i at visit j; t_{ij} is the duration in the unit of year after first drug intake for participant i at visit j; $I_{x,i} = 1$ if participant i in a *PSEN1* E280A mutation carrier is in the crenezumab mITT group, and $I_{x,i} = 0$ if participant i is a *PSEN1* E280A mutation carrier is in the placebo mITT group; u_{0i} and u_{1i} are random intercept and random slope, respectively, for the i^{th} participant; ε_{ij} is the random error for the i^{th} participant at visit j. The crenezumab treatment effect on the rate of change is quantified by β_x; the 95% CI and p-value for crenezumab treatment effect will be presented.
The crenezumab treatment effect compared with the placebo group on the relative reduction (RR) scale is calculated as $-\beta_x/\beta \times 100\%$; its estimate will also be presented; its 95% CI using bootstrap approach will also be provided.

5.3.4 Sensitivity Analyses for Primary Endpoints

The following sensitivity analyses for the primary endpoints are planned.

- To relax the assumption of the linearity in disease trajectory, a mixed model with repeated measures (MMRM) will be used. The response variable is change from baseline. The effects in the model will include baseline score, randomization factors, treatment group, visit, baseline score-by-visit interaction, and treatment-by-visit interaction. Visits will be treated as the repeated variable within a participant. Participant, treatment, and visit will be treated as class variables. An unstructured variance-covariance structure will be applied to model the within-participant errors; in the case of non-convergence, compound symmetry will be used. If non-convergence still exists, data from Year 8 (longest duration in the trial based on enrollment) will be removed, likely caused by a small number of datapoints at Year 8 due to the common-close study design. If non-convergence still exists, visit data will be removed in reverse chronological order until a convergence (either by unstructured or compound symmetry assumption, with unstructured always being tested first) is reached.

  For instances where the score is assessed outside of the visit window described in the protocol, it will be mapped to the closest scheduled visit as described in the schedule of activities in Appendices A–1 through A–4 of the protocol. Baseline score is defined as the last non-missing score on or before the first study drug exposure. After the mapping, if multiple scores exist at a visit, the score closest to the scheduled visit date is chosen in the MMRM analysis.

  The difference between the crenezumab group and the placebo group in the least squares (LS) mean change from baseline will be estimated at each timepoint. The 95% CIs and p-values for LS mean treatment difference will be presented.

The MMRM LS mean plot with standard errors (SEs) will be made to visually examine the linearity of the disease trajectory.

- To quantitatively examine the impact of the linearity assumption and the assumption of the variance decomposition in the RCRM, the robust variance estimator (Huber 1967; White 1982) will be used in the RCRM to re-calculate the SEs and; therefore, CIs and p-values in the primary analysis model
- Analysis assuming participants who *discontinue* from *crenezumab treatment* will lose the treatment effect instantaneously.

  Immediately after a participant *discontinues in the first occurrence* from the *crenezumab treatment*, the population-level annualized rate of change will be assumed to be the same as the population-level annualized rate of change in the placebo arm. This approach will likely yield a conservative estimate of the treatment effect. In particular, the model is written as:

$$Y_{ij} = \alpha + (\text{Randomization Factors})_i + u_{0i} + (\beta + u_{1i})t_{ij} + \left(\beta_x \left(t_{ij} I_{t_{ij} \le t_{i,discon}} + t_{DROP} I_{t_{ij} > t_{i,discon}}\right)\right) I_{x,i} + \epsilon_{ij}$$

  where $t_{i,DROP}$ represents the time of *the first occurrence of discontinuation from the crenezumab treatment* for participant i, and $I_{t_{ij} \le t_{i,discon}}$ and $I_{t_{ij} > t_{i,discon}}$ are indicator variables similarly defined previously in this document. If a participant has not *discontinued from the crenezumab treatment* during Study Period A, then $t_{i,discon}$ is defined as an arbitrarily large number (e.g., 100 with "year" as the unit).
  The 95% CIs and p-values for estimates in β_x will be presented.

- Analysis using control-based mean imputation
  The underlying assumption of this sensitivity analysis is that the missing value mean from the crenezumab arm is the same or worse than the estimated overall mean from the placebo arm. Consequently, this approach yields a conservative treatment effect estimate in crenezumab. The details of this approach can be found in Mehrotra et al. (2017). The model setup and more statistical details can be found in Appendix 3.
- Analysis using hypothetical strategy for all intercurrent events, except for the dosing regimen change. The intercurrent events and the handling strategy in this analysis is defined as follows:
  - Events that lead to treatment withdrawal due to an adverse event.
    A hypothetical strategy where all values will be censored after the first occurrence of the intercurrent event.
  - Death: A hypothetical strategy where all values will be censored after the occurrence of the intercurrent event.
  - Pregnancy that leads to treatment withdrawal. A hypothetical strategy will be applied where all values will be censored after the occurrence of the intercurrent event.

- Resumption of dosing after treatment discontinuation post-protocol amendment, v8. A hypothetical strategy where all values will be censored after 1 day before the first occurrence of the intercurrent event.
- Dose change from SC to IV administration. A treatment policy strategy will be applied where all observed values will be used regardless of the occurrence of the intercurrent event.
- Events that lead to treatment interruption or withdrawal due to COVID-19 diagnosis. A hypothetical strategy where all values will be censored after the first occurrence of the intercurrent event.
- *Concurrent use of protocol-defined prohibited medications that can potentially impact cognition (e.g., anti-psychotics, anticonvulsants, benzodiazepines, etc.). A hypothetical strategy will be applied where all values will be censored from the start date of the occurrence of the intercurrent event to 2 days after the occurrence of the intercurrent event.*

5.3.5 Supplementary Analyses for Primary Endpoints

The following supplementary analyses will be performed for the primary endpoints to provide further understanding of the treatment effect.

- Analysis investigating efficacy in SC crenezumab versus efficacy in IV crenezumab:

  In this analysis, the treatment effect will be assumed to be different between the 2 administration periods. This analysis intends to gain insight on the impact of dose on the treatment effect. Since the IV dose is approximately 4–fold higher than the SC dose, we can use the administration method as a proxy for the dose. In this analysis, the change in administration method is assumed to change the population-level outcome trajectory for participants in the placebo group. For participants in the crenezumab group, the population-level outcome trajectory will also change once an administration method occurs. Furthermore, the switch can only go from SC to IV. In particular, the statistical model is constructed as:

$$\begin{aligned} Y_{ij} = \alpha + (\text{Randomization Factors})_i + u_{0i} + (\beta + u_{1i})t_{ij} + \beta_p(t_{ij} \\ - t_{i,s})\big(1 - I_{x,i}\big)t_{ij}I_{t_{ij}>t_{i,s}} \\ + \Big(\big(\beta_{SC,x}t_{ij}\big)I_{t_{ij}\le t_{i,s}} + \big(\beta_{IV,x}(t_{ij} - t_{i,s}) + \beta_{SC,x}t_{i,s}\big)I_{t_{ij}>t_{i,s}}\Big)I_{x,i} + \epsilon_{ij}. \end{aligned}$$

  In this model, $\boldsymbol{t_{i,s}}$ denotes the switching time for participant i after baseline; if a participant never switched in Study Period A of the study, then $\boldsymbol{t_{i,s}}$ is defined as an arbitrarily large number (e.g., 100 with "year" as the unit). The parameter $\boldsymbol{\beta_p}$ denotes the effect of the change in administration method for participants in the placebo group. The variables $\boldsymbol{I_{t_{ij}\le t_{i,s}}}$ and $\boldsymbol{I_{t_{ij}>t_{i,s}}}$ are indicator variables similarly defined as $\boldsymbol{I_{x,i}}$. The annualized treatment benefit of crenezumab by SC is quantified by $\boldsymbol{\beta_{SC,x}}$; the annualized treatment benefit of crenezumab by IV is quantified by $\boldsymbol{\beta_{IV,x}}$. For model simplicity, up to 1 switch is assumed. The 95% CIs and p-values for estimates in $\boldsymbol{\beta_{SC,x}}$ and $\boldsymbol{\beta_{IV,x}}$ will be presented. The point estimates for the SC and IV treatment effect in the RR scale will also be presented.

- Analysis on each component of the API ADAD Composite Cognitive Test total score

  To understand the treatment effect on each component of one of the primary endpoints (API ADAD Composite Cognitive Test total score), the same primary analysis model will be run on each component. The 95% CIs and p-values for the treatment effect estimator will be presented. A forest plot will be made.

- To explore whether decline accelerates or decelerates, a quadratic RCRM where leading t^2 terms (treatment and time^2 interaction terms and random quadratic term) are added to the primary analysis model. The covariance matrix of $\{u_{0i}, u_{1i}, u_{2i}\}$ is assumed to be unstructured and all random effect terms are assumed to be independent of the pure error term. Specifically, this model is written as:

$$Y_{ij} = \alpha + (\text{Randomization Factors})_i + u_{0i} + (\beta + \beta_x I_{x,i} + u_{1i})t_{ij} + (\theta + \theta_x I_{x,i} + u_{2i})t_{ij}^2 + \epsilon_{ij}.$$

All statistical inferences as in the primary analysis model along with the point estimates of θ_x and 95% CI will be presented.

5.3.5.1 Subgroup Analyses for Primary Endpoints

The following subgroups will be analyzed with respect to the primary endpoints using the same primary analysis. Forest plots will be presented to summarize the results.
The subgroup categories may be combined if there is not enough representation of a specific subpopulation or if the statistical model fails to converge.

- Age ($\leq$38 vs. $>$38)
- Education ($<$9 years vs. $\geq$9 years)
- APOE4 carrier status (carrier vs. non-carrier)
- Baseline CDR zero vs. CDR non-zero
- Amyloid status at baseline (amyloid positive vs. amyloid negative) using a standardized uptake value ratio (SUVr) threshold *(details of the threshold will be pre*specified *in the biomarker analysis plan)*

5.4 SECONDARY ENDPOINTS ANALYSES

The continuous secondary efficacy endpoints will be analyzed using the same primary statistical model as for the primary endpoints. All time-to-event endpoints will be analyzed with a stratified log-rank test using the primary analysis stratification factors. Within each time-to-event endpoint, the hazard ratio (HR) for recurrence will be estimated using a stratified Cox proportional hazards model, and the 95% CI for the HR will be provided. Stratification factor(s) may be removed from the stratified analyses if there is risk of non-convergence or sparse strata *(in the case of non-convergence, stratification factors will be removed in the order of the smallest number of observations in any cell caused by that stratification factor).*

5.4.1 Key/Confirmatory Secondary Endpoints

The key secondary endpoints are:

- *Annualized rate of change in mean cerebral fibrillar amyloid accumulation using florbetapir PET from a predefined ROI*
  - *SUVrs will be derived from florbetapir PET images using* a composite region *of interest* (including frontal, temporal, parietal, and cingulate cortices) with a subcortical white matter reference region
- Time to progression from preclinical AD to MCI due to AD or from preclinical AD to dementia due to AD
- Annualized rate of change in the CDR Scale-Sum of Boxes
- Time to progression to non-zero in CDR global score
- Annualized rate of change in a measure of overall neurocognitive functioning: RBANS total score

Time to progression is calculated from the time at baseline, which is the last record on or prior to the first baseline reference date; the baseline reference date is either the first exposure date or the randomization date if the first exposure date is missing.
A gatekeeping strategy will be used for testing hypotheses of the key secondary outcomes using the graph-based type I error approach described in Section 3.1, the key secondary outcomes will be tested only if both the primary endpoints are statistically significant. All the tests will be done based on a two-sided $\alpha = 0.05$. The tests in the key secondary outcomes will be done in sequential order specified in the following:

1. Annualized rate of change in mean fibrillar amyloid accumulation in a composite region (including frontal, temporal, parietal, and cingulate cortices) with a subcortical white matter reference region using florbetapir PET
2. Log-rank analysis of time to progression from preclinical AD to MCI due to AD or from preclinical AD to dementia due to AD
3. RCRM analysis of CDR Scale-Sum of Boxes
4. Log-rank analysis of time to progression to non-zero in CDR global score
5. RCRM analysis of RBANS total score

At initial study readout if the study is positive (at least one of the primary endpoints is statistically significant), the Sponsor will analyze all available data from the key secondary biomarker endpoint: annualized rate of change in mean cerebral fibrillar amyloid accumulation in a composite region (including frontal, temporal, parietal, and cingulate cortices) with a subcortical white matter reference region using florbetapir PET. If the study is negative (neither *of the* primary endpoints are statistically significant), the Sponsor will analyze at least the baseline and last-available (in time) data from the florbetapir PET endpoint at the initial study readout, which is at the *primary analysis.*

5.4.2 Supportive Secondary Endpoints

The supportive secondary biomarker endpoints include:

- *Annualized rate of change in regional cerebral metabolic rate of glucose (CMRgl) using FDG-PET in a predefined ROI*
  - Annualized rate of change in regional (CMRgl) *will be derived from FDG PET images using an empirically predefined statistical ROI (sROI) known to be preferentially affected by CMRgl decline in persons with AD (Chen et al. 2010; Van Dyck et al. 2019)*
- *Annualized rate of change in volumetric measurements using MRI*
  - *Regions will include* the whole brain, bilateral hippocampus, and *bilateral* ventricles
- Annualized rate of change in CSF tTau and pTau biomarkers

Similar to the approach from the key secondary biomarker efficacy endpoint, the Sponsor will analyze all available data on the FDG-PET and volumetric MRI data if the study is positive. Otherwise, the Sponsor will analyze at least the baseline and last-available (in time) data on FDG-PET and volumetric MRI at the initial study readout.

5.5 EXPLORATORY ENDPOINTS ANALYSIS

Exploratory outcome measures are listed below. The continuous endpoints will be analyzed using the RCRM from the primary analysis and/or using an MMRM; the correlation analysis among outcomes will be modeled using a Spearman correlation estimation. The "mutation non-carriers treated with placebo" in this section is defined as all randomized mutation non-carriers that received at least 1 dose of study drug.

5.5.1 Clinical

Changes from baseline over time in the following cognitive measures:

- Trail Making Test (Armitage 1946)
- MMSE (Folstein et al. 1975)
- RBANS Index Scores (Randolph 1998)
- Scores from each of the components of the API ADAD Composite Cognitive Test Battery
- Preclinical Alzheimer Cognitive Composite (PACC; Donohue et al. 2014) modified for the API ADAD trial; PACC is calculated as the standardized sum of MMSE, RBANS story recall, RBANS coding, and FCSRT total score (sum of 3 free and cued recall trials)
- Clinical endpoints not examined in secondary outcome measures
- Changes in the Neuropsychiatric Inventory (NPI) total score, items, and factors (Cummings et al. 1994; 1997)

- Changes in the Geriatric Depression Scale (GDS) total score (Sheikh and Yesavage 1986)
- Changes in Functional Assessment Staging of Alzheimer's Disease (FAST) total score (Sclan and Reisberg 1992)
- Changes in Subjective Memory Checklist (Acosta-Baena et al. 2011)

5.5.2 Fluid Biomarkers

- CSF levels of $A\beta_{1-42}$ and $A\beta_{1-40}$
- Changes in other blood and CSF measures such as neurofilament light (NfL) and plasma tau markers

5.5.3 Imaging Biomarkers

- Analysis of other *regional or global measures* not selected in secondary outcome measures. *See the biomarker analysis plan for more details.*
- Other imaging outcome measures and analytic methods, alone, or in conjunction with other imaging modalities not explored in secondary outcome measures
- Change in tau burden over time, as measured by *Genentech Tau Probe 1 (*GTP1*)* PET in an optional substudy (GN28352-1/BN40199). See Section 5.8 for analysis details.

5.5.4 Other Endpoints

See the biomarker analysis plan for more details on the following endpoints:

- Changes in primary, secondary, and exploratory outcomes in mutation non-carriers treated with placebo
- Comparisons of clinical and biomarker outcomes between carriers treated with placebo and non-carriers
- Changes in primary, secondary, and exploratory outcomes in carriers and non-carriers as functions of the APOE genotype and other genetic variations
- Short-term changes in imaging measures as functions of initiation (e.g., baseline to 12 weeks)
- Analyses of outcome measures in relation to one another and in relation to baseline characteristics

5.6 SAFETY ANALYSES

All safety analyses of Study Period A will be performed in the safety-evaluable population (see Section 4 for definition).

As the study participants are considered treatment-exposed during the 16-week (approximately 5 half-lives) follow-up period, adverse events reported within 16 weeks after the last dose of study drug administered in Study Period A will be included in the safety analyses from Study Period A (contrary to the description in the protocol stating they will be summarized separately). At the end of Study Period B, safety analyses on

Study Period A will be refreshed *(due to the additional 16-week safety follow-up data accrued in Period A after the primary analysis)* and reported in the final CSR.

Adverse events reported within 16 weeks after the last dose of study drug administered in Study Period B will be included in the safety analyses on Study Period B.
Safety analyses of Study Period B and/or of the potential OLE data are summarized in Section 5.10.2.

Safety will be assessed through summaries of exposure to study treatment, adverse events, changes in laboratory test results, MRI findings, changes in vital signs and ECGs, and changes in suicidality assessment.

Study treatment exposure (such as treatment duration, total dose received, and dose modifications) will be summarized with descriptive statistics.

All verbatim adverse event terms will be mapped to MedDRA thesaurus terms, and adverse event severity will be graded according to National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE), v4.0. All adverse events, serious adverse events, adverse events leading to death, adverse events of special interest, and adverse events leading to study treatment discontinuation that occur on or after the first dose of study treatment (i.e., treatment-emergent adverse events) will be summarized. All adverse events will be summarized by severity and by relationship to study drug. For events of varying severity, the highest grade will be used in the summaries.

Relevant laboratory, vital signs (pulse rate, respiratory rate, blood pressure, pulse oximetry, and temperature), and ECG data will be displayed by time, with grades identified where appropriate. Additionally, a shift table of selected laboratory tests will be used to summarize the baseline and maximum post-baseline severity grade.
Changes in vital signs and ECGs will be summarized. Proportion of participants with suicidal ideation or behavior, as reported in the suicidality assessment, will be analyzed by treatment groups.

5.6.1 Extent of Exposure

Study treatment exposure (such as treatment duration, total dose received, and dose modifications) will be summarized with descriptive statistics.

Exposure to study drug (number of study drug administrations and duration of treatment), regardless of SC or IV administration, will be summarized by treatment group in the safety-evaluable population.

Exposure will also be summarized by the administration route (SC vs. IV).

If a participant did not resume treatment after a treatment discontinuation, then the duration of treatment is the time from first study drug to the earlier of:

- Date of treatment discontinuation or date of study treatment completion in Study Period A
- The analysis cutoff date

Otherwise, if a participant resumed treatment after a treatment discontinuation, then the duration is the summation of duration of each treatment period, where the duration of the first treatment period is defined as the time from first study drug to the date of first treatment discontinuation and the duration of the last treatment period is defined as the time from the last treatment resumption date to the earlier of:

- Date of treatment discontinuation or date of study treatment completion in Study Period A
- The analysis cutoff date

The duration of the rest of the treatment periods is defined as the time from the treatment resumption date to the date of treatment discontinuation.

See Section 5.7.4 for PK/PD analyses.

5.6.2 Adverse Events

All treatment-emergent adverse events, serious adverse events, adverse events leading to death, protocol-specified adverse events of special interest (including amyloid-related imaging abnormalities–edema/effusion [ARIA-E], amyloid-related imaging abnormalities–hemosiderin deposition [ARIA-H], cerebral macrohemorrhages, pneumonia, *potential drug-induced liver injury, and suspected transmission of infectious agents by the investigational product*) and adverse events leading to study treatment discontinuation will be summarized by MedDRA Preferred Term, appropriate thesaurus level by treatment arm. Adverse events will be summarized by severity and by relationship to study drug. For events of varying severity, the highest grade will be used in the summaries.

In addition, treatment-emergent injection reactions, infusion-related reactions and COVID-19 infections will be summarized by treatment group.

Only treatment-emergent adverse events will be included in the analyses, defined as any adverse event reported during or after the first dose of study drug. Adverse events with a missing onset date will be considered to be treatment-emergent. Adverse events with a partially missing onset date will also be included as treatment-emergent if the month (if it was recorded) and the year occur on or later than the month and year of the study treatment start date.

A separate listing of adverse events reported during the screening period (before the first dose of study drug) will also be produced.

5.6.3 Clinical Laboratory Data

Clinical laboratory data (serum chemistry, hematology evaluations, including CBC with differential and platelet counts, and urinalysis values) will be summarized over time by descriptive statistics by treatment group, with grades identified where appropriate. Additionally, a shift table of selected laboratory tests will be used to summarize the baseline and maximum post-baseline severity grade.

5.6.4 MRI Evaluations

Neuroradiologic evaluation with respect to the occurrence of cerebral vasogenic edema (ARIA-E); superficial siderosis of the CNS or cerebral microhemorrhages (ARIA-H); or of cerebral macrohemorrhages will be performed during the *Study Period A* treatment period and will be summarized by treatment group using descriptive statistics.

5.6.5 Vital Signs

Vital signs (pulse rate, blood pressure, body temperature, and respiratory rate) will be displayed over time and changes from baseline will be summarized by treatment group.

5.6.6 Electrocardiograms

Electrocardiogram data will be displayed over time and changes from baseline for relevant ECG intervals will be summarized by treatment group *up to Week 17*.

5.6.7 Suicidal Ideation and Behavior

The proportion of participants with suicidal ideation or behavior, as reported in the suicidality assessment, will be analyzed by treatment group.

5.6.8 Anti-Drug Antibodies

The number and percentage of participants with confirmed positive anti-drug antibodies (ADAs) will be reported for each treatment group at baseline (prevalence of ADAs in the crenezumab and placebo groups) and after treatment with crenezumab (treatment-emergent ADAs). *At the primary analysis, the Sponsor will have only analyzed partial ADA data. If the study is negative and results from the partial ADA data are consistent with the extensive ADA data collected in previous Phase I–III studies, no additional ADA analyses are planned. The rationale is that ADA data from approximately 1900 samples that have already been tested is sufficient to confirm the low immunogenicity risk of crenezumab in this population without impacting the interpretation of the overall study results. However, if the study is positive, (at least one of the primary endpoints is statistically significant) the Sponsor will then analyze the complete ADA dataset in the final analysis and report it.*

5.7 OTHER ANALYSES

5.7.1 Summaries of Conduct of Study

In Study Period A, the number of participants who enroll, discontinue from treatment, discontinue from study, and complete the study will be tabulated by treatment group. Reasons for early discontinuation of treatment or early termination from the study will be listed and summarized by treatment group. Any eligibility criteria and other major protocol deviations will also be summarized by treatment group. The summaries of conduct of study will be based on the ITT population; randomized non-carriers will also be grouped and presented in this section.

5.7.2 Summaries of Treatment Group Comparability/Demographics and Baseline Characteristics

Demographic and baseline characteristics such as age, sex, race, APOE4 status, and baseline cognitive scores will be summarized for the ITT population by treatment group using descriptive statistics; randomized non-carriers will also be grouped and presented in this section. Baseline is defined as the last-available measurement obtained on or prior to randomization.

5.7.3 Summaries COVID-19 Impact on the Trial

The study was ongoing during the COVID-19 pandemic. Consequently, to monitor the potential impact of the pandemic on the trial, we will provide a specific set of descriptive analyses related to COVID-19, including:

- COVID-19 adverse event
- COVID-19-related protocol deviations
- Missed doses due to COVID-19
- Study discontinuations due to COVID-19
- Remote scale administrations
- Site actions and site closures

5.7.4 Pharmacokinetic and Pharmacodynamic Analyses

Mean serum and CSF crenezumab concentration versus time data will be tabulated and plotted, as appropriate. In serum, trough plasma concentration ($C_{trough,\ ss}$) will be tabulated and summarized (e.g., mean, standard deviation, minimum, and maximum). Additional PK analyses, such as evaluating relationships between crenezumab concentrations, PD biomarkers, efficacy and safety endpoints, may be conducted as appropriate. Mean plasma total $A\beta_{1\text{-}40}$ and $A\beta_{1\text{-}42}$ concentrations versus time will be tabulated and plotted, as appropriate. The relationship between crenezumab and total $A\beta_{1\text{-}40}$ and $A\beta_{1\text{-}42}$ concentrations will be tabulated and plotted to explore the peripheral PK/PD relationship*, as appropriate*.

The schedule for PK and PD analyses will be independent of the safety and efficacy endpoints.

5.8 ANALYSES IN TAU PET LONGITUDINAL SUBSTUDY BN40199

In this section, we outline the main statistical details in the tau PET longitudinal Substudy BN40199. The analysis populations in this substudy are described in detail in Section 4.

The primary endpoint of the tau PET longitudinal substudy is the annualized rate of change in the tau PET SUVr in the entorhinal cortex (Braak Stage 1), and other ROIs, including whole cortical gray matter, may be explored based on evidence in presymptomatic ADAD populations. The analysis population is based on mITT among the mutation carriers. The primary analysis in the tau PET longitudinal substudy will be performed using an RCRM. This model will include the actual tau PET SUVr as the response variable. The model set up is similar to the RCRM in the primary analysis model in the main study.

The 95% CI and p-value for crenezumab treatment effect will be presented. The estimated crenezumab treatment effect compared with the placebo group on the RR scale will also be presented. *Similar to secondary imaging endpoints (Sections 5.4.1 and 5.4.2)*, at the initial study readout, the Sponsor will analyze at least the *first* and last-available (in time) tau PET data. If the result from the main Study GN28352 is positive (at least one of the primary endpoints is statistically significant), the Sponsor will further analyze all available tau PET data and the p-value will be instead based on the full data.

The MMRM analysis may also be used to explore the treatment benefit over time. Additional descriptive summaries that describe the relationship between tau PET SUVr and biomarker and clinical endpoints are specified in Section 1.1.

Summaries of conduct of the substudy and summaries of demographic and baseline characteristics will be presented in a similar fashion to the main study; the analyses will be based on the ITT population; randomized non-carriers will also be grouped and presented in this section.

The main safety analyses planned for Study Period A (see Section 5.6) of the safety data collected during the tau substudy will also be performed in the tau substudy safety-evaluable population (see Section 4 for definition).

5.9 INTERIM ANALYSES

No interim analyses to be conducted.

5.10 ANALYSES FOR PERIOD B AND OPEN-LABEL EXTENSION STUDY

This section will pre-specify the main analyses planned for Study Period B of this study and the potential OLE *study*.

If the Sponsor decides to terminate the study based on the primary analysis results from Study Period A, all the data and analyses from Study Period B will be considered as exploratory in nature; a final CSR will be written, including these extra data. However, if the Sponsor decides to *extend Study Period B or* initiate *an* OLE *study*, a planned *data snapshot* will occur *2 years after finished the last efficacy visit in Study Period A.*

Three analysis populations are defined for this snapshot:

- Study Period A mITT population: the mITT population defined in Section 4: all randomized participants who received at least 1 dose of study drug; participants are grouped according to the treatment assignment at randomization; all efficacy analyses in this snapshot are based on mITT population among the mutation carriers
- Study Period B safety population: all randomized participants that received at least 1 dose of study drug in Study Period B; participants are grouped according to the actual treatment received; participants on placebo who received more than 2 doses of crenezumab will be grouped in the crenezumab arm
- OLE safety population: all OLE enrolled participants that received at least 1 dose of study drug (crenezumab)

5.10.1 Efficacy Analyses

All efficacy analyses will include data from Study Period A, Study Period B, and the OLE *study*; the analysis population will be based on mITT. *The delayed start efficacy analyses will be triggered 2 years after the last efficacy visit in Study Period A.*
Two types of efficacy analyses are prespecified depending on the endpoint type: delayed-start analyses for continuous endpoints and rank-preserving structural failure time (RPSFT) analyses for time-to-event endpoints.

5.10.1.1 Delayed-Start Analyses

The delayed-start analysis will be used for all continuous efficacy endpoints that will be collected in *Study Period B and* the *potential* OLE *study*. Since *the potential* OLE *will not be published by the finalization of this SAP*, the description in this section is in general terms and does not focus on any specific endpoint. However, statistical testing will be performed only on the API ADAD Composite Cognitive Test total score and the FCSRT Cueing Index. Other continuous efficacy endpoints, including clinical endpoints, fluid biomarkers, and imaging biomarkers, will be considered exploratory and thus, formal statistical testing will not be done.

Figure 4 illustrates the concept of the delayed-start analysis. As shown in the figure, a treatment effect is assumed in Period A of Study GN28352; the disease progression slope of mutation carriers from the Study Period A crenezumab arm is assumed to be the same in Study Period B and in the OLE; the disease progression slope of mutation carriers from the Study Period A placebo group; however, will have a different slope in Study Period B and in the OLE. The statistical model details can be found in

Appendix 3. To control the type I error within the delayed-start analyses, the fixed testing sequence is specified in the following figure. The αlevel in each testing is a two-sided 0.05. If at any stage the null hypothesis H_0 is not rejected, then the formal testing procedure will stop.

Figure 4 Illustration of Delayed-Start Analyses Using Data from the Beginning of Period A of Study GN28352 to the End of *Study Period B or* OLE

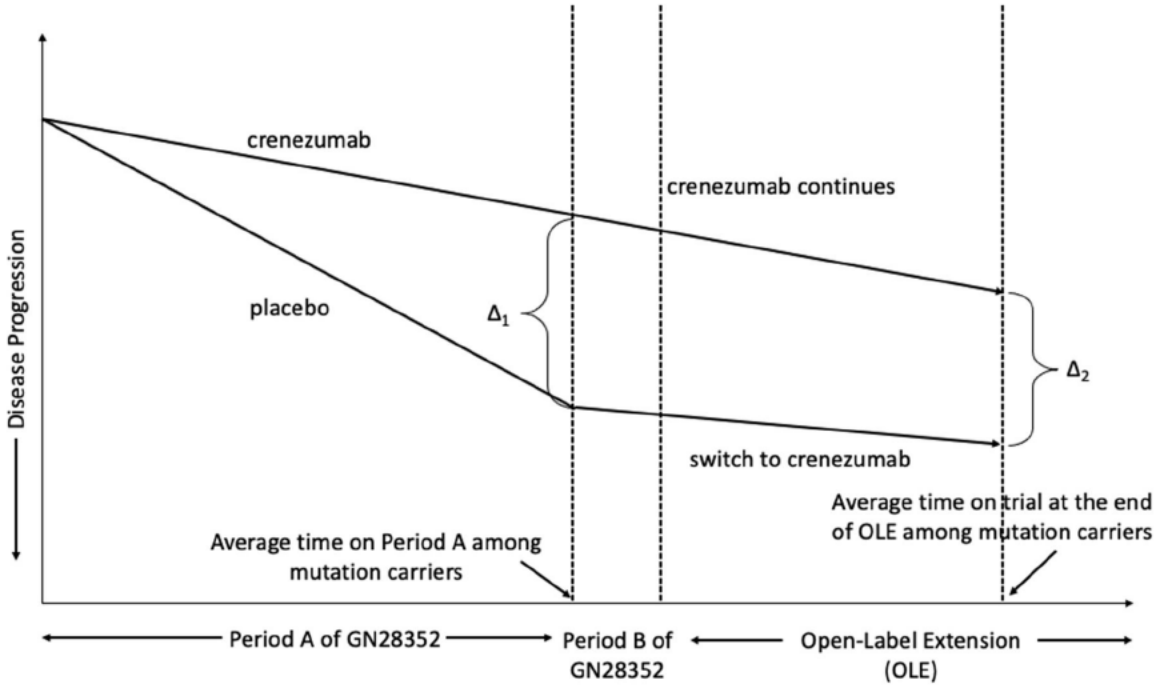
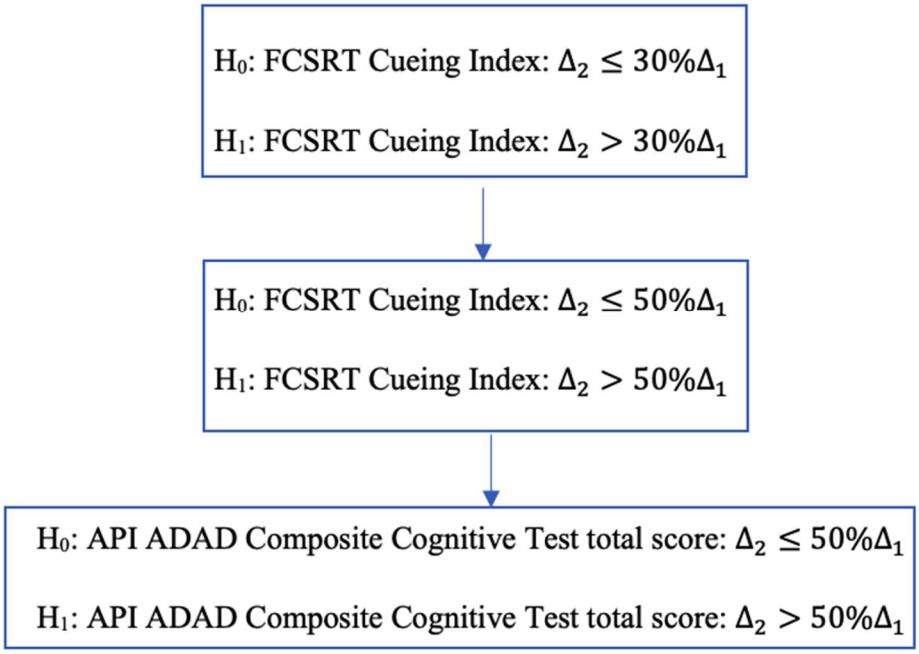


Figure 5 Testing Hierarchy for Delayed-Start Analyses



API = Alzheimer's Prevention Initiative; ADAD = autosomal-dominant Alzheimer's disease; FCSRT = Free and Cued Selective Reminding Task.

Note: For a definition of H0 and H1, please see Section 2.

5.10.1.2 Rank-Preserving Structural Failure Time Analyses

The RPSFT analysis will be used for the 2 time-to-event (TTE) endpoints below, also collected in Period B of Study GN28352 and the OLE. The TTE endpoints will anchor the randomization date from Period A of Study GN28352 as the reference date.

- Time to progression to non-zero in CDR global score
- Time to progression from preclinical AD to MCI due to AD or from preclinical AD to dementia due to AD

The Sponsor will ensure that the TTE endpoints will remain as objective/unbiased as possible in Period B of Study GN28352 and the OLE. *During* Period B of Study GN28352, the Sponsor will be unblinded to the mutation status and treatment assignment when analyzing the Study Period A results. *However,* participants, contract research organization staff, and blinded site staff will remain blinded *to treatment allocation in Study Period A*. Furthermore, an independent adjudication committee will still be in place to adjudicate cases deemed by the site that have progressed from preclinical AD to MCI due to AD or from preclinical AD to dementia due to AD. If the Sponsor initiates the OLE after the results of Study GN28352 are available, we anticipate that *only mutation carriers will be eligible for* the OLE.

The RPSFT method as introduced by Robins and Tsiatis (1991) provides an estimate of the overall event time for participants randomized into placebo group in Study Period A had treatment switching not occurred. With this model, we assume that the crenezumab treatment effect on event time is the same regardless of when crenezumab is given. It estimates overall event time measured from the time of treatment switching by applying an estimate of the benefit of crenezumab (derived iteratively and referred to as the inverse of the acceleration factor). The treatment switching does not happen for participants randomized into the crenezumab group in Study Period A; therefore, their event time can be directly used. The stratified log rank test, where the stratification factors are from Study Period A, will then be used to calculate the two-sided p-value; the HR for recurrence will be estimated using a stratified Cox proportional hazards model, and the 95% CI for the HR will be provided. Some stratification factor(s) may be removed from the stratified analyses if there is risk of non-convergence or sparse strata.

5.10.2 Safety Analyses

If the Sponsor decides *to terminate the study based on the primary analysis results from Study Period A*, then safety data from Study Period B will be reported in the final CSR. The Study Period B safety population is defined in the beginning of Section 5.10.

If the Sponsor decides to *extend Study Period B or* initiate *an* OLE, *then safety data from Study Period B will be reported in the final CSR and* safety data from the OLE will be summarized *in the OLE CSR*. The OLE safety population is defined in the beginning of Section 5.10.

All the safety analyses described in Section 5.6 will be reported.

5.10.3 Other Analyses

Summaries of conduct of study and summaries of treatment group comparability/demographics and baseline characteristics will also be presented. The population is based on either the Study Period B ITT population or the OLE ITT population, depending on whether the Sponsor initiates the OLE. The analyses are similar to what are specified for Study Period A.

Pharmacokinetic and PD analyses will be handled in a similar fashion to Study Period A.

6. SUPPORTING DOCUMENTATION

This section is not applicable, since there is no additional supporting document.

Appendix 1 Changes to Protocol-Planned Analyses

The following changes have been made in this Statistical Analysis Plan (SAP) from the latest Protocol Version 10:

- The definition of the safety population for Study Period B has been updated in the SAP
- Adverse events reported during the 16-week follow-up period in Study Period A will not be summarized separately

Appendix 2 Blinded Power Analysis and its Impact on Type I Error

This section outlines the methodology that the Sponsor used for the blinded power (BP) analysis. We then provide the evidence of family-wise type I error control when a BP analysis is performed. *More details of this methodology can be found in (Hu et al. 2022).*

METHODOLOGY

Assuming a BP analysis is performed on p continuous endpoints: $X^{(1)}, X^{(2)}, \ldots, and\ X^{(p)}$, right before study unblinding. The goal of the BP analysis is to help ensure that the primary analysis meets the statistical power criteria defined in the protocol.
For simplicity and without loss of generality, we also assume the following:

- All endpoints are expected to decrease overtime in the mutation carrier placebo arm
- n/2 participants randomized to the placebo arm and n/2 participants randomized to the mutation carrier experimental arm, with no dropout
- Change from baseline scores at the last visit are calculated from each participant (statistic denoted as $X_i^{(k)}$ for the k[th] endpoint for the i[th] participant), with $X_i^{(k)} \sim N\left(\mu_{pbo}^{(k)}, \sigma^{(k)^2}\right)$ if participant i is from the placebo arm and $X_i^{(k)} \sim N\left(\mu_{exp}^{(k)}, \sigma^{(k)^2}\right)$ if participant i is from the experimental arm
- **The same treatment effect in relative reduction (RR) scale R_a is assumed in the analysis across endpoints** (note that R_a differs from the true underlying treatment effect $R^{(k)} = \frac{\mu_{pbo}^{(k)} - \mu_{exp}^{(k)}}{\mu_{pbo}^{(k)}} * 100\%$)

From Friede and Kieser (2013), the scaled sample variance $\frac{n-1}{\sigma^{(k)^2}} S^{(k)^2} \sim$ Noncentral chi-squared distribution with n-1 degrees of freedom and noncentrality parameter $\frac{n\left(\mu_{pbo}^{(k)} R^{(k)}\right)^2}{4\sigma^{(k)^2}}$ and

$$E\left(S^{(k)^2}\right) = \sigma^{(k)^2} + \frac{n}{4(n-1)}\left(\mu_{pbo}^{(k)} R^{(k)}\right)^2.$$

Since the number of participants from the placebo and experimental arms both consist of ½ of the total population, respectively, then we can construct the following equation.

$$\frac{\hat{\mu}_{pbo}^{(k)} + \hat{\mu}_{pbo}^{(k)}(1 - R_a)}{2} = \overline{X^{(k)}}.$$

From the previous 2 equations, we can estimate the placebo mean decline and standard deviation for endpoint k:

$$\hat{\mu}_{pbo}^{(k)} = \frac{2\overline{X^{(k)}}}{2 - R_a}$$

$$\hat{\sigma}^{(k)^2} = S^{(k)^2} - \frac{n}{4(n-1)}\left(\frac{2R_a\overline{X^{(k)}}}{2-R_a}\right)^2.$$

When we assume the same relative treatment effect R_a across p continuous endpoints, finding the endpoint with the highest statistical power is equivalent to finding the endpoint with the smallest CV in the placebo arm. We have:

$$\widehat{CV}^{(k)} = \frac{\sqrt{\hat{\sigma}^{(k)^2}}}{\hat{\mu}_{pbo}^{(k)}} = \sqrt{\left(\frac{2-R_a}{2}\right)^2 \frac{S^{(k)^2}}{\overline{X^{(k)}}^2} - \frac{nR_a^2}{4(n-1)}}$$
$$= \sqrt{\left(\frac{2-R_a}{2}\right)^2 \widehat{CV}_{pooled}{}^{(k)^2} - \frac{nR_a^2}{4(n-1)}}.$$

We can see that $\widehat{CV}^{(k)}$ is a monotonically increasing function of $\widehat{CV}_{pooled}{}^{(k)^2}$, which implies that endpoints can be evaluated based on their relative $\widehat{CV}_{pooled}{}^{(k)^2} = \frac{S^{(k)^2}}{\left(\overline{X^{(k)}}\right)^2}$ using the blinded data.

The BP analysis just described in the proceeding bears similarities to blinded sample size re-estimation. In blinded sample size re-estimation, the updated sample size depends on the targeted treatment effect and observed pooled sample variance. Likewise, our BP analysis depends on multiple targeted treatment effects and the observed pooled CVs. It is important to note that neither case requires unblinded data.

It has been reported that blinded sample size re-estimation does not lead to substantive type I error inflation (Friede and Kieser 2006; Glimm and Läuter 2013). Next, we prove that analytically, BP analysis controls family-wise error rate (FWER) in *both the strong* the weak sense. We then provide simulations which support FWER control in the strong sense.

TYPE I ERROR CONTROL UNDER BP ANALYSIS

We present the theorem under the assumption that *at least one of the* p endpoints *has* no treatment effect.

Theorem 1

Among the p endpoints $\{X^{(1)}, X^{(2)}, \ldots, X^{(p)}\}$ *in* the *BP analysis, if there is at least one endpoint that has no treatment effect, the family-wise* type I error is controlled under the BP analysis.

Proof: We start by writing out the density function of the endpoint vector for participant i. If we define the endpoint vector $X_i = \left[X_i^{(1)}, X_i^{(2)}, \ldots, X_i^{(p)}\right]^T$ *and if we assume*

the endpoints considered follow a multi-normal distribution, then we have that $X_i \sim N(L_i\mu, \Sigma)$*, where:*

$$L_i = \begin{bmatrix} l_i & 0 & \dots & 0 & 1-l_i & 0 & \dots & 0 \\ 0 & l_i & \dots & 0 & 0 & 1-l_i & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & l_i & 0 & 0 & \dots & 1-l_i \end{bmatrix},$$

$$\mu = \left[\mu_{pbo}^{(1)}, \mu_{pbo}^{(2)}, \dots, \mu_{pbo}^{(p)}, \mu_{exp}^{(1)}, \mu_{exp}^{(2)}, \dots, \mu_{exp}^{(p)}\right]^T,$$

$$\Sigma = \begin{bmatrix} {\sigma^{(1)}}^2 & \dots & \sigma^{(1)(p)} \\ \vdots & \ddots & \vdots \\ \sigma^{(1)(p)} & \dots & {\sigma^{(p)}}^2 \end{bmatrix}.$$

The constant l_i *takes the value of either 0 or 1 (participant i in the experimental group or placebo group). We denote the parameter vector* $\theta = (\mu^T, vec(\Sigma)^T)^T$*. We also denote* $\mu_{pbo} = \left[\mu_{pbo}^{(1)}, \mu_{pbo}^{(2)}, \dots, \mu_{pbo}^{(p)}\right]^T$ *and* $\mu_{exp} = \left[\mu_{exp}^{(1)}, \mu_{exp}^{(2)}, \dots, \mu_{exp}^{(p)}\right]^T$*. The density function of* X_i *is:*

$$f(X_i|\theta, L_i) = \frac{1}{\sqrt{(2\pi)^p|\Sigma|}} exp\left(-\frac{1}{2}(X_i - L_i\mu)^T\Sigma^{-1}(X_i - L_i\mu)\right).$$

We define functions:

$$\eta(\theta) = \begin{pmatrix} \begin{bmatrix} \Sigma^{-1}\mu_{pbo} \\ \Sigma^{-1}\mu_{exp} \end{bmatrix} \\ -\frac{1}{2}\Sigma^{-1} \end{pmatrix},$$

$$T(X_i) = \begin{pmatrix} L_i^T X_i \\ X_i X_i^T \end{pmatrix},$$

$$h(X_i) = (2\pi)^{-\frac{p}{2}},$$

$$A(\theta) = \frac{1}{2}(L_i\mu)^T\Sigma^{-1}L_i\mu + \frac{1}{2}log\,|\Sigma|.$$

The dot product between $\eta(\theta)$ *and* $T(X_i)$ *is calculated as:*

$$\eta(\theta) \cdot T(X_i) = \begin{bmatrix} \Sigma^{-1}\mu_{pbo} \\ \Sigma^{-1}\mu_{exp} \end{bmatrix}^T L_i^T X_i + < \left(-\frac{1}{2}\Sigma^{-1}\right)^T, X_i X_i^T >_F = \mu^T L_i^T \Sigma^{-1} X_i - \frac{1}{2} tr\left(\Sigma^{-1} X_i X_i^T\right),$$

where $<,>_F$ is the Frobenius inner product and "tr(.)" stands for the trance of a matrix. The density function can then be written as:

$$f(X_i|\theta, L_i) = (2\pi)^{-\frac{p}{2}} exp\left(-\frac{1}{2}(X_i - L_i\mu)^T\Sigma^{-1}(X_i - L_i\mu) - \frac{1}{2}log|\Sigma|\right)$$
$$= h(X_i)\, exp\big(\eta(\theta)\cdot T(X_i) - A(\theta)\big).$$

Therefore, the multivariate normal distribution is in the exponential family with the parameter defined as $\theta = (\mu^T, vec(\Sigma)^T)^T$. The statistic $\sum_{i=1}^n T(X_i) = \begin{pmatrix}\sum_{i=1}^n L_i^T X_i \\ \sum_{i=1}^n X_i X_i^T\end{pmatrix}$ is then both complete and sufficient for θ (Theorems 6.2.10 and 6.2.25 from Casella and Berger 2021). The final t-test statistic t_k is clearly free of the parameter vector θ; therefore, t_k is an ancillary statistic to θ. Based on Basu's theorem, the testing statistic t_k is independent of $\sum_{i=1}^n T(X_i)$. Since the statistics $\frac{S^{(k)^2}}{\left(\overline{X^{(k)}}\right)^2}$ $(k = 1,2,\dots,p)$ in the BP analysis are all functions of $\sum_{i=1}^n T(X_i)$, then the testing statistic t_k is also independent of $\frac{S^{(k)^2}}{\left(\overline{X^{(k)}}\right)^2}$ $(k = 1,2,\dots,p)$.

We assume the final t-test is a two-sided test with the significance level of 0.05. If the experimental drug has no effect on all p endpoints, then the family-wise type I error is:

$$\alpha = \sum_{k=1}^{p} P\left(\frac{S^{(k)^2}}{\left(\overline{X^{(k)}}\right)^2} = min\left\{\frac{S^{(1)^2}}{\left(\overline{X^{(1)}}\right)^2}, \frac{S^{(2)^2}}{\left(\overline{X^{(2)}}\right)^2}, \dots, \frac{S^{(p)^2}}{\left(\overline{X^{(p)}}\right)^2}\right\}, |t_k| \geq t_{0.05,n-2}\right)$$
$$= \sum_{k=1}^{p} P\left(\frac{S^{(k)^2}}{\left(\overline{X^{(k)}}\right)^2} = min\left\{\frac{S^{(1)^2}}{\left(\overline{X^{(1)}}\right)^2}, \frac{S^{(2)^2}}{\left(\overline{X^{(2)}}\right)^2}, \dots, \frac{S^{(p)^2}}{\left(\overline{X^{(p)}}\right)^2}\right\}\right)$$
$$* P\big(|t_k| \geq t_{0.05,n-2}\big) = 1 * 0.05 = 0.05.$$

If the experimental drug has no effect only in the first l endpoint (without loss of generality) but has effect in the remaining endpoints, then the family-wise type I error is:

$$\alpha = \sum_{k=1}^{l} P\left(\frac{S^{(k)^2}}{\left(\overline{X^{(k)}}\right)^2} = min\left\{\frac{S^{(1)^2}}{\left(\overline{X^{(1)}}\right)^2}, \frac{S^{(2)^2}}{\left(\overline{X^{(2)}}\right)^2}, \dots, \frac{S^{(p)^2}}{\left(\overline{X^{(p)}}\right)^2}\right\}, |t_k| \geq t_{0.05,n-2}\right)$$
$$= \sum_{k=1}^{l} P\left(\frac{S^{(k)^2}}{\left(\overline{X^{(k)}}\right)^2} = min\left\{\frac{S^{(1)^2}}{\left(\overline{X^{(1)}}\right)^2}, \frac{S^{(2)^2}}{\left(\overline{X^{(2)}}\right)^2}, \dots, \frac{S^{(p)^2}}{\left(\overline{X^{(p)}}\right)^2}\right\}\right)$$
$$* P\big(|t_k| \geq t_{0.05,n-2}\big) \leq 1 * 0.05 = 0.05.$$

We have proved that the family-wise type I error is not inflated under the BP analysis, both in the strong and weak sense.

Next, we show through simulations *to support the proof.* For simplicity, we assume the BP analysis is performed on 2 endpoints X and Y, where the experimental drug has no treatment effect on X but *could have* a treatment effect on Y.

Mimicking study GN28352, we assume the following:

- 84 participants randomized to the placebo arm and 84 participants randomized to the experimental arm, with no dropout
- Endpoints X and Y both follow a normal distribution within each arm
- Within endpoints X and Y, without loss of generality, the standard deviation in both arms is 1
- Endpoint X has no treatment effect and the placebo arm mean is 1/0.65 (protocol CV assumption on API ADAD Composite Cognitive Test total score = 0.65)

In the simulation study, we vary the following, which yields 135 scenarios:

- Treatment effect in endpoint Y in RR scale: *0*%, 50%, and 100%
- Placebo mean in endpoint Y: 0, 1, 1.5, 2, 10
- *Endpoints X and Y follow a bivariate normal distribution with*
  - Correlation between endpoint X and Y within the placebo group: $\rho_{pbo} = 0.1$, 0.5, 0.9
  - Correlation between endpoint X and Y within the treatment group: $\rho_{exp} = 0.1$, 0.5, 0.9

The family-wise type I error in this case equals the probability of rejecting the null H_0 in endpoint X. *When there is a treatment effect in endpoint Y, the* only scenario where a type I error can happen is when endpoint X is chosen by the BP analysis. The number of iterations in the simulation study was 100,000. We summarize the simulation results on type I error (along with the probability of choosing endpoint Y *in scenarios where endpoint Y has a treatment effect*) in Tables 1–3.

Table 1 Family-Wise Strong Type I Error, Assuming *0*% Treatment Effect in Endpoint *Y (Type I Error in the Weak Sense)*

		Placebo Mean in Endpoint Y				
		0	**1**	**1.5**	**2**	**10**
$\rho_{pbo} = 0.1$	$\rho_{exp} = 0.1$	0.05090	0.04968	0.*0583*	0.*04965*	0.*04862*
	$\rho_{exp} = 0.5$	0.04928	0.04963	0.*04995*	0.*05029*	0.*04983*
	$\rho_{exp} = 0.9$	0.05042	0.04981	0.*05012*	0.*04894*	0.*05068*
$\rho_{pbo} = 0.5$	$\rho_{exp} = 0.1$	0.04949	0.04898	0.*04953*	0.*04982*	0.*05075*

		Placebo Mean in Endpoint Y				
	$\rho_{exp} = 0.5$	0.05052	0.05048	0.*05014*	0.*04953*	0.*04962*
	$\rho_{exp} = 0.9$	0.04913	0.05105	0.*05016*	0.*04927*	0.*04969*
$\rho_{pbo} = 0.9$	$\rho_{exp} = 0.1$	0.05074	0.05068	0.*05055*	0.*04959*	0.*05099*
	$\rho_{exp} = 0.5$	0.05010	0.05002	0.*05012*	0.*04951*	0.*05088*
	$\rho_{exp} = 0.9$	0.04969	0.05063	0.*04900*	0.*05004*	0.*04970*

Table 2 Family-Wise Strong Type I Error with Probability of Choosing Endpoint Y under the Blinded Power Analysis, Assuming 50% Treatment Effect in Endpoint Y *(Type I Error in the Strong Sense)*

		Placebo Mean in Endpoint Y				
		0	**1**	**1.5**	**2**	**10**
$\rho_{pbo} = 0.1$	$\rho_{exp} = 0.1$	0.04927 (0%)	0.05127 (0%)	0.05072 (0.03%)	0.04689 (9.0%)	0.00000 (100%)
	$\rho_{exp} = 0.5$	0.04948 (0%)	0.05005 (0%)	0.04937 (0%)	0.04570 (6.5%)	0.00000 (100%)
	$\rho_{exp} = 0.9$	0.04993 (0%)	0.05038 (0%)	0.05005 (0%)	0.04308 (2.9%)	0.00000 (100%)
$\rho_{pbo} = 0.5$	$\rho_{exp} = 0.1$	0.04936 (0%)	0.04950 (0%)	0.05061 (0.01%)	0.04590 (8.1%)	0.00000 (100%)
	$\rho_{exp} = 0.5$	0.04934 (0%)	0.04995 (0%)	0.04996 (0%)	0.04585 (5.4%)	0.00000 (100%)
	$\rho_{exp} = 0.9$	0.05039 (0%)	0.05019 (0%)	0.04961 (0%)	0.04601 (1.9%)	0.00000 (100%)
$\rho_{pbo} = 0.9$	$\rho_{exp} = 0.1$	0.04943 (0%)	0.05038 (0%)	0.05132 (0%)	0.04672 (5.8%)	0.00000 (100%)
	$\rho_{exp} = 0.5$	0.04975 (0%)	0.05020 (0%)	0.04959 (0%)	0.04746 (3.0%)	0.00000 (100%)
	$\rho_{exp} = 0.9$	0.05086 (0%)	0.04916 (0%)	0.04940 (0%)	0.04895 (0.4%)	0.00000 (100%)

Table 3 Family-Wise Strong Type I Error with Probability of Choosing Endpoint Y under the Blinded Power Analysis, Assuming 100% Treatment Effect in Endpoint Y *(Type I Error in the Strong Sense)*

		Placebo Mean in Endpoint Y				
		0	**1**	**1.5**	**2**	**10**
$\rho_{pbo} = 0.1$	$\rho_{exp} = 0.1$	0.05030 (0%)	0.05097 (0%)	0.05008 (0%)	0.05070 (0%)	0.04979 (0%)
	$\rho_{exp} = 0.5$	0.04990 (0%)	0.04985 (0%)	0.04988 (0%)	0.05013 (0%)	0.04838 (0%)
	$\rho_{exp} = 0.9$	0.04908 (0%)	0.04976 (0%)	0.05006 (0%)	0.05013 (0%)	0.05044 (0%)
$\rho_{pbo} = 0.5$	$\rho_{exp} = 0.1$	0.04964 (0%)	0.05094 (0%)	0.05063 (0%)	0.05004 (0%)	0.04949 (0%)
	$\rho_{exp} = 0.5$	0.05021 (0%)	0.05089 (0%)	0.05002 (0%)	0.05042 (0%)	0.05016 (0%)
	$\rho_{exp} = 0.9$	0.04965 (0%)	0.04976 (0%)	0.05033 (0%)	0.04960 (0%)	0.05064 (0%)
$\rho_{pbo} = 0.9$	$\rho_{exp} = 0.1$	0.05039 (0%)	0.05153 (0%)	0.04978 (0%)	0.05014 (0%)	0.04924 (0%)
	$\rho_{exp} = 0.5$	0.05148 (0%)	0.04994 (0%)	0.04925 (0%)	0.04939 (0%)	0.05031 (0%)
	$\rho_{exp} = 0.9$	0.04936 (0%)	0.05032 (0%)	0.05038 (0%)	0.04999 (0%)	0.05082 (0%)

As seen from Tables 1–3, the simulated family-wise type I errors are at or substantially below 0.05. *Also in scenarios where the treatment effect in Y is non-zero*, when the probability of choosing endpoint Y approaches to 100%, the family-wise type I error approaches to 0*%; this can be seen in the proof in Theorem 1*.

We conclude that the BP analysis does not inflate family-wise type I error both weakly or strongly.

Appendix 3 Details on Selected Statistical Models

In this section, we present the details on selected statistical models from this Statistical Analysis Plan (SAP).

CONTROL-BASED MEAN IMPUTATION SENSITIVITY ANALYSIS MODEL FROM SECTION 5.3.3

In the analysis using control-based mean imputation, the treatment effect in annualized rate of change is quantified as

$$\hat{\beta}_x[c] = f_{cren}^{comp}\left(\hat{\beta}_{cren}^{comp} - \hat{\beta}_{pbo,MAR} - c\right) + c,$$

where f_{cren}^{comp} is the percentage of participants from the crenezumab arm who have the score collected from the end of treatment visit A (ET-A)/end of treatment visit B (ET-B) visit; $\hat{\beta}_{cren}^{comp}$ is the RCRM estimated annualized rate of change when using the data only from the participants from the crenezumab arm who have the primary endpoint collected from the ET-A/ET-B visit; $\hat{\beta}_{pbo,MAR}$ is the RCRM estimated annualized rate of change when using the data only from the participants from the placebo arm; c is a non-positive constant (note that the sign of c is different from Mehrotra et al. (2017) because our primary endpoints are expected to decline over time) which will be addressed in the following paragraph. The variance of this treatment effect estimator is estimated as:

$$\hat{V}\left(\hat{\beta}_x[c]\right) = \left(f_{cren}^{comp}\left(1 - f_{cren}^{comp}\right)N_{cren}^{-1} + \left(f_{cren}^{comp}\right)^2\right)\left(\hat{V}_{cren}^{comp} + \hat{V}_{pbo,MAR}\right) + \left(f_{cren}^{comp}\left(1 - f_{cren}^{comp}\right)N_{cren}^{-1}\left(\hat{\beta}_{cren}^{comp} - \hat{\beta}_{pbo,MAR} - c\right)^2\right),$$

where N_{cren} is the number of participants in the crenezumab arm from the mITT population; $\hat{V}_{cren}^{comp}$ is the variance estimate on the RCRM estimated annualized rate of change when using the data only from the participants from the crenezumab arm who have the primary endpoint collected from the ET-A/ET-B visit; $\hat{V}_{pbo,MAR}$ is the variance estimate on the RCRM estimated annualized rate of change when using the data only from the participants from the placebo arm. The degrees of freedom corresponding to $\hat{V}_{cren}^{comp}$ and $\hat{V}_{pbo,MAR}$ using the Kenward and Roger method (Kenward and Roger 1997) are denoted as λ_{cren}^{comp} and $\lambda_{pbo,MAR}$, respectively. The degrees of freedom of the treatment effect estimator can be calculated as:

$$\lambda[c] = \frac{\left(\hat{V}\left(\hat{\beta}_x[c]\right)\right)^2}{\frac{\left(\hat{V}_{cren}^{comp}\right)^2}{\lambda_{cren}^{comp}} + \frac{\left(\hat{V}_{pbo,MAR}\right)^2}{\lambda_{pbo,MAR}}}.$$

We will then perform the imputation based on c=0. This scenario assumes that the mean slope from the participants with missing values from the crenezumab arm is the same as the estimated overall placebo slope. We note that this assumption is already conservative in estimating the treatment effect because it implicitly assumes participants

on treatment will even do worse than an average placebo participant, after dropping out of the study.

DELAYED-START ANALYSIS MODEL FROM SECTION 5.10.1.1

The statistical model for this delayed-start analysis can be written as:

$$Y_{ij} = \alpha + (\text{Randomization Factors})_i + u_{0i} + (\beta + \beta_x I_{x,i} + u_{1i})t_{ij} + (\beta_{x,B/CAS} t_{ij} - \beta_{x,B/OLE} T_i)(1 - I_{x,i}) I_{t_{ij} > T_i} + \epsilon_{ij},$$

where β_x quantifies the crenezumab annualized treatment effect in slowing down the disease progression in Period A of Study GN28352 and $\beta_{x,B/OLE}$ quantifies the treatment effect by switching from placebo to crenezumab in Period B of Study GN28352 and the OLE; T_i represents the time of the first endpoint assessment after participant i enters into Study Period B or the OLE.

Due to the common-close design of Study GN28352, mutation carriers will have varied duration of follow-up at the end Study Period of A. To calculate the crenezumab treatment effect at the end of Study Period A, we first calculate the arithmetic mean of duration in Study Period A (last endpoint assessment date – randomization date in Study Period A + 1) among mutation carriers and we call this duration $\bar{T}_1$. On average, the treatment benefit in slowing down the disease progression at the end of Study Period A can be written as:

$$\Delta_1 = \beta_x \bar{T}_1.$$

Similarly, we also calculate the arithmetic average of overall duration at the end of the OLE among mutation carriers and we call this duration $\bar{T}_2$. The average treatment benefit remained at the end of OLE can be written as:

$$\Delta_2 = (\beta_x - \beta_{x,B/OLE})\bar{T}_2 + \beta_{x,B/OLE}\bar{T}_1.$$

Note that both $\bar{T}_1$ and $\bar{T}_2$ are treated as constants and are estimated from the data. A non-inferiority test with different margins will be applied (Liu-Seifert et al. 2015). A 90% CI will be calculated for $\Delta_2 - 50\%\Delta_1$ or $\Delta_2 - 30\%\Delta_1$, depending on the specified margin in the test; if the lower limit of the CI is greater than 0, then the H_0 will be rejected, meaning that by the end of the OLE at least 50% (or 30%) of the treatment effect at the end of Study Period A is preserved.

7. REFERENCES

Acosta-Baena N, Sepulveda-Falla D, Lopera-Gómez CM, et al. Pre-dementia clinical stages in presenilin 1 E280A familial early-onset Alzheimer's disease: a retrospective cohort study. Lancet Neurol 2011;10:213–20.

Albert MS, DeKosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011;7:270–9.

Armitage SG. An analysis of certain psychological tests used for the evaluation of brain injury. Psychological Monographs 1946;60:i–48.

Bateman RJ, Xiong C, Benzinger TL, et al. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. N Engl J Med 2012;367:795–804.

Bretz F, Maurer W, Brannath W, et al. A graphical approach to sequentially rejective multiple test procedures. Stats Med 2009;28:586–604.

Bretz F, Posch M, Glimm E, et al. Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes, or parametric tests. Biom J 2011;53:894–913.

Casella G, Berger RL. Statistical inference. Cengage Learning 2021, 2nd edition.

Caselli RJ, Langlais BT, Dueck AC, et al. Neuropsychological decline up to 20 years before incident mild cognitive impairment. Alzheimers Dement 2020;16:512–23.

Chen K, Langbaum JB, Fleisher AS, et al. Twelve-month metabolic declines in probable Alzheimer's disease and amnestic mild cognitive impairment assessed using an empirically pre-defined statistical region-of-interest: findings from the Alzheimer's Disease Neuroimaging Initiative. Neuroimage 2010;51:654–64.

Cummings JL, Mega M, Gray K, et al. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. Neurology 1994;44:2308–14.

Cummings JL. The Neuropsychiatric Inventory: assessing psychopathology in dementia patients. Neurology 1997;48:S10–6.

Donohue MC, Sperling RA, Salmon DP, et. al. The preclinical Alzheimer cognitive composite: measuring amyloid-related decline. JAMA Neurol 2014;71:961–70.

Flaherty KR, Wells AU, Cottin V, et al. Nintedanib in progressive fibrosing interstitial lung diseases. New Engl J Med 2019;381:1718–27.

Folstein MF, Folstein SE, McHugh PR. "Mini-mental state." A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189–98.

Friede T, Kieser M. Blinded sample size re-estimation in superiority and noninferiority trials: bias versus variance in variance estimation. Pharm Stat 2013;12:141–6.

Friede T, Kieser M. Sample size recalculation in internal pilot study designs: a review. Biom J 2006;48:537–55.

Glimm E, Läuter J. Some notes on blinded sample size re-estimation. arXiv 2013 [resource on the Internet]. Available from: https://arxiv.org/pdf/1301.4167v1.pdf.

Gollan TH, Weissberger G, Runnqvist E, et al. Self-ratings of spoken language dominance: a multi-lingual naming test (MINT) and preliminary norms for young and aging Spanish-English bilinguals. Biling (Camb Engl) 2012;15:594–615.

Hu N, Mackey H, Thomas R. Power and sample size for random coefficient regression models in randomized experiments with monotone missing data. Biom J 2021;63:806–24.

[redacted], *et al. Adapative endpoint selection in randomized clinical trials. Manuscript in preparation. 2022.*

Huber PJ. The behavior of maximum likelihood estimates under nonstandard conditions. Berkeley Symposium on Mathematical Statistics and Probability 1967:221–3.

Kenward MG, Roger JH. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 1997;53:983–97.

Liu-Seifert H, Andersen SW, Lipkovich I, et al. A novel approach to delayed-start analyses for demonstrating disease-modifying effects in Alzheimer's disease. PloS One 2015;10:e0119632.

McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011;7:263–9.

Mehrotra DV, Liu F, Permutt T. Missing data in clinical trials: control-based mean imputation and sensitivity analysis. Pharm Stat 2017;16:378–92.

Mohs RC, Knopman D, Petersen RC, et al. Development of cognitive instruments for use in clinical trials of antidementia drugs: additions to the Alzheimer's Disease Assessment Scale that broaden its scope. Alzheimer Dis Assoc Disord 1997;11:S13–21.

Morris JC, Heyman A, Mohs RC, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. Neurology 1989;39:1159–65.

Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology 1993;43:2412–14.

Randolph C, Tierney MC, Mohr E, et al. Repeatable Battery for the Assessment of Neuropsychological Status. San Antonio, TX: The Psychological Coporation, 1998.

Raven JC. Revised Manual for Raven's Progressive Matrices and Vocabulary Scale. New York: The Psychological Corporation, Harcourt Brace Jovanovich, Inc., 1976.

Richeldi L, du Bois RM, Raghu G, et al. Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. New Engl J Med 2014;370:2071–82.

Robins JM, Tsiatis AA. Correcting for non-compliance in randomized trials using rank preserving structural failure time models. Communications in Statistics-Theory and Methods 1991;20:2609–31.

Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. Am J Psychiatry 1984;141:1356–64

Sclan SG, Reisberg B. Functional assessment staging (FAST) in Alzheimer's disease: reliability, validity, and ordinality. Int Psychogeriatr 1992;4:55–69.

Sheikh JI and Yesavage JA. Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. Clinical Gerontology 1986;5:165–73.

Sperling RA, Aisen PS, Beckett LA, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines of Alzheimer’s disease. Alzheimers Dement 2011;7:280–92.

Tounsi H, Deweer B, Ergis AM, et al. Sensitivity to semantic cuing: an index of episodic memory dysfunction in early Alzheimer disease. Alzheimer Dis Assoc Disord 1999;13:38–46.

U.S. Food and Drug Administration (FDA). Multiple endpoints in clinical trials: guidance for industry [resource on the Internet] 2017. Available from: https://www.fda.gov/files/drugs/published/Multiple-Endpoints-in-Clinical-Trials-Guidance-for-Industry.pdf.

van Dyck CH, Nygaard HB, Chen K, et al. Effect of AZD0530 on cerebral metabolic decline in Alzheimer disease: a randomized clinical trial. JAMA Neurol 2019;76:1219–29.

White H. Maximum likelihood estimation of misspecified models. Econometrica 1982;50:1–25.