NOVARTIS

Clinical Development

CADPT06

CADPT06A12201 / NCT04795466

Exploratory Platform trial of Anti-Inflammatory agents in Alzheimer's Disease (EXPLAIN-AD): A randomized, placebo-controlled, multicenter platform study to evaluate efficacy, safety, tolerability and pharmacokinetics of various anti-inflammatory agents in patients with mild cognitive impairment due to Alzheimer's disease and mild Alzheimer's disease

Statistical Analysis Plan (SAP)

Author:	Personal Protected Data (PPD)
Document type:	SAP Documentation
Document status:	Final – Amendement 1
Release date:	03-May-2024
Number of pages:	27

Template Version 4.0, Effective from 23-Apr-2021

Commercially Confidential Information (CCI)

Table of contents

Commercially Confidential Information

List of tables

Commercially Confidential Information

List of figures

List of abbreviations

AD	Alzheimer's Disease
ADL	Activities of daily living
AE	Adverse Event
	Commercially Confidential Information
ANCOVA	Analysis of Covariance
CFT	Category Fluency Test
COWAT	Controlled Oral Word Association Test
CRF	Case Report Form
CSF	Cerebrospinal Fluid
C-SSRS	Columbia Suicide Severity Rating Scale
CSR	Clinical Study Report
CV	Coeficient of Variation
DBP	Diastolic Blood Pressure
DMC	Data Monitoring Committee
DMS	Document Management System
DNA	Deoxyribonucleic Acid
DSST	The Digit Symbol Substitution Test
ECG	Electrocardiogram
e-Cog	Everyday Cognition Scale
FAS	Full Analysis Set
IA	Interim Analyses
IG	Immunogenicity
IL	Interleukin
LLOQ	Lower Limit of Quantification
MCI	Mild Cognitive Impairment
MedDRA	Medical Dictionary for Drug Regulatory Affairs
MMRM	Mixed Model for Reapeted Measurements
MMSE	Mini-Mental Status Examination
	Commercially Confidential Information
NPI	Neuropsychiatric Inventory
NPI-D	Neuropsychiatric Inventory-Caregiver Distress Scale
NTB	Neuropsychological Test Batter
	Commercially Confidential Information
PD	Pharmacodynamic
PET	Positron Emission Tomography
PK	Pharmacokinetics
PRO	Patient-reported Outcomes
P-Tau	Phosphorylated tau protein
QTcF	QT interval corrected by Fridericia's formula
RAP	Reporting & Analysis Process

RAVLT	Rey Auditory Verbal Learning Test
REML	Restricted Maximum Likelihood
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SAS	Statistical Analysis System
SBP	Systolic Blood Pressure
s.c	Subcutaneous
SD	Standard Deviation
	Commercially Confidential Information
TFLs	Tables, Figures, Listings
T-Tau	Total tau protein
WHO	World Health Organization

1 Introduction

This Statistical Analysis Plan (SAP) describes the implementation of all statistical analyses planned in the platform study ADPT06A12101 protocol.

The content of this SAP is based on the protocol version v04 dated 15-Sep-2022.

For each completed cohort, the data analysis will be performed upon final database lock for the cohort, and the results will be reported upon completion of each cohort in an end of cohort report. This analysis will include all available study data collected from the completed treatment arms within the cohort (and data from previous cohorts where available and appropriate). The purpose of this analysis is to formally assess safety and efficacy and to ensure data cleaning and reporting occurs on a periodic basis.

Commercially Confidential Information

This study will include an external Data Monitoring Committee (DMC). The DMC will function independently of site investigators participating in the study. The purpose of the DMC is to assess the progress of the EXPLAIN-AD treatment cohorts and the safety data.

Clinical Study Report (CSR). CCI DMC deliverables (tables, figures, listings - TFLs) and further programming specifications are described in separate documents.

1.1 Study design

This is a randomized, placebo-controlled, participant- and investigator-blinded platform study in participants with MCI due to AD or mild AD who have evidence of peripheral inflammation.

The EXPLAIN-AD study uses a platform type design to investigate “multiple targeted therapies in the context of a single disease in a perpetual manner”. Each investigational agent or batch of agents and matching placebo entered into the trial at a given time will be considered a unique cohort. The decision to include future cohorts is not dependent on performance of prior agents or cohorts. Each time a new cohort is introduced, agent-specific information will be added to the protocol as a substantial amendment and submitted to Health Authorities and Ethical Committees, as required by local regulations.

Each cohort in the study will undergo the same study evaluations and assessments. Each cohort will include a screening period (Day -60 to Day -8), followed by a baseline period of 7 days (Day -7 to Day -1), a treatment period of 20 weeks (Day 1 to Day 141), and a study completion evaluation [End of Cohort (EOC1)] approximately 30 days after the last agent administration (Day 171). For agents with longer half-lives, such as monoclonal antibodies, the same study design will be followed; however, an additional post-treatment follow-up visit (EOC2) will occur on Day 281.

Participants who meet the eligibility criteria at screening will have baseline assessments performed. All baseline safety evaluation results must be available prior to randomization. All participants will undergo assessments of cognition, neuropsychiatric symptoms and function (activities of daily living) at various time points throughout the duration of the trial.

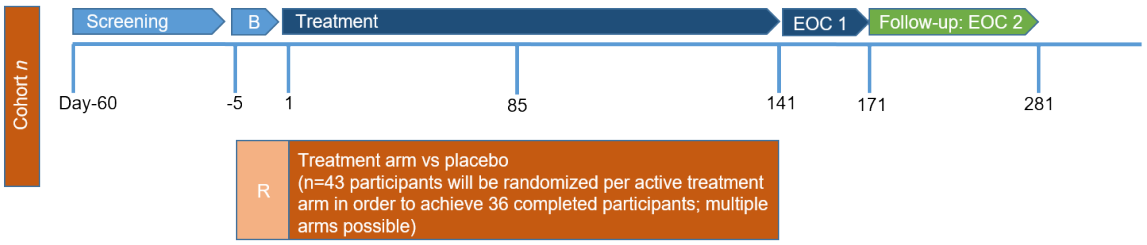
Moreover, all participants will undergo CSF sampling at screening and following completion of 12 weeks treatment. In addition, blood samples will be collected throughout the study.

In addition, a sub-set of these same participants (up to 22 per investigational agent/matching placebo) will undergo PET CCI imaging at baseline and following completion of the first 12 weeks treatment. Alterations in activated microglia and astrocytes due to therapeutic intervention will be measured using a PET CCI . PET imaging in this subset of participants and analysis of CSF and/or serum samples in all participants will provide measures of both central and peripheral inflammation following treatment with an investigational agent/matching placebo. An interim analysis of the PET CCI and CSF soluble biomarker data may be conducted following completion of the first 12 weeks of treatment.

For each investigational agent, one or more interim analyses may be conducted, as appropriate, to support decision making concerning the sponsor's clinical development projects in general, or in case of any safety concerns. The results of these analyses will not have any impact on study conduct.

The study design is described in Figure 1-1.

Figure 1-1 Study design schematic



B = Baseline
R = Randomization
EOC = End of cohort

Randomization for any participant can occur after informed consent is obtained and eligibility is confirmed, ideally as close to Day 1 as possible. Participants will be randomized into any of the treatment arms open to enrolment for which the participant meets the eligibility criteria.
Treatment arm(s) will be defined for each cohort.
EOC2 visit only required for an extended half-life. For all other treatments, EOC1 will serve as End of Cohort visit.

Approximately n=86 participants were planned to be randomized in a 1:1 ratio (active:placebo) using an IRT system. The study was prematurely terminated after 34 participants randomized due to sponsor's decision.

The primary endpoint is the change in NTB total z-score at week 24.

1.2 Study objectives, endpoints and estimands

Objective(s)	Endpoint(s)
Primary objective(s)	**Endpoint(s) for primary objective(s)**
• To compare the effects of each individual agent vs. placebo on cognition in early AD	• The Neuropsychological Test Battery (NTB) score at 24 weeks
Secondary objective(s)	**Endpoint(s) for secondary objective(s)**
• To investigate the safety (AEs) and tolerability of each individual agent vs placebo	• ECG measurements and findings • Adverse events and serious adverse events • Laboratory measurements • Vital signs • Prospective suicidality assessment (ideation and behavior) from electronic Columbia Suicide Severity Rating Scale (eC-SSRS)
• To investigate the effects of each individual agent vs placebo in lowering central inflammation	• Reduction of microglial activation measured by Positron-Emission Tomography CCI 18 kDa (PET CCI) following the initial 12-weeks of treatment in a subset of participants
• To explore the effects of each individual agent vs placebo on neuropsychiatric symptoms	• The Neuropsychiatric Inventory (NPI-D) total score at 24 weeks • eNeuropsychiatric At-Home Caregiver assessment score at 12 weeks
• To compare the effects of each individual agent vs placebo on function (activities of daily living) • To compare the effects of each individual agent vs placebo on memory and executive function	• The Everyday Cognition (ECog) scale at 24 weeks • The NTB memory and executive function composites at 24 weeks • DSST at 24 weeks • eCognitive at-home assessment at 12 weeks
• To determine the pharmacokinetics of each individual agent	• Agent concentration in serum (for biotherapeutic agents) or plasma (for low molecular weight agents) and in cerebrospinal fluid (CSF) • Ratio of agent concentration in CSF to that in serum (for biotherapeutic agents) or plasma (for low molecular weight agents)
• To determine the total target and immunogenicity of each individual biotherapeutic agent	• Commercially Confidential Information • Anti-agent antibodies in serum (when applicable)

1.2.1 Primary estimand(s)

The primary aim of the study is to demonstrate the effect on cognition of each anti-inflammatory agent vs. placebo, as measured by the NTB response after the 24 week treatment period.

The primary estimand is the change in NTB total z-score at weeks 24.

2 Statistical methods

2.1 Data analysis general information

Data will be analyzed by Novartis Biostatistics and Statistical Programming personnel according to the data analysis section 12 of the study protocol as detailed in this analysis plan.

SAS® version 9.4 (or later version if available at time of analysis) will be used for all analyses.

Data from all patients who signed informed consent will be used in the analysis; no center effect will be assessed as it is expected to have small sample size of enrollment at each individual centers.

General analysis conventions

Unless otherwise specified: categorical data will be presented as frequencies and percentages; continuous data will be presented as n, mean, standard deviation (SD), median, 25th and 75th percentiles, minimum, and maximum.

2.1.1 General definitions

Investigational treatment or **investigational drug** refers to the investigational agent tested within each cohort.

Study treatment or **study drug** refers to the investigational agent or matching placebo.

Treatment or **treatment group**

For presentation in the outputs, **treatment** refers to:

- **ACZ885** or **Placebo** in cohort 1

Date of first administration of study treatment

The date of first administration of study treatment is derived as the first date when a nonzero dose of study treatment was administered as per the Dosage Administration CRF. The date of first administration of study treatment will also be referred as ***start of study treatment***.

Date of last administration of study treatment

The date of last administration of study treatment is derived as the last date when a nonzero dose of study treatment was administered as per the Dosage Administration CRF. The date of last administration of study treatment will also be referred as ***end of study treatment***.

Study day

Study day 1 for all assessments is taken to be the start of study treatment.

The study day for all assessments will be calculated as follows:

1. If date of assessment occurred on or after the start of study treatment, then
   Study day = Date of assessment - Start of study treatment + 1.
2. If date of assessment occurred before the start of study treatment, then
   Study day = Date of assessment - Start of study treatment.

The study day will be displayed in the data listings. If an event starts before the reference start date, the study day displayed on the listing will be negative.

Baseline

For safety evaluations, the last available assessment on or before the date of start of study treatment is taken as baseline assessment. In case time of assessment and time of treatment start is captured, the last available assessment before the treatment start date/time is used for baseline.

For safety parameters (e.g. ECGs), where the study requires multiple replicates per time point, the average of these measurements will be calculated (if not already available in the database) to determine the baseline. The associated time will be the time of the last replicate.

In rare cases where multiple measurements meet the baseline definition, with no further flag or label that can identify the chronological order, then the following rule should be applied: If values are from central and local laboratories, the value from central assessment should be considered as baseline. If multiple values are from the same laboratory (local or central) or collected for ECGs or vital signs, then the median should be considered as baseline.

If participants have no value as defined above, the baseline result will be missing.

For safety parameters other than ECG, scheduled pre-dose collections as well as unscheduled collections on Day 1 for which no time is available will be considered as pre-dose.

For ECG, if dosing time on study Day 1 or ECG time for scheduled pre-dose ECGs is missing, the pre-dose ECG will be considered as baseline. If a scheduled pre-dose measurement actually occurred post-dose, then the corresponding measurement will be treated and analyzed similar to an unscheduled post-dose measurement.

Unscheduled Assessment:

In general, descriptive summary statistics will not include unscheduled/repeat assessment.

For all analyses regarding abnormal assessments or analyses based on worst post–baseline value (e.g. laboratory, ECGs, vital signs), all post-baseline values will be included (scheduled, unscheduled, repeat). All unscheduled and repeat measurements will be included in listings.

Analysis visit window:

No time-window will be applied in this study.

In summary by visit the scheduled EOC1 and EOC2 (if available) visit will be included.

On-treatment assessment/event

The overall observation period will be divided into 2 mutually exclusive segments:

1. ***pre-treatment period***: from day of participant's informed consent to before date/time of first administration of study treatment
2. ***on-treatment period***: from date/time of first administration of study treatment to the last visit date EOC1 (or EOC2 if applicable)

Note: If dates are incomplete in a way that clear assignment to pre-, on-, post-treatment period cannot be made, then the respective data will be assigned to the on-treatment period.

Safety summaries (tables, figures) include only data from the on-treatment period with the exception of baseline data which will also be summarized where appropriate (e.g. change from baseline summaries). In particular, summary tables for adverse events (AEs) will summarize only on-treatment events, with a start date during the on-treatment period (***treatment-emergent*** AEs).

For all safety data listings (e.g. ECG, lab, vital signs, AEs), data collected during the post-treatment period will be flagged.

2.2 Analysis sets

For each completed cohort, the data analysis will be performed upon final database lock for the cohort, and the results will be reported while the study may be ongoing for subsequent cohorts. For all analysis sets, participants will be analyzed according to the study treatment(s) received.

The full analysis set (FAS) will include all participants who received any study drug.

The safety analysis set is defined similarly as the FAS.

The PK analysis set will include all participants with at least one available valid (i.e., not flagged for exclusion) PK concentration measurement, who received any study drug and with no protocol deviations that impact PK data.

The PD analysis set will include all participants with available PD data and no protocol deviations with relevant impact on PD data. This set will be used for the primary endpoint analysis.

The IG analysis set will include all participants with at least one available valid (i.e., not flagged for exclusion) IG concentration measurement, who received any biotherapeutic study drug and with no protocol deviations that impact IG data.

Table 2-1 Protocol deviations and exclusions

Deviation code	Text description	Data exclusion
TRT08	Canakinumab/placebo wrong medication pack number administered at Day 85 visit to participant CCI	If a participant randomized on placebo took active drug, the adverse events recorded after the date/time of administration of the active drug will be excluded from the analysis, They will be summarized separately. Conversely, If a participant randomized on active took placebo, the data will be analyzed as planned. The participant took half of the planned dose.
EXCL02A	Ongoing medical history impacting cognition(Not AD) – Participant CCI	The participant will be excluded from the PD analysis set.
M-OTH38	Participant CCI has a CSF collection not per protocol at Day 85 (Backup sample used, using a different tube)	An additional analysis of biomarkers in CSF may be conducted excluding the impacted CSF sample from the analysis.

Deviation code	Text description	Data exclusion
M-OTH38	Participant CCI has a CSF collection not per protocol at Day 85 (Day85 CSF sample was done 2 weeks after the other Day 85 assessments)	-PK: the PK ratio CSF/serum will not be calculated (see Section 2.8, ratios will be calculated only if serum and CSF samples were taken within a window +/- 2 days). -Biomarkers: no impact on the analysis of biomarkers in CSF.

2.2.1 Subgroup of interest

Subgroups of participants MCI or Mild AD and participants CCI positive or negative may be used for supplementary analyses of efficacy, if the sample size in the subgroup is sufficient.

2.3 Patient disposition, demographics and other baseline characteristics

2.3.1 Patient disposition

Participant disposition will be presented using the Safety set for all participants and by actual treatment received. The following summaries will be provided:

- Number (%) of participants who completed treatment and those who discontinued the study treatment phase along with the primary reason for study discontinuation
- At IA only: Number (%) of participants who are still on-treatment (based on the 'EOC1 (or EOC2 if applicable) disposition page not completed)

Participant disposition data will be listed.

2.3.2 Demographics and other baseline characteristics

Demographic and other baseline data including disease characteristics will be listed and summarized descriptively by treatment group for all participants in the FAS.

Categorical data will be presented as frequencies and percentages. For continuous data, mean, standard deviation, median, minimum, and maximum will be presented.

Relevant medical histories and current medical conditions at baseline will be listed by treatment, participant, system organ class and preferred term.

2.4 Treatments (study treatment, rescue medication, concomitant therapies, compliance)

2.4.1 Study treatment / compliance

The duration of exposure in months to each treatment arm as well as dose intensity (computed as the ratio of actual cumulative dose received and actual duration of exposure) and the relative dose intensity (computed as the ratio of dose intensity and planned dose intensity) will be summarized by means of descriptive statistics using the safety set.

2.4.2 Prior, concomitant and post therapies

Concomitant medications and significant non-drug therapies prior to and after the start of the study treatment will be listed according to the Anatomical Therapeutic Chemical classification system, by treatment group and for all participants.

2.5 Analysis supporting primary objective(s)

The primary endpoint, NTB total score, includes nine validated components: 1-2) Wechsler Memory Scale Visual-Paired Associates immediate and delayed scores; 3-4); Wechsler Memory Scale Verbal-Paired Associates immediate and delayed scores; 5-6) Rey Auditory Verbal Learning Test (RAVLT) immediate and delayed scores; 7) Wechsler Memory Scale Digit Span; 8) Controlled Word Association Test (COWAT); and 9) Category Fluency Test (CFT). In order to reduce burden on sites and participants, components 1-4 will not be administered in EXPLAIN-AD.

For each of the remaining five components, a raw score is first converted to a standardized z-score using baseline mean and SD (calculated from all participants in the PD analysis set with baseline scores), and a total z-score is derived by averaging all resulting z-scores. A change from baseline is calculated as post-baseline z-score minus pre-treatment z-score, whereby a positive change indicates an improvement from baseline.

The primary analysis will be performed on the PD analysis set.

2.5.1 Statistical hypothesis, model, and method of analysis

The study research hypothesis is that each investigated anti-inflammatory agent provides an improvement over placebo in NTB response at 24 weeks.

The change from baseline in NTB total z-score will be analyzed using a restricted maximum likelihood (REML)-based mixed model with repeated measurements (MMRM). The model will include fixed, categorical effects of treatment, visit, and treatment-by-visit interaction, as well as the continuous fixed covariates of baseline and baseline-by-visit interaction. Unstructured covariance will be used to model within-patient errors, with Kenward-Roger approximation to estimate denominator degrees of freedom. If the unstructured covariance causes model convergence issues, then other simpler covariance structures will be considered. Least squares mean, the associated 2-sided 90% confidence interval and the p-value will be obtained for each treatment at each visit. The primary comparison will be the contrast between active treatment and placebo at 24 weeks. Both cohort-wise and combined cohort analysis (i.e., pooling placebo data across different cohorts) will be performed.

2.5.2 Handling of intercurrent events

Not applicable.

2.5.3 Handling of missing values not related to intercurrent event

For an NTB outcome at any visit, if more than three out of five NTB items are missing, the total z-score will not be derived and will be set to missing. Completely missing visits will be handled

through the MMRM model, which implicitly imputes missing data under missing at random (MAR) assumption.

2.5.4 Sensitivity analyses

To assess the robustness of the primary analysis results, several sensitivity analyses may be explored, using the primary model adjusted for several important baseline prognostic factors, such as age and educational level.

2.5.5 Supplementary analyses

A supplementary analysis will be performed by averaging the screening and baseline visits to serve as new baseline, the visits at day 57 and Day 85 to serve as new day 85, and the last 2 visits (day 141 and Day 171) to serve as new Day171.

A supplementary analysis may involve estimation of treatment effects and contrasts (anti-inflammatory agent vs. placebo) within subgroups (e.g., MCI or mild AD CCI positive versus negative) if the sample size in the subgroups is sufficient.

2.6 Analysis supporting secondary objectives

Key secondary objectives of EXPLAIN-AD are to determine the safety of each individual agent and to demonstrate how each individual agent affects imaging markers of inflammation in early AD.

Safety and tolerability of each agent will be evaluated via adverse events, laboratory measurements, vital signs, physical examination, ECG findings and suicidality assessment. Details are provided in Section 2.7.

2.6.1 Secondary endpoint(s)

2.6.2 Memory, cognition, executive function, and neuropsychiatric outcomes

Memory

The "memory function" composite score is obtained by averaging the following z-scores from the NTB: RAVLT immediate and delayed scores.

Executive function

The "executive function" composite score is obtained by averaging the following three z-scores from the NTB: Wechsler Memory Scale Digit Span, COWAT, and CFT.

Cognition

The DSST is an attention-demanding component of the Wechsler Adult Intelligence Scale-IV. The DSST score is the number of digits coded correctly in a fixed amount of time, whereby higher scores denote better performance.

The eCognitive assessments will include tests on CCI administered on a tablet. The primary derived metrics of these tests are accuracy (obtained through the number of errors made) and speed (obtained through response time in milliseconds).

The Everyday Cognition (ECog) scale is a validated informant-based measure comprised of 39 items covering six cognitively relevant domains. Each item is scored on a 4-point scale (1=better or no change compared to 10 years earlier, 2=questionable/occasionally worse, 3=consistently a little worse, 4=consistently much worse). An "I don't know" response is also included. The total ECog score is calculated as the sum of all 39 items, and the average score is derived from the mean average of all responses. If the response is "I don't know", then the item is not included in the calculation. The lower total ECog and average ECog scores indicate better performance.

Neuropsychriatric inventory (NPI)

The eNeuropsychiatric At-Home and the NPI-D assessments are study partner reported outcome measures of 12 neuropsychiatric symptoms and the distress associated with these disturbances. The NPI total score is in the range from 0 to 144, with higher values indicating greater disturbance.

Ten behavioral and two neurovegetative domains are included in the NPI:

The 10 behavioral domains are:

A. Delusions
B. Hallucinations
C. Agitation/Aggression
D. Depression/Dysphoria
E. Anxiety
F. Elation/Euphoria
G. Apathy/Indifference
H. Disinhibition
I. Irritability/Lability
J. Aberrant motor behavior

The 2 neurovegetative domains are:

K. Sleep and Nighttime Behavior Disorders
L. Appetite and Eating Disorders

For each domain:

- Frequency is rated as:
  1. Rarely – less than once per week
  2. Sometimes – about once per week
  3. Often – several times per week but less than every day
  4. Very often – once or more per day
- Severity is rated as:

1. Mild – produces little distress in the patient
2. Moderate – more disturbing to the patient but can be redirected by the caregiver
3. Severe – very disturbing to the patient and difficult to redirect

The total score for each domain is:

- domain score = frequency x severity

Caregiver distress is scored as:

0. Not at all
1. Minimally (almost no change in work routine)
2. Mildly (some change in work routine but little time rebudgeting required)
3. Moderately (disrupts work routine, requires time rebudgeting)
4. Severely (disruptive, upsetting to staff and other residents, major time infringement)
5. Very Severely or Extremely (very disruptive, major source of distress for staff and other residents, requires time usually devoted to other residents or activities)

Thus, for each domain there are four scores:

- Frequency
- Severity
- Total (frequency x severity)
- Caregiver distress

The NPI total score will be calculated by adding the scores of the 12 domain total scores together. The distress score is not included in the total NPI total score.

The total distress score (NPI-D) is calculated by adding together the scores of the 12 individual NPI distress questions.

PET CCI

The PET CCI will be examined at baseline and upon completion of week 12.

Imaging analyses will include, but are not limited to:

- Standard uptake values (SUV), SUV ratio (SUVR) and volume of distribution (Vt) in different brain regions (i.e. occipital lobe, whole brain and cerebral white matter) for each CCI PET scan, and change after canakinumab treatment, compared to baseline
- Summary statistics by dose cohort of the relative change from baseline in % Vt, SUV and SUVR CCI after treatment with canakinumab
- Plot of change from baseline of individuals with arithmetic mean overlaid in Vt, SUV and SUVR CCI

Total target

For biotherapeutic agents, measurement of total target concentrations (the sum of free and drug bound target) is used as a PD marker to evaluate target engagement.

In cohort 1, measurement of total IL-1β concentrations CCI in serum and CSF is used as a PD marker to evaluate CCI. Pharmacodynamic (PD) samples will be obtained and evaluated in all participants, including the placebo group.

Immunogenicity (Anti-ACZ885 Antibodies)

IG samples will be obtained and evaluated in serum of all agent treated participants.

To assess potential immunogenicity, serum samples for determination of anti-canakinumab antibodies will be collected at the timepoints defined in the PK/PD/IG Assessment Schedule.

2.6.3 Statistical hypothesis, model, and method of analysis

Memory, cognition, executive function, and neuropsychiatric outcomes

The memory function composite score, the executive function composite score, the DSST score, the overall eCognitive test scores, the total ECog score and the average ECog score, and the total NPI and NPI-D scores will be analyzed using the same model as for the primary endpoint (See Section 2.5).

PET CCI

As only one participant has data available, the PET CCI data will be only listed.

Total CCI

For biotherapeutic agents, total CCI concentration data will be listed by treatment, participant, and visit/sampling time point. Summary statistics will be provided by treatment and visit/sampling time point.

In cohort 1, concentrations of serum and CSF total Commercially Confidential Information will be given in mass per volume units. Concentrations below the LLOQ will be reported as "zero".

Immunogenicity (Anti-ACZ885 Antibodies)

Summary statistics of the number of positive participants (n, %) will be provided by timepoint.

2.6.4 Handling of intercurrent events

NA

2.6.5 Handling of missing values not related to intercurrent event

NA

2.6.6 Sensitivity analyses

Sensitivity analyses may be conducted by adjusting the models to important prognostics factors such as age or educational level.

2.6.7 Supplementary analyses

Supplementary analysis of NTB memory, cognition, executive function, DSST and NPI will be conducted using the primary model, by averaging the screening and baseline visits to serve as new baseline, the visits at day 57 and Day 85 to serve as new day 85, and the last 2 visits (day 141 and Day 171) to serve as new Day171.

At home assessments will have their 2 baseline data-points averaged to serve as baseline; data prior to visit 29 will be averaged to serve as Day 29 and data prior to Day 85 will be averaged to serve as Day 85.

In these analyses averaging 2 visits/timepoints, if one data is missing, the other single data will be used as the average.

2.7 Safety analyses

For all safety analyses, the safety set will be used.

2.7.1 Adverse events (AEs)

All information obtained on adverse events will be displayed by treatment and participant.

The number (and percentage) of participants with treatment emergent adverse events (events started after the first dose of study medication or events present prior to start of study treatment but increased in severity based on preferred term) will be summarized in the following ways:

- by treatment, primary system organ class and preferred term.
- by treatment, primary system organ class, preferred term and maximum severity.

Separate summaries will be provided for study drug related adverse events, death, serious adverse events, other significant adverse events leading to discontinuation.

A participant with multiple adverse events within a primary system organ class is only counted once towards the total of the primary system organ class.

In AE summaries, the primary system organ class will be presented alphabetically and the preferred terms will be sorted within primary SOC in descending frequency. The sort order for the preferred term will be based on their frequency in all participants.

In addition, for EudraCT requirements a summary of (1) Serious AEs and deaths, with number of occurrences and (2) Non-serious AEs, with number of occurrences will be produced (an occurrence is defined as >1 day between start and prior end date of record of same preferred term).

2.7.2 Deaths

All deaths will be listed using Safety set and post treatment deaths will be flagged. A separate listing of deaths prior to starting treatment will be provided for all screened participants.

2.7.3 Laboratory data

All laboratory data will be listed by treatment group, participant, and visit/time and if normal ranges are available abnormalities will be flagged. Summary statistics of values and changes from baseline will be provided by treatment and visit/time.

2.7.4 Other safety data

2.7.4.1 ECG

All ECG data will be listed by treatment group, participant and visit/time; abnormalities will be flagged. Summary statistics will be provided by treatment and visit/time.

The number and percentage of participants with notable ECG values will be presented as follows:

- QT, QTcF
  - New value of > 450 and ≤ 480 ms
  - New value of > 480 and ≤ 500 ms
  - New value of > 500 ms
  - Increase from baseline of > 30 ms to ≤ 60 ms
  - Increase from baseline of > 60 ms
- HR
  - Increase from baseline >25% and to a value > 100 bpm
  - Decrease from baseline >25% and to a value < 50 bpm
- PR
  - Increase from baseline >25% and to a value > 200 ms
  - New value of > 200 ms
- QRS
  - Increase from baseline >25% and to a value > 110 ms
  - New values of QRS > 110 ms

ECG data will be summarized by presenting summary statistics of observed data and change from baseline by treatment and timepoint. The definition of baseline is provided in Section 2.1.1.

A listing will be provided for participants with any notable values (defined above). If there is any notable value of an ECG for a participant, all measurements of this ECG value for the participant will be presented in this listing with the notable values flagged. A listing will also be provided of ECG findings for which there is a clinically significant ECG abnormality.

2.7.4.2 Vital signs

All vital signs data will be listed by treatment, participant, and visit/time and if ranges are available, abnormalities (and relevant orthostatic changes) will be flagged. Summary statistics will be provided by treatment and visit/time.

The following parameters were collected: height (cm), weight (kg), body temperature (°C), pulse rate (beats per minute), systolic and diastolic blood pressure (mmHg).

Weight, pulse rate, systolic and diastolic blood pressure will be summarized by presenting summary statistics of raw data and change from baseline by treatment and timepoint.

A listing will be provided for participants with any notable values (defined below). If there is any notable value of a vital sign for a participant, all measurements of this vital sign for the participant will be presented in this listing with the notable values flagged.

Notable criteria (High/Low):

- Systolic blood pressure [mmHg]: >140/<90 mmHg
- Diastolic blood pressure [mmHg]: >90/<50 mmHg
- Pulse rate [bpm]: >90/<45 bpm
- Weight [kg]: >120/<50 kg
- Temperature [°C]: >37.5/<35 °C

2.7.4.3 Suicidality assessment

C-SSRS data will be listed by treatment, participant, and visit/time; abnormalities will be flagged. Summary statistics will be provided by treatment and visit/time, if deemed relevant.

2.8 Pharmacokinetic endpoints

Agent serum concentration data will be listed by treatment, participant, and visit/sampling time point. Descriptive summary statistics will be provided by treatment and visit/sampling time point, Commercially Confidential Information

Summary statistics will include mean (arithmetic and geometric), SD, CV (arithmetic and geometric), median, minimum, and maximum. Commercially Confidential Information

Pharmacokinetic parameters will be listed by treatment and participant. Descriptive summary statistics will include mean (arithmetic and geometric), SD, and CV (arithmetic and geometric), median, minimum, and maximum. An exception to this is *Tmax* where median, minimum, and maximum will be presented, as applicable.

CSF to serum concentration ratio (expressed in %) will be calculated by the ratio of agent exposure in CSF to that in serum (for biotherapeutic agents) for post dose samples. The ratio will be calculated only if the collection dates of CSF and serum samples are on same day +/-2 days.

In cohort 1, Canakinumab concentrations will be assessed in serum and in CSF of all agent-treated participants. Serum and CSF PK samples will be collected at the timepoints defined in the PK/PD/IG Assessment Schedule (Predose before start of study treatment at screening visit and Day 85 in CSF and on Day 1 and Day 85 in serum). Any CSF samples having more than

500 red blood cells/µL will be documented, as such high cell counts may impact PK analysis. Concentrations of canakinumab in serum and CSF will be given in mass per volume units.

A previously utilized and validated mixed effects modeling approach may be used to characterize the exposure-related PK parameters (i.e., AUC and CL) of canakinumab. These results may be reported in a separate report.

2.9 PD and PK/PD analyses

The relationships between individual PK profiles or derived PK parameters and efficacy/PD measurements may be explored using graphical approaches (e.g., scatter plots) and regression analysis as appropriate.

Commercially Confidential Information

Commercially Confidential Information

Commercially Confidential Information

2.12 Interim analysis

Commercially Confidential Information

2.13 Data monitoring committee

This study will include an external Data Monitoring Committee (DMC). The DMC will function independently of site investigators participating in the study. The DMC will assess the progress of the EXPLAIN-AD treatment cohorts and the safety data. The Sponsor will continue, modify, or terminate investigational agents based on these reviews.

The analysis will be conducted by the trial statistician and programmer(s).

Specific details regarding composition, responsibilities, data monitoring, meeting frequency, and documentation of DMC reports, minutes, and recommendations are described in the Safety Surveillance Plan for each agent and the DMC Charter.

3 Sample size calculation

3.1 Primary endpoint(s)

For the primary endpoint, change in NTB total z-score at week 24, data from 36 evaluable participants per arm (investigational agent and placebo) provides ~80% power to detect a statistically significant difference between the two groups at a 1-sided α=0.10 when the true standardized mean difference is 0.5. The effect size of 0.5 corresponds a moderate treatment effect which is thought to represent a clinically meaningful improvement at 24 weeks. To account for 15% dropout rate, approximately n=86 participants randomized in a 1:1 ratio would be needed to address the primary objective. These calculations take into account results of the published study by Frölich et al. (2019), where the standard deviation of the change in NTB total z-score was estimated to be ~0.38 for both active and placebo groups.

For smaller true effect sizes, the power is lower (Table 3-1 below).

Table 3-1 Sensitivity of power to changes in assumptions

SD (common to 2 arms)	True mean difference (Δ)	True effect size (Δ/SD)	alpha (1-sided)	Power
0.38	0.19	0.5	10%	80%
0.38	0.152	0.4	10%	66%
0.38	0.114	0.3	10%	50%
0.38	0.095	0.25	10%	41%

It is assumed that n=86 participants are randomized in a 1:1 ratio and the dropout rate is 15% (thus we have 72 evaluable participants, 36 per arm)

3.2 Secondary endpoint(s)

For the secondary endpoint, PET CCI , a comparison between an investigational agent and placebo will be based on data from approximately 10 participants per arm. Assuming inter-participant variability of 20%-25%, the power to detect statistically significant treatment difference (using 1-sided α=-0.05) is shown in Table 3-2 below. For the 25% true mean reduction in microglial activation due to an anti-inflammatory agent vs. placebo, there is at least 87% power (assuming coefficient of variation is 25% or less).

Table 3-2 Power for treatment comparison using log-transformed PET CCI at week 12

Coefficient of variation (CV)	True mean reduction (active vs. placebo)	α (1-sided)	Power
0.2	30%	5%	99%
0.2	25%	5%	93%
0.2	20%	5%	78%
0.25	30%	5%	93%
0.25	25%	5%	81%
0.25	20%	5%	62%

The assumed sample size is n=20 (10 per arm)

4 Change to protocol specified analyses

The study was CCI stopped. Therefore, the IA planned after 24 patients complete PET TPSO at 24 weeks will not be conducted.

5 Reference

ICH E9(R1) Harmonized Guideline: addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. Final version on 20 November 2019.