Official Title: A Phase III, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group, Multicenter, Efficacy and Safety Study of Gantenerumab in Patients With Mild Alzheimer's Disease: Part 1

A Phase III, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group, Multicenter, Efficacy and Safety Study of Gantenerumab in Patients With Mild Alzheimer's Disease; Part II: Open-Label Extension for Participating Patients

NCT Number: NCT02051608

Document Date(s): SAP Version 1: 23-March-2018

SAP Version 1: 31-May-2021

STATISTICAL ANALYSIS PLAN

TITLE: **A PHASE III, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-GROUP, MULTICENTER, EFFICACY AND SAFETY STUDY OF GANTENERUMAB IN PATIENTS WITH MILD ALZHEIMER'S DISEASE: PART 1**

PROTOCOL NUMBER: WN28745

STUDY DRUG: Gantenerumab (RO4909832)

VERSION NUMBER: 1.0

IND NUMBER: 102,266

EUDRACT NUMBER: 2013-003390-95

SPONSOR: F. Hoffmann-La Roche Ltd

PLAN PREPARED BY: [redacted]

DATE FINAL: 23 March 2018

STATISTICAL ANALYSIS PLAN APPROVAL

[redacted], [redacted] _____ Date

CONFIDENTIAL

This is an F. Hoffmann-La Roche Ltd document that contains confidential information. Nothing herein is to be disclosed without written consent from F. Hoffmann-La Roche Ltd.

TABLE OF CONTENTS

LIST OF APPENDICES

LIST OF ABBREVIATIONS

Abbreviation	Definition
Aβ	Amyloid β
AD	Alzheimer's disease
ADAS-Cog-13	Alzheimer's Disease Assessment Scale Cognition 13
ADCS-ADL	Alzheimer's Disease Cooperative Study−Activities of Daily Living Inventory
ADL	Activity of Daily Living
AE	Adverse Event
APOE	Apolipoprotein E
ARIA	Amyloid−Related Imaging Abnormalities
ARIA-E	Amyloid−Related Imaging Abnormalities-Oedema/Effusion
ARIA-H	Amyloid−Related Imaging Abnormalities-Hemosiderin Deposition
AUC	Area Under the Curve
CDR	Clinical Dementia Rating
CDR-GS	Clinical Dementia Rating-Global Score
CDR-SB	Clinical Dementia Rating-Sum of Boxes
Cmax	Maximum concentration
Cmin	Minimum concentration
CSF	Cerebrospinal Fluid
CSR	Clinical Study Report
CRF	Case Report Form
C-SSRS	Columbia-Suicide Severity Rating Scale
ECG	Electrocardiogram
iDMC	Independent Data Monitoring Committee
ITT	Intent To Treat
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed Model Repeated Measure
MMSE	Mini Mental State Examination
MRI	Magnetic Resonance Imaging
NPI	Neuropsychiatric Inventory

OLE	Open-Label Extension
PK	Pharmacokinetic
PT	Preferred Term
pTau	Phosphorylated tau
Q1	First quartile
Q3	Third quartile
Q4W	Every 4 weeks
SAP	Statistical Analysis Plan
SD	Standard Deviation
STREAM	Standard Reporting and Analysis Modules
Tmax	Time at maximum concentration
tTau	Total Tau

1. BACKGROUND

This document describes the statistical data that will be reported in the Clinical Study Report (CSR) for double blind treatment (Part 1) of Study WN28745 and will focus on the statistical methodology underlying the report. The efficacy and safety endpoints that will be the basis for comparing treatments will be defined in full in this document along with the populations of patients that are to be used in the analyses.

Dosing in study WN25203, investigating gantenerumab in prodromal Alzheimer's Disease (AD), was stopped in December 2014 due to a futility analysis. Consequently, recruitment into study WN28745 was stopped early. The final patient was randomized into study WN28745 on the 9th December 2015. Patients active in study WN28745 continued to receive study treatment until they discontinued from the study, finished the study, or transferred into the open label extension (OLE).

Although there will be reporting of study data in a variety of contexts beyond that of the CSR, such as the independent Data Monitoring Committee (iDMC), this document will not cover those contexts.

The description of layouts for the CSR outputs, the details about the underlying analysis datasets and programs, and the linking of production outputs to sections in the CSR are not within the scope of this document and will be covered in separate documents.

The language used in this statistical analysis plan (SAP) supersedes that in the protocol and protocol synopsis.

2. STUDY DESIGN

WN28745 is a Phase III, multicenter, randomized, double-blind, placebo-controlled, parallel group study to evaluate the safety and efficacy of gantenerumab in patients with mild AD. In part 1 of the study, patients were randomized to receive gantenerumab or placebo in a 1:1 ratio. Patients randomized to gantenerumab received 105mg every 4 weeks (Q4W) for 24 weeks and 225mg from Week 28 to the end of the study provided no amyloid related imaging abnormality-oedema (ARIA-E) and no more than one new amyloid related imaging abnormality-microbleed/hemosiderosis (ARIA-H) were seen on the post Week 24 dose MRI scan. Treatment assignment was stratified by

Apolipoprotein E (*APOE*) genotype (presence or absence of ε4 allele), anti-dementia medications (present or absent) at baseline, and geographic region of the study center (Europe or Rest of World).

The study initially consisted of three phases (Part 1):

1. A screening period of up to 8 weeks.
2. A double-blind treatment period of 100 weeks. A final efficacy and safety assessment to be performed 4 weeks following the patient's last dose (Week 104).
3. A 52 week follow-up period.

The WN28745 protocol was amended in October 2015, following a pre-planned futility analysis in study WN25203, to allow for higher doses of gantenerumab to be studied in Part 2: an OLE. If and when the amended protocol was approved in each country, patients could transition from the main study to the open label extension. This SAP covers data from Part 1 only.

2.1 PROTOCOL SYNOPSIS

The Protocol Synopsis and Schedule of Assessments can be found in the study protocol.

2.2 ANALYSIS TIMING

The analysis covered in this SAP will include all data from Part 1 for all patients enrolled in study WN28745.

For patients who did not transition into the OLE this will include all data from all visits up to and including the final follow-up visit (Week 152). For patients who transitioned into the OLE, data from all visits up to and including the Week 104/early termination visit (Follow up 1), or last visit prior to first OLE dose, will be included. No data from the OLE will be used. Time windowing will be used to ensure that data from all Week 104/ early termination visits are summarized at the correct time point.

Due to the early transition of this study to an OLE, many patients will not have reached the primary analysis time point as defined in the protocol (Week 104). Consequently, all

analysis results (descriptive and MMRM) will be summarized for Week 48, Week 72 and Week 104.

An interim analysis was originally planned to occur when 50% of patients had completed 104 weeks of double-blind treatment. Due to stopping recruitment into the study and the protocol amendment leading to OLE transition, no such interim analysis will be performed.

2.3 OUTCOME MEASURES

For all questionnaire-based outcome measures, scores from the electronic capture system (ie. Bracket tablet) should be used wherever possible.

2.3.1 Primary Efficacy Outcome Measures

The co-primary efficacy outcome assessments include one measure of cognition and one of function: the difference between gantenerumab and placebo in mean change from baseline to Week 104 in Alzheimer's Disease Activity Scale – Cognitive 13 item sub-scale (ADAS-Cog13) and Alzheimer's Disease Cooperative Study – Activities of Daily Living (ADCS-ADL).

2.3.1.1 Alzheimer's Disease Assessment Scale–cognition (13 item sub-scale)

ADAS-Cog is a cognitive assessment designed to measure the severity of some of the most important symptoms of AD, including impairment of memory and language. The 13-item total score is calculated using the sum of all the 13 item scores (see Appendix 1 for more information). This scale is administered electronically and if any item score is missing or is invalid, the total score will be set to missing. The 13-item ADAS-cog total score ranges from 0 to 85. Higher scores indicate greater impairment.

2.3.1.2 Alzheimer's Disease Cooperative Study – Activities of Daily Living (ADCS-ADL).

The ADCS-ADL is the scale most widely used to assess functional outcome in patients with AD. The ADCS-ADL covers both basic ADL (e.g., eating and toileting) and more

complex ADL or instrumental ADL (e.g., using the telephone, managing finances, preparing a meal). The ADCS-ADL consists of 32 items (questions and sub-questions) each with multiple choice answers. For example, questions with three possible responses often have the following response structure (3 points = without supervision or help, 2 points = with supervision, 1 point = with physical help). Total score is calculated by summing all sub-scores. The ADCS-ADL total score ranges from 0 - 78 with lower values indicating greater impairment. The basic ADL (bADL) score is calculated by summing questions 1 to 5 and 6b inclusive. The instrumental ADL (iADL) score is calculated by summing questions 6a and 7 to 23 inclusive. This scale is administered electronically with no option to miss a question. Therefore, missing data will only be present if the entire scale is missing. However, due to a translation error in Spanish for Spain, question 18 will be set to missing for all patients from Spain. This will allow the bADL score to be calculated but the iADL and total scores will be missing for these patients.

2.3.2 Secondary Efficacy Outcome Measures

The change from baseline will be assessed for the following secondary efficacy outcome measures.

2.3.2.1 Clinical Dementia Rating–Global Score

The Clinical Dementia Rating global (CDR-global) score is being used in the study as a measure of dementia severity. The CDR-global score is calculated on the basis of Washington University's CDR assignment algorithm. Investigators enter scores from each of the six categories: memory, orientation, judgment and problem solving, community affairs, home and hobbies and personal care ("box scores"). Memory (M) is considered the primary category, and all others are secondary. The CDR-global score is calculated using a calculator available at http://www.biostat.wustl.edu/~adrc/cdrpgm/index.html. Details of how the global score is derived are given in Appendix 1. The electronic administration of this scale does not allow missing data.

2.3.2.2 Clinical Dementia Rating – Sum of Boxes

The CDR-SB score is obtained by summing each of the domain box scores. Thus the CDR-SB score for a patient ranges from 0 to 18 with a higher score indicating greater impairment. The electronic administration of this scale does not allow missing data.

2.3.2.3 Neuropsychiatric Inventory

The Neuropsychiatric Inventory (NPI) is being used in this study as a measure of the behavioral and neuropsychological symptoms of AD. The NPI is an informant-based instrument in which twelve behavioral domains (see below) are evaluated on a 3-point scale for severity (1-mild, 2-moderate, 3-marked) and a 4-point scale for frequency (1-occasionally, 2-often, 3-frequently, 4-very frequently). A behavioral domain is scored 0 if AD symptoms in the domain are absent. The score for each domain is calculated as the frequency score multiplied by the severity score. The 10 item total score and 12 item total score will be summarized. The two domains not included in the 10 item score are indicated with a * in the list below. The total scores are calculated as the sum of all domain scores with any missing sub-scores set to 0.

- Delusions
- Hallucinations
- Agitation or aggression
- Depression or dysphoria
- Anxiety
- Elation or euphoria
- Apathy or indifference
- Disinhibition
- Irritability or lability
- Motor disturbance
- Nighttime behaviors*
- Appetite and eating*

2.3.2.4 Mini Mental State Exam Total Score

The Mini Mental State Exam (MMSE) is being used in the study as a measure of cognition. The MMSE comprises 11 items assessing a patient's mental status and the individual's general level of impairment in five areas: orientation, short-term memory retention, attention, short-term recall, and language. The total score is calculated by summing scores of all 11 items. If any item score is missing or invalid, the item score will

be set to 0. The total score ranges from 0 – 30 with lower values indicating greater impairment.

2.3.3 Exploratory Efficacy Outcome Measures

2.3.3.1 Time to clinically evident decline

Time to clinically evident decline, as determined by:

- Confirmed decline of ≥ 2 points (at two consecutive visits) on the MMSE, AND
- Loss of ≥ 1 points on one or more basic ADL, as measured on the ADCS-ADL

  OR
- Loss of ≥ 2 points on one or more iADL, as measured on the ADCS-ADL

2.3.4 Safety Outcome Measures

- Incidence and nature of MRI safety findings: ARIA-E and ARIA-H
- Incidence, nature, and severity of serious adverse events
- Incidence, nature, and severity of adverse events
- Incidence of treatment discontinuations due to adverse events
- Mean changes in clinical laboratory tests from baseline over time and incidence of abnormal laboratory values and abnormal laboratory values reported as adverse events
- Mean changes in ECG assessments from baseline over time and incidence of abnormal ECG assessments
- Incidence of anti-gantenerumab antibodies
- Physical and examination abnormalities
- Mean change in vital signs assessment from baseline over time and incidence of abnormal vital signs measurements

- Suicidal ideation, suicidal behavior, and self-injurious behavior without suicidal intent, as determined using the Columbia−Suicide Severity Rating Scale (C-SSRS)

2.3.5 Biomarker Outcome Measures

- Change from baseline in t-tau levels in cerebrospinal fluid (CSF) measured by the Elecsys® Total-Tau immunoassay
- Change from baseline in p-tau levels in CSF measured by the Elecsys® Phospho-Tau immunoassay
- Change from baseline in $A\beta_{1-42}$ levels in CSF measured by the Elecsys® β-Amyloid (1-42) immunoassay
- Change from baseline in $A\beta_{1-40}$ levels in CSF measured by the Elecsys® β-Amyloid (1-40) immunoassay
- Change from baseline in the $A\beta_{1-42}$ / $A\beta_{1-40}$ ratio in CSF
- Change from baseline in the pTau / tTau ratio in CSF
- Change from baseline in the pTau / $A\beta_{1-42}$ ratio in CSF
- Change from baseline in $A\beta_{1-42}$ levels measured by PET scan
- The change from baseline in hippocampal volume measured by MRI scan
- The change from baseline in whole brain volume measured by MRI scan
- The change from baseline in ventricular volume measured by MRI scan

2.3.6 Pharmacokinetic Outcome Measures

- Primary population PK parameters (e.g., apparent total CL [CL/F] and apparent volume of distribution [V/F]) and relevant covariates as necessary to describe the plasma gantenerumab concentration−time course
- Secondary population estimates of plasma gantenerumab exposure at steady state to include peak plasma concentration (Cmax), time to peak concentration (Tmax), trough plasma concentration (Cmin), and AUC

2.3.7 Protocol Deviations

- Major protocol deviations summarized by four main categories (inclusion criteria, exclusion criteria, medication and procedural).

2.4 DETERMINATION OF SAMPLE SIZE

The sample size for this study was planned to be 500 patients per arm (Gantenerumab and Placebo). This calculation was based on the following assumptions:

Test	Two-sided
Alpha	0.05
Power	80%
Expected difference between treatment arms in absolute change from baseline at Week 104 in ADAS-Cog 13 (SD)	2.3 (10.4)
Expected difference between treatment arms in absolute change from baseline at Week 104 in ADCS-ADL (SD)	2.5 (10.9)
Expected withdrawal rate	30%

However, due to the conversion of study WN28745 to an OLE, recruitment into Part 1 of the study was stopped early. Final numbers recruited, enrolled and dosed in study WN28745 will be presented. Number of patients who completed the study, discontinued from the study and transferred to the OLE will also be summarised.

3. STUDY CONDUCT

3.1 RANDOMIZATION ISSUES

To maintain a balanced number of patients enrolled in each treatment arm, randomization was stratified by *APOE* status (presence or absence of ε4 allele) and anti-dementia medications (present or absent) at baseline, and geographic region.

3.2 DATA MONITORING

The incidence and nature of adverse events, serious adverse events, ARIA-E and ARIA-H, and laboratory abnormalities has been assessed on a regular basis by an un-blinded independent Data Monitoring Committee (iDMC). The iDMC charter describes the roles of the iDMC and the frequency and conduct of the planned review meetings.

4. STATISTICAL METHODS

In the following sections, for all continuous variables for which descriptive statistics are indicated, the following statistics will be reported: the number of observations, the mean, median, standard deviation, and minimum and maximum. The 25th and 75th percentiles (Q1 and Q3) will also be reported for selected tables.

In all analysis, baseline is defined as the last available assessment up to and including the first day of study drug intake in the double-blind period (Part 1).

4.1 ANALYSIS POPULATIONS

4.1.1 Intent-to-Treat Population

The intent-to-treat (ITT) population will include all randomized patients. The ITT analysis will be performed by randomized treatment.

4.1.2 Safety Population

The safety population will consist of all patients who received at least one dose of study drug, regardless of whether the patient withdrew prematurely or not. All safety data will be analyzed according to study drug actually received. Any patients who have received at least two doses of gantenerumab will be summarized as having received gantenerumab.

The safety population will be the primary population for all analyses of primary and secondary outcome variables.

4.1.3 CSF Intent-to-Treat Population

The CSF intent-to-treat (CSF-ITT) population will include all patients in the ITT who had at least one post-baseline CSF measure of $A\beta_{1-40}$, $A\beta_{1-42}$, pTau or tTau. The CSF ITT analysis will be performed by randomized treatment.

4.1.4 MRI Intent-to-Treat Population

The MRI intent-to-treat (MRI-ITT) population will include all patients in the ITT who had at least one post-baseline volumetric MRI scan or rs-fMRI scan. The MRI ITT analysis will be performed by randomized treatment.

4.1.5 Safety Follow-Up Population

The safety follow-up population will include all patients who did not enter the OLE who had a vital signs visit date, AE start date or study completion or discontinuation date greater than 28 days after their last intake of study-drug in double-blind. It will also include all patients who entered the OLE whose first dose in the OLE was greater than 28 days after their last dose in double-blind. The time each patient spent in follow-up is also defined by these dates.

4.2 ANALYSIS OF STUDY CONDUCT

4.2.1 Disposition of Patients

The number of patients randomized, and the number and percentage of patients who prematurely withdrew from the study (including the reasons for discontinuation and the distribution of these discontinuations by time-windowed visit) will be summarized using descriptive statistics.

4.3 ANALYSIS OF TREATMENT GROUP COMPARABILITY

4.3.1 Demographic Data and Baseline Characteristics

Baseline comparability of the placebo and gantenerumab treatment groups will be assessed for the Safety and ITT populations. The comparisons will be made on patient demographic and baseline variables, including the following:

- Sex
- Age
- Race
- Years of education
- Weight
- Height

- Body mass index
- *APOE* status
- MMSE, total score at baseline
- ADAS-cog (13 items) at baseline
- ADCS-ADL, total score at baseline
- CDR-SB at baseline
- CDR-Global score at baseline
- NPI at baseline

All comparisons will be based on descriptive statistics.

4.3.2 Previous and Concomitant Medications and Medical History

A patient's medical history is defined as a record of all medical events (for example: procedures and diseases) that start and finish prior to the date of the first dose of study treatment. Similarly, medications a patient has received with an end date prior to the date of the first dose of study treatment will be classified as previous medications.

Any medication ongoing on the date of the first dose of study treatment is classified as previous-concomitant, while any medication starting after the first dose of study treatment is classified as concomitant. Any medical event occurring after the first dose of study treatment is classified as an AE and handled according to the rules in Section 4.5.2.

If the medical event (for example: treatment, procedure or disease) does not have complete or valid start and end dates, the following rules will be applied. The reference date is defined as the date of the first dose of study treatment.

- If the medical event start date is (a) completely missing, or (b) only has the year specified and this is the same as the year of the reference date, or (c) only has the month/year specified and these are the same as the month/year of the reference date, then the medical event will be assumed to have started before the reference date.
- If the medical event end date is (a) completely missing, or (b) only has the year specified, and this is the same as the year of the reference date, or (c) only has the

month/ year specified, and these are the same as the month/ year of the reference date, then the medical event will be assumed to have ended after the reference date.

- Otherwise, partial medical start and end dates (month/year or year only) will be compared with the corresponding part of the reference date to determine whether strictly before or strictly after the reference date (no other possibility exists, because this is covered by the two previous bullet points).

All listings and summaries will include only the preferred term or super class term.

4.4 EFFICACY ANALYSIS

Efficacy analysis will use time-windowed visits details of which are given in the DAP Module 3. In brief, time window borders are in the middle between scheduled visit times. If more than one visit falls in a window, the visit closest to the per-protocol visit date is used. The baseline visit window includes all assessments performed on Day 1 or before and consequently, the time window for the first post-baseline visit begins on study day 2. Therefore, data collected at an early termination visit will be summarized at the appropriate time in the trial and not at Week 104.

Not all patients will reach the primary time point (Week 104) as the trial was converted to an OLE. No imputation will be performed for individuals who did not reach Week 104. However, summaries by time-windowed visit will be provided to allow comparisons between arms at earlier time points.

There may be instances when different questionnaires are performed on different study days within the time window of one visit. In such cases, the earlier study day will be used for the time to clinically evident decline endpoint.

4.4.1 Primary Efficacy Endpoint

The co-primary efficacy outcome measures are:

- The difference between Gantenerumab and Placebo treated patients in the mean change in ADAS-Cog 13 from baseline (Day 1) to Week 104 in the population of all randomized patients.

- The difference between Gantenerumab and Placebo treated patients in the mean change in ADCS-ADL from baseline (Day 1) to Week 104 in the population of all randomized patients.

For each of the co-primary endpoints, the absolute values, change from baseline and percentage change from baseline will be summarized by time-windowed visit.

In addition, the co-primary endpoints will also be analysed using a mixed model with repeated measures (MMRM). The model will use an unstructured variance – covariance matrix for the repeated measures allowing the inclusion of individuals with incomplete data. The following variables will be included in the modelling:

Dependant variable	ADAS-Cog 13 or ADCS-ADL
Independent variables	Treatment (placebo or Gantenerumab) *APOE* status (carrier or non-carrier) Baseline ADAS-Cog 13 or ADCS-ADL score Time (weeks relative to first dose in study) Treatment (placebo or Gantenerumab) by time (weeks relative to first dose in study) interaction

Patient, treatment and time will be treated as class variables. All data from all patients (as defined in Section 2.2) will be used, including data from Week 104/early termination visits.

A treatment-by-time interaction contrast will be used to estimate the difference between gantenerumab and placebo in the mean change from baseline at each visit. Least square means, standard errors, 95% confidence intervals will be reported for gantenerumab relative to placebo.

4.4.2 Secondary Efficacy Endpoints

The absolute values, change from baseline and percentage change from baseline in the continuous secondary efficacy endpoints listed in Section 2.3.2 will be summarized by time-windowed visit.

4.4.2.1 Biomarker Outcome Measures

- Change in t-tau levels in CSF from baseline at Week 104
- Change in p-tau levels in CSF from baseline at Week 104
- Change in $A\beta_{1-42}$ levels in CSF from baseline at Week 104
- Change in $A\beta_{1-40}$ levels in CSF from baseline at Week 104
- Change in $A\beta_{1-42}$ / $A\beta_{1-40}$ ratio in CSF from baseline at Week 104
- Change in pTau / tTau ratio in CSF from baseline at Week 104
- Change in pTau / $A\beta_{1-42}$ ratio in CSF from baseline at Week 104
- Change in $A\beta_{1-42}$ levels as measured by PET scan from baseline at Week 104
- The change from baseline in hippocampal volume at Week 104
- The change from baseline in whole brain volume at Week 104
- The change from baseline in ventricular volume at Week 104

4.5 SAFETY ANALYSES

Descriptive statistics will be used to analyze all safety data. The outputs defined in this section will be produced using the safety analysis population, unless otherwise specified.

For all safety analyses, Day 1 will be defined as the start date of double-blind randomized treatment.

4.5.1 Exposure of Study Medication

Exposure to study drug will be summarized using the following descriptive statistics:

- Duration of treatment (defined as the time from the date of first treatment to date of last treatment in Part 1 of the study)
- Duration of treatment at the 105mg dose for patients whose highest dose was 105mg
- Duration of treatment at the 105mg dose for patients who reached the 225mg dose
- Duration of treatment at the 225mg dose for patients who reached the 225mg dose
- Cumulative dose
- Total patient years on gantenerumab: Time from first dose in double-blind until 28 days after last dose in double-blind or first dose in OLE (whichever occurs first).

- Total patient years on 105mg gantenerumab
- Total patient years on 225mg gantenerumab

4.5.2 Adverse Events

Adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version that is current at the time of the analysis (Version 20). For each treatment group, the frequency of each adverse event preferred term will be defined as the number of patients experiencing at least one occurrence of the event. As patients will have had different amounts of time in the study, the incidence rate will be calculated as the frequency count divided by the total number of patient years exposure for the patients in the population and treatment group specified. Each table will also present the overall number of patients experiencing at least one adverse event and the total number of adverse events reported. Patient listings will contain preferred terms and comments for each event. In summary tables, adverse events will be sorted by body system (in decreasing order of overall incidence), then by preferred term (in decreasing order of overall incidence).

The following safety information will be summarized:

- Adverse events
- Serious adverse events
- Study drug–related adverse events
- Study drug–related serious adverse events
- Adverse events leading to discontinuation of study treatment
- Adverse events leading to dose reduction
- Adverse events leading to no up-titration from 105mg to 225mg at Week 28
- Adverse events leading to dose held

The following data handling rules will be applied for all adverse event summary tables:

- Multiple occurrences of the same adverse event in 1 patient will be counted only once.
- Missing preferred terms (PTs) will be presented as the investigator's term, preceded by "NO_PT_Found" in all outputs. Missing super-class terms (SCTs) will be

presented as “<NO_SCT_FOUND>” in all outputs. All cases of missing PTs should be resolved at the time of the final database closure.

- Events that are missing both onset and end dates will be considered to have started after the first dose of study treatment and the duration will be set to missing.
- If the onset date is missing, and the end date is on or after the first dosing date or unresolved or missing, then the event will be to have started after the first dose of study treatment.

The following data handling rules will also be applied for specific tables:

- An adverse event will be included in the summary table of adverse events leading to study drug discontinuation if the “action taken with blinded gantenerumab” drop-down menu on the Adverse Event CRF is checked “drug withdrawn”.
- In the summary table of adverse events by intensity, if a patient has more than one occurrence of an event, the event with the most severe intensity will be counted. If the intensity of an adverse event is missing, then the adverse event will be included only in the total number of events column, and not in the count of patients with the event by intensity.

For scientific interpretation it is important to separate ARIA-H and ARIA-E events in listing and tables. However, using MedDRA standards, both get mapped to the preferred term 'AMYLOID RELATED IMAGING ABNORMALITIES'. Furthermore, there are some additional terms that the Clinical Science and Safety Science teams deem to be ARIA-E or ARIA-H that have separate MedDRA preferred terms. To overcome this, baskets will be implemented to summarize all ARIA-H events together and all ARIA-E events together. These baskets are defined below:

- ARIA-E
  - Preferred term: ‘VASOGENIC CEREBRAL OEDEMA’ or
  - Lower level terms: ‘AMYLOID RELATED IMAGING ABNORMALITY-EDEMA/EFFUSION’ or ‘AMYLOID RELATED IMAGING ABNORMALITY-OEDEMA/EFFUSION’ or ‘ARIA-E’
- ARIA-H

- Preferred term: ‘CEREBELLAR MICROHAEMORRHAGE’ or ‘CEREBRAL HAEMOSIDERIN DEPOSITION’ or ‘CEREBRAL MICROHAEMORRHAGE’
- Lower level terms: ‘AMYLOID RELATED IMAGING ABNORMALITY-MICROHAEMORRHAGE AND HAEMOSIDERIN DEPOSITS’ or ‘AMYLOID RELATED IMAGING ABNORMALITY-MICROHEMORRHAGES AND HEMOSIDERIN DEPOSITS’ or ‘ARIA-H’

Instances where an AE has MedDRA preferred term ‘AMYLOID RELATED IMAGING ABNORMALITIES’ but does not fit into one of the two baskets above, will be assessed on a case-by-case basis.

In the summary table of adverse events by relationship to study drug, if a patient has more than one occurrence of an event, the most closely related event will be counted. If the relationship of an adverse event is missing, then the adverse event will be included only in the total number of events column, and not in the count of patients with the event by relationship.

4.5.3 MRI Safety Findings and Injection-Site Reactions

The incidence of ARIA-E, ARIA-H and other MRI findings as well as injection-site reactions will be listed and summarized by treatment group. In addition, the MRI data will be summarized by *APOE* genotype and by time-window.

Sites were asked to also capture both ARIAs leading to dose modification and injection site reactions as AEs.

4.5.4 Laboratory Data

Laboratory data will be summarized for each assessment using descriptive statistics of absolute values, change from baseline values and percent change from baseline. The proportion of patients who develop marked abnormalities during study treatment, on the basis of the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE v4.0) and the International Standard for Handling and Reporting of Laboratory Data (COG 2.2) will be summarized.

4.5.5 Vital Signs

Vital signs assessments include systolic blood pressure, diastolic blood pressure, and pulse measured throughout the study. Summaries of vital sign measurements at each assessment time-window using descriptive statistics will be generated for the actual values, the change from baseline and the percent change from baseline.

4.5.6 Electrocardiograms

Descriptive statistics will be generated for the actual values, the change from baseline and the percent change from baseline for the following parameters:

- Heart rate
- PQ (PR) interval
- QRS interval
- QT interval
- RR interval
- QTcB and QTcF

Duration of the QT interval is inversely correlated with the corresponding heart rate, so it should be corrected relative to heart rate to allow a reliable interpretation of changes in the QTc interval independent of the changes in the heart rate. The QT intervals will be corrected for the heart rate using Bazett's and Fridericia's formulae (QTcB and QTcF, respectively) as follows:

- The QTcB value will be calculated as the QT interval (in milliseconds) divided by the squared root of the RR interval (in seconds).
- The QTcF value will be calculated as the QT interval (in milliseconds) divided by the cube root of the RR interval (in seconds).

In some cases, multiple ECGs were performed in error. In these cases, the mean values across the ECGs should be used.

4.5.7 Deaths

Deaths will be listed.

4.5.8 Sub-group analysis

All safety and efficacy analysis will also be performed in the sub-group of patients who were enrolled in Japan and by APOE e4 carrier/non-carrier status.

4.6 INTERIM ANALYSES

No efficacy interim analyses were performed in this study.

Appendix 1
Endpoint Details

ADAS-Cog 13

The item scores for individual items are derived from the following:

- Item 1 - Word recall: Three trials of reading and recall are given. The item score is calculated using the mean number of "total not recalled" recorded on the electronic system for the three trials. The item score ranges from 0 to 10. If the "total not recalled" in any of the three trials is missing, then the item score will be set to missing.
- Items 2 - Following commands: The item score is the number of incorrect responses on commands (i.e., no) for Questions 2a–e on the electronic system. The score ranges from 0 to 5.
- Items 3 - Constructional praxis: The item score is the number of incorrect drawings (i.e., Drawn correct? = no) for Questions 3a–d on the electronic system. The score ranges from 0 to 5 as follows:

  0 = All 4 drawings correct
  1 = 1 form drawn incorrectly
  2 = 2 forms drawn incorrectly
  3 = 3 forms drawn incorrectly
  4 = 4 forms drawn incorrectly
  5 = No figures drawn, scribbles; parts of forms; words instead of forms

- Items 4 - Delayed word-recall task: The item score is "total not recalled" on the electronic system. The score ranges from 0 to 10.
- Item 5: Naming objects and fingers: The item score is derived on the basis of "total incorrect" recorded on the electronic system and determined as follows:

  0 = 0 to 2 incorrect
  1 = 3 to 5 incorrect
  2 = 6 to 8 incorrect
  3 = 9 to 11 incorrect
  4 = 12 to 14 incorrect
  5 = 15 to 17 incorrect

- Item 6 - Ideational praxis: The item score equals the “total incorrect” recorded on the electronic system for this item. The score ranges from 0 to 5.
- Item 7 – Orientation: The item score equals “total incorrect” recorded on the electronic system for this item. The score ranges from 0 to 8.
- Item 8 - Word recognition: The item score equals “total incorrect” recorded on the electronic system when “total incorrect” is ≤ 12 and is 12 if “total incorrect” is > 12. The item score ranges from 0 to 12.
- Item 9 - Remembering test instructions: The item score is derived on the basis of replies recorded on the electronic system and determined as follows:

  0 = None
  1 = Very mild
  2 = Mild
  3 = Moderate
  4 = Moderately severe
  5 = Severe.

- Item 10 - Comprehension: The item score is derived on the basis of replies recorded on the electronic system and determined as in Item 9.
- Item 11 - Word finding difficulty: The item score is derived on the basis of replies recorded on the electronic system and determined as in Item 9.
- Item 12 - Spoken language ability: The item score is derived on the basis of replies recorded on the electronic system and determined as in Item 9.
- Item 13 - Number cancellation: The cancellation score equals the number of target hits (minus) the number of errors (minus) number of times reminded of task, for which:

  0 = Cancellation score > 23
  1 = Cancellation scores 18–22
  2 = Cancellation scores 13–17
  3 = Cancellation scores 9–12
  4 = Cancellation scores 5–8
  5 = Cancellation scores ≤4

CDR GS

The global CDR is derived from the scores in each of the six categories ("box scores") as follows:

a) CDR-global score = M if at least three secondary categories are given the same scores as memory.
b) If at least three secondary categories are on one side of M, then CDR-global score = score of these secondary categories. Except in the case of c.
c) CDR-global score = M when three secondary categories are scored greater than M and two secondary categories are scored less than M, or vice versa.
d) If M = 0.5, CDR-global score cannot be 0, it can only be 0.5 or 1. CDR-global score = 1 if at least three of the other categories are scored ≥ 1.
e) If M = 0, CDR-global score = 0, unless there is impairment (≥ 0.5) in two or more secondary categories, in which case, CDR-global score = 0.5.
f) If there are ties in the secondary categories on either side of M, choose the tied scores closest to M for the CDR-global score. (For example, if M and another secondary category = 3, two secondary categories = 2 and two secondary categories = 1 then CDR-global score = 2).
g) When only one or two secondary categories are given the same score as M, CDR-global score = M, as long as no more than two secondary categories are on either side of M.
h) When $M \geq 1$, the CDR-global score cannot be 0. In such a circumstance, CDR-global score = 0.5 when the majority of secondary categories are 0.

STATISTICAL ANALYSIS PLAN

STUDY TITLE: **A PHASE III, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-GROUP, MULTICENTER, EFFICACY AND SAFETY STUDY OF GANTENERUMAB IN PATIENTS WITH MILD ALZHEIMER'S DISEASE; PART II: OPEN-LABEL EXTENSION FOR PARTICIPATING PATIENTS**

STUDY NUMBER: WN28745

VERSION NUMBER: 1

ROCHE COMPOUND(S): Gantenerumab (RO4909832)

EUDRACT NUMBER: 2013-0033990-95

IND NUMBER: 102266

NCT NUMBER: NCT02051608

PLAN PREPARED BY: [redacted] PhD

STATISTICAL ANALYSIS PLAN APPROVAL

Date and Time(UTC)	Reason for Signing	Name
31-May-2021 15:37:20	Company Signatory	[redacted]

SPONSOR: F. Hoffmann-La Roche Ltd
LEGAL REGISTERED ADDRESS: Grenzacherstrasse 124
4070 Basel, Switzerland

DATE FINAL: See electronic date stamp above

STATISTICAL ANALYSIS PLAN VERSION HISTORY

This SAP was developed based on Roche SAP model document updated on 26 October 2020.

SAP Version	Approval Date	Based on Protocol (Version, Approval Date)
1	|see electronic date stamp on title page| |day month year|	Version 5, 28 February 2020

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation or Term	Description
AD	Alzheimer's Disease
ADA	anti-drug antibodies
ADAS-Cog	Alzheimer Disease Assessment Scale-Cognition
ADCS-ADL	Alzheimer's Disease Cooperative Study–Activity of Daily Living Scale
ADRDA	Alzheimer's Disease and Related Disorders Association
AE	adverse event
APOE	apolipoprotein E
ARIA-E	amyloid-related imaging abnormality – oedema/effusion
ARIA-H	amyloid-related imaging abnormality – microhaemorrhages and hemosiderin deposits
$AUC_{0-\tau}$	area under plasma concentration time curve
BGTS	Barkhof Grand Total Score
CCOD	clinical cutoff date
CDR	Clinical Dementia Rating Scale
CDR-SB	Clinical Dementia Rating – Sum of Boxes
C_{max}	maximum plasma concentration
C_{min}	minimum plasma concentration
COVID-19	Corona Virus Disease 2019
CSF	cerebrospinal fluid
CSR	Clinical Study Report
C-SSRS	Columbia−Suicide Severity Rating Scale
DB	double blind
MedDRA	Medical Dictionary for Regulatory Activities
MMSE	Mini Mental State Exam
MRI	magnetic resonance imaging
MRI-C	MRI review committee-Charter
OLE	open-label extension
PD	pharmacodynamic(s)
PET	positron emission tomography
PK	pharmacokinetic(s)
PT	preferred term
QTcB	QT corrected using Bazett's formula
QTcF	QT corrected using Fridericia's formula

SAE	serious adverse events
SAP	Statistical Analysis Plan
SE	safety-evaluable
SOC	system organ class
SUVR	standard uptake value ratio

1. INTRODUCTION

This document describes the planned statistical analyses that will be reported in the final Clinical Study Report (CSR) for open-label extension (OLE) treatment of Study WN28745. The corresponding protocol version is Version 5.

Study WN28745 was designed as a double-blind, Phase III study (Part I) investigating the effect of 2 years of treatment with subcutaneous (SC) gantenerumab 225 mg versus placebo every 4 weeks (Q4W) on cognition and function in patients with mild Alzheimer's disease (AD). The protocol for this study was amended to an open label extension (OLE) (Part II) to evaluate the safety and tolerability of SC doses of gantenerumab up to 1200 mg Q4W. Enrolment into the double blind (DB) part of the study was halted in December 2015 after 389 patients had been randomized; however, the patients from this study continued dosing while the decision to transition to OLE was made. Enrolment into the WN28745 OLE started on 03 March 2016 and was closed on 31 July 2017; overall, 230 of 389 patients originally randomized into the study entered the OLE part.

The duration of the OLE was up to 5 years. In addition to the initial 2 years in the OLE part, patients were given the option to continue receiving open-label gantenerumab treatment until the end of 2020. Patients who discontinued study drug at any time during the OLE part, or who completed the first 2 years of OLE only were asked to complete follow-up visits at 4 and 16 weeks from their last dose (Follow-Up 1 and 2 respectively)

The description of layouts for the CSR outputs, the details about the underlying analysis datasets and programs, and the linking of production outputs to sections in the final CSR are not within the scope of this document.

The language used in this Statistical Analysis Plan (SAP) supersedes that in the protocol and protocol synopsis.

1.1 OBJECTIVES AND OUTCOME MEASURES

1.1.1 Objectives

The primary objective of the OLE was to evaluate the safety and tolerability of gantenerumab at doses up to 1200mg every 4 weeks, focusing on physical and neurologic examinations, vital signs, blood safety tests, ECGs, and adverse event (AE) monitoring.

The secondary objectives included the following:

- To evaluate the effect of higher doses of gantenerumab on imaging biomarkers (positron emission tomography [PET] and magnetic resonance imaging [MRI]), cerebrospinal fluid (CSF) biomarkers, and on clinical outcome measures (cognition and function) over time
- To explore gantenerumab pharmacokinetics (PK) at higher doses.

1.1.2 Outcome Measures

The references for the assessments and procedures can be found in the study protocol, Sections 4.5.

1.1.2.1 Efficacy Measures

The clinical and cognitive efficacy measures are exploratory. During the first two years of the OLE, they included:

- Alzheimer Disease Assessment Scale-Cognition (ADAS-Cog13),
- Mini Mental State Exam (MMSE),
- Clinical Dementia Rating Scale (CDR)
- Alzheimer's Disease Cooperative Study–Activity of Daily Living Scale (ADCS-ADL).

Only the MMSE was collected during the subsequent years of the OLE.

1.1.2.2 Biomarker Outcome Measures

The biomarker outcome measures for this study were as follows:

- Change from baseline (DB screening visit) to OLE Week 104 in MRI volumetry, as assessed on structural MRI:
  - Change from baseline in hippocampal volume
  - Change from baseline in whole brain volume
  - Change from baseline in cortical thickness
  - Change from baseline in ventricular volume

Change from baseline (DB screening visit) to OLE Week 104 in functional brain connectivity, as measured by resting state-MRI, was an exploratory biomarker outcome measure, and its analysis will not be covered in this SAP.

- Change in brain amyloid load over time using a florbetapir F18 injection, a PET radioligand selective to β-amyloid
- The analysis of the CSF biomarker will not be covered in this SAP.

1.1.2.3 Pharmacokinetic Outcome Measures

The PK outcome measures were as follows:

- Plasma concentrations at each scheduled PK sample time point.

A population PK model previously developed using nonlinear mixed-effects modeling (1105289) will be used to derive individual measures of exposure at steady state such as:

- Maximum plasma concentration (C_{max})
- Minimum plasma concentration (C_{min})
- The area under the plasma concentration-time curve (AUC_T).

1.1.2.4 Safety Outcome Measures

The safety outcome measures were as follows:

- Incidence and nature of MRI safety findings: amyloid-related imaging abnormality – oedema/effusion (ARIA-E) and amyloid-related imaging abnormality – microhaemorrhages and hemosiderin deposits (ARIA-H)
- Incidence, nature, and severity of serious adverse events (SAEs)
- Incidence, nature, and severity of adverse events (AEs)
- Incidence of treatment discontinuations due to AEs
- Mean changes in clinical laboratory tests from the OLE baseline over time and incidence of abnormal laboratory values and abnormal laboratory values reported as AEs
- Mean changes in ECG assessments from baseline over time and incidence of abnormal ECG assessments
- Incidence of anti-gantenerumab antibodies
- Physical and neurologic examination abnormalities
- Mean change in vital signs assessment from baseline over time and incidence of abnormal vital signs measurements
- Suicidal ideation, suicidal behavior, and self-injurious behavior without suicidal intent, as determined using the Columbia−Suicide Severity Rating Scale (C-SSRS).

1.2 STUDY DESIGN

Study WN28745 was a Phase III, multicenter, randomized, double-blind, placebo-controlled, parallel-group study to evaluate the safety and efficacy of gantenerumab in patients with mild AD.

All patients who were actively enrolled in Study WN28745 (i.e., not discontinued from study drug) were invited to participate to the OLE once the study protocol was approved in their country. Overall, 230 patients who had enrolled in Study WN28745 continued to the OLE. DB treatment allocation of patients and apolipoprotein E (APOE) ε4 status were revealed, and the up-titration schedules were determined depending on APOE ε4 status (0ε4 as non-carriers and 1ε4 or 2ε4 as carriers) and whether patients were previously on placebo or active gantenerumab. There were four titration schedules as follows:

- Carrier patients previously on 225 mg gantenerumab
- Non-carrier patients previously on 225 mg gantenerumab
- Carrier patients previously on 105 mg gantenerumab or on placebo
- Non-carrier patients previously on 105 mg gantenerumab or placebo.

See Table 1 for titration schedules in detail.

At every dose increase, MRI confirmation of no symptomatic ARIA-E, no ARIA-E with Barkhof Grand Total Score (BGTS) > 1, and no more than eight ARIA-H cumulatively as per Section 5.1.4 in the study protocol was required. Patients who did not meet the criteria for up-titration either remained on the same dose or interrupted the study treatment until satisfactory resolution of MRI finding (see Section 5.1.4 in the study protocol).

Table 1 Up-Titration Schedule for Open-Label Extension

Previous Treatment Assignment (APOE ε4 status)	OLE [a] Day 1	OLE Week 4	OLE Week 8	OLE Week 12	OLE Week 16	OLE Week 20	OLE Week ≥ 24
225 mg Gantenerumab (Carriers)	450	450	900	900	1200	1200	1200
225 mg Gantenerumab (Non-carriers)	600	600	1200	1200	1200	1200	1200
105 mg Gantenerumab or Placebo (Carriers)	225	225	450	450	900	900	1200
105 mg Gantenerumab or Placebo (Non-carriers)	300	300	600	600	1200	1200	1200

APOE = apolipoprotein E; OLE = Open-Label Extension.

Note: All numerical amounts presented above are in milligrams (mg).

[a] First open-label gantenerumab dose administered following signature of Part 2 ICF.

Patients received open-label gantenerumab for an initial 2 years treatment and up to an additional 3 years beyond the initial 2 years of OLE (5 years of OLE in total), followed by a safety and limited efficacy assessment 4 weeks after the last dose. Patients in the OLE were given the option to continue receiving open-label gantenerumab treatment until the end of 2020. The first dose of open-label gantenerumab was administered after the patient had signed the OLE informed consent form.

At the end of the study, patients were given the option of enrolling in the open-label rollover study (WN41874), aimed at evaluating the safety and tolerability of long-term administration of doses up to 1200 mg of gantenerumab. Patients who did not enroll in Study WN41874 had one additional follow-up visit at 16 weeks after the final dose for safety and limited efficacy assessments.

1.2.1 Treatment Assignment

Patients enrolled in the OLE were not randomized. Depending on their APOE ε4 status and previous treatment assignment (gantenerumab or placebo) in the DB part, OLE patients were assigned to one of the four up-titration dosing schemes presented in Table 1.

1.2.2 Independent Imaging Review Facility

All MRI scans were read and reported by the central reader, in alignment with the processes from both parts of the study. When the central reader identified a new MRI finding, the study center medical staff and the Sponsor were notified, and in the case of relevant new MRI findings as defined in the MRI Review Committee (MRI-C) Charter, the Sponsor, in turn, notified the MRI-C concurrently. The MRI-C was actively involved in the review and had the remit to make recommendations that might deviate from the protocol guidelines, as deemed necessary. The committee reviewed relevant MRI findings as defined in the MRI-C Charter.

1.2.3 Data Monitoring

During the OLE, ARIA findings were monitored on an ongoing basis by the Sponsor's Internal Monitoring Committee (IMC) and an internal MRI-C (iMRI-C). The independent Data Monitoring Committee (iDMC) reviewed relevant safety findings on a regular basis until the majority of patients had reached the target dose. The last iDMC was hold on 10 April 2018. Afterwards, the iMRI-C and the IMC continued to review safety findings on a regular basis.

2. STATISTICAL HYPOTHESES

Not Applicable.

3. SAMPLE SIZE DETERMINATION

All patients who have been randomized and were actively enrolled in the study at the time of the amendment approval in their respective country will be eligible to participate in the OLE. Patients who have been discontinued from the study will not be allowed to enroll in the OLE.

4. ANALYSIS POPULATIONS

All Enrolled Population

All enrolled population includes all patients that were enrolled in the OLE.

Safety Evaluable Population

The safety-evaluable (SE) population includes patients who received of at least one dose of gantenerumab during the OLE. Analysis using this population will be performed by previous randomization, which is the treatment assignment in the DB part, and overall or by actually received up-titration scheme (as needed), and overall.

Safety Population with Post Open-Label Extension Baseline Magnetic Resonance Imaging

The safety population with post-OLE baseline MRI includes patients in the SE population who had at least one post-OLE baseline MRI.

5. STATISTICAL ANALYSES

5.1 GENERAL CONSIDERATION

In the following sections, for all continuous variables for which descriptive statistics are indicated, the following statistics will be reported: the number of observations, the mean, median, standard deviation, and minimum and maximum. The 25th and 75th percentiles (Q1 and Q3) will also be reported for selected tables.

Statistical analyses will primarily focus on safety aspects in the OLE. The efficacy analyses, including clinical/cognitive measures, will be descriptive and strictly exploratory in nature:

- The start of OLE is defined as the date when patients received the first dose of gantenerumab in the OLE.
- The OLE baseline is defined as the last available assessment up to and including the first day of study drug intake in the OLE. Thus, the OLE baseline could be in the DB part.
- The end of the study is defined as the date of the last visit (including the last scheduled follow-up visit according to the study protocol) of the last participant in the study.

5.2 PARTICIPANT DISPOSITION

Descriptive statistics will be used to summarize study conduct measures for the all enrolled population, such as study disposition (including premature withdrawals from treatment), incidence of protocol deviations, percentage of intended dose administered, and incidence of incorrect treatment allocation.

All protocol deviations will be listed. Furthermore, the protocol deviations related to epidemic/pandemic will be listed separately.

5.3 DEMOGRAPHIC DATA AND BASELINE CHARACTERISTICS

Baseline characteristics will be analyzed descriptively for the SE population by previous randomization and by actually received up-titration scheme and overall. The patients' demographics and DB screening or OLE baseline variables include, but are not limited to the following:

- Sex
- Age at DB screening
- Age at OLE baseline
- Race

- Years of education
- Weight
- Height
- Body mass index
- Region
- APOE ε4 carrier status (yes vs no)
- APOE genotype (0ε4, 1ε4, or 2ε4)
- ARIA-E during DB period (yes vs no)
- Cumulative ARIA-H at OLE baseline
- MMSE total score at OLE baseline
- ADAS-cog (13-items score) at OLE baseline
- ADCS-ADL at OLE baseline
- Clinical Dementia Rating – Sum of Boxes (CDR-SB) at OLE baseline
- Clinical Dementia Rating - Global score at OLE baseline

5.4 PREVIOUS AND CONCOMITANT MEDICATION AND MEDICAL HISTORY

The medical history, previous medication and concomitant medication data will be coded using the WHO Drug Dictionary version that is current at the time of the analysis (WHO Drug 1 March, 2021), and will be summarized for the SE populations by previous randomization and overall:

- Medical history is defined as a record of all medical events (for example: procedures and diseases) of a patient that started prior to the date of the first dose of study treatment in the DB part.
- Previous medication is defined as all received medications of a patient that ended prior to the date of the first dose of study treatment in the DB part.
- Concomitant medications are classified into the following four categories.
  - Medications ongoing at first dose in the DB is defined as all received medications of a patient that started prior to and were still ongoing on the date of the first dose of study treatment in the DB part.
  - Medications used after first dose in the DB is defined as all received medications of a patient used after the first dose of study treatment in the DB part.
  - Medications ongoing at first dose in the OLE is defined as all received medications of a patient that were still ongoing on the date of the first dose of study treatment in the OLE.
  - Medications used after first dose in the OLE is defined as all received medications of a patient that started after the date of the first dose of study treatment in the OLE.

If the medical event (for example: treatment, procedure or disease) does not have complete or valid start and end dates, the following rules will be applied. The reference date is defined as the date of the first dose of study treatment.

If the medical event start date is (a) completely missing, or (b) only has the year specified, and this is the same as the year of the reference date, or (c) only has the month/year specified, and these are the same as the month/year of the reference date, then the medical event will be assumed to have started before the reference date.

If the medical event end date is (a) completely missing, or (b) only has the year specified, and this is the same as the year of the reference date, or (c) only has the month/ year specified, and these are the same as the month/ year of the reference date, then the medical event will be assumed to have ended after the reference date.

Otherwise, partial medical event start and end dates (month/year or year only) will be compared with the corresponding part of the reference date to determine whether strictly before or strictly after the reference date (no other possibility exists, because this is covered by the two previous points).

5.5 EFFICACY ANALYSES

Efficacy outcomes of CDR, MMSE, ADAS-Cog13, ADCS-ADL at OLE baseline, OLE Week 52/104 were assessed and collected in the initial two years of the OLE, while only MMSE was assessed and collected every 6 months thereafter. The rating scale ADAS-Cog, ADCS-ADL, CDR and MMSE were captured electronically and transferred to the database directly from the core laboratory. During the OLE additional years, the MMSE was performed as a paper-and-pencil scale and recorded on the eCRF form. Definition and schedules of these measurements can also be found in the Section 4.5.5 and Appendix 2 of the study protocol.

For each of these outcome measures, the total scores and change from the OLE baseline will be summarized by previous randomization and overall at each visit. All efficacy analyses will be descriptive.

5.5.1 Alzheimer Disease Assessment Scale-Cognition

The modified ADAS-Cog with a 13-item total score was used in the study. The 13-item total score is calculated using the sum of all the 13-item scores.

If any item score is missing or is invalid, the total score will be set to missing. The 13-item ADAS-Cog total score has a range of 0−85. Higher scores indicate greater impairment.

5.5.2 Mini Mental State Exam

The MMSE comprises 11 items to assess a participant's mental status and identifies the individual's general level of impairment in five areas: orientation, short-term memory

retention, attention, short-term recall, and language. The total score is calculated by summing the scores of all 11-items; the maximum total score is 30.

If any item score is missing or invalid, the item score will be set to 0. The total score ranges from 0 to 30 with lower values indicating greater impairment.

5.5.3 Clinical Dementia Rating Scale−Global Score

The CDR global score is calculated based on Washington University's CDR assignment algorithm (https://biostat.wustl.edu/~adrc/cdrpgm/index.html). Investigators enter scores from each of the six categories ("box scores"), and the CDR global score is calculated using a calculator available at http://www.biostat.wustl.edu/~adrc/cdrpgm/index.html.

The global CDR score is derived from the scores in each of the six categories ("box scores") and is determined as follows (Memory [M] is considered the primary category, and all others are secondary):

- CDR = M if at least three secondary categories give the same scores as memory.
- CDR = score of majority of secondary categories with at least three secondary categories given a score greater or less than the memory score on one side of M, except in the case of CDR = M (see the subsequent point).
- CDR = M when three secondary categories are scored on one side of M and two secondary categories are scored on the other side of M.
- If M = 0.5, then CDR cannot be 0 and it can only be 0.5 or 1; CDR = 1 if at least three of the other categories are scored ≥ 1.
- If M = 0, then CDR = 0; if there are impairments (≥ 0.5) in two or more secondary categories, then CDR = 0.5.
- When there are ties in the secondary categories on one side of M, the chosen tied score for CDR is the closest score to M (e.g., If M and another secondary category = 3, two secondary categories = 2, and two secondary categories = 1, then CDR = 2).
- When only one or two secondary categories are given the same score as M, then CDR = M as long as no more than two secondary categories are on either side of M.
- When M ≥ 1, then CDR cannot be 0; in such a circumstance, CDR = 0.5 when the majority of secondary categories are 0.

5.5.4 Clinical Dementia Rating – Sum of Boxes

The CDR-SB score is obtained by summing each of the domain box scores. Thus, the CDR-SB score for a patient ranges from 0 to 18 with a higher score indicating greater impairment. The electronic administration of this scale does not allow missing data.

5.5.5 Alzheimer's Disease Cooperative Study – Activities of Daily Living

The ADCS-ADL is the scale most widely used to assess functional outcome in patients with AD. The ADCS-ADL consists of 32 items (questions and sub-questions) each with multiple choice answers. For example, questions with three possible responses often have the following response structure (3 points = without supervision or help, 2 points = with supervision, 1 point = with physical help). Total score is calculated by summing all sub-scores. The ADCS-ADL total score ranges from 0–78 with lower values indicating greater impairment.

This scale is administered electronically with no option to miss a question. Therefore, missing data will only be present if the entire scale is missing. However, due to a translation error in Spanish for Spain, Question 18 will be set to missing for all patients from Spain. Total scores will be missing for these patients.

5.6 PHARMACOKINETIC AND PHARMACODYNAMIC ANALYSES

5.6.1 Pharmacokinetic and Anti-Drug Antibody Analysis

Plasma concentration data for gantenerumab will be descriptively summarized by nominal time/visit including, but not limited to the arithmetic mean, median, standard deviation and range.

A population PK model previously developed using nonlinear mixed-effects modeling (1105289) will be used to derive individual measures of exposure including C_{max}, C_{min}, and AUC_T at steady state. These parameters will be descriptively summarized including, but not limited to, arithmetic mean, median, standard deviation, range, and coefficient of variation.

Data on anti-drug antibodies (ADAs) will be summarized from the first dose of gantenerumab to the start of OLE and from the start of OLE to the end of study. Furthermore, their effect on gantenerumab plasma PK will be assessed graphically, as appropriate.

5.6.2 Pharmacodynamic and Exploratory Biomarker Analysis

The PET substudy (WN28745-PET) was initiated in the double-blind period. Florbetapir^{18}F (Amyvid TM) was the tracer used for all patients in this substudy. Brain amyloid imaging was considered a pharmacodynamic (PD) measurement of potential effects of gantenerumab on amyloid load in the brains of patients with mild AD. There were up to 4 assessments included in the analysis: OLE Baseline, OLE Week 52, OLE Week 104 and OLE Week 156. Note that the OLE baseline measurement should occur within 9 months prior to the start of OLE. All patients participating in the DB part of the PET substudy (approximately 100 patients in the WN28745 PET substudy) were invited to continue in the OLE part to assess changes of brain amyloid load over time with gantenerumab treatment. No patients started the PET substudy in the OLE.

A prespecified standard uptake value ratio (SUVR) method, identical to the method report for analyses from the DB part of the study, was used. SUVR values were also converted to centiloid values to provide a more intuitive interpretation of amyloid reduction. Centiloid zero is the mean amyloid burden for a typical population of young healthy controls, 100 centiloid is the typical mean of a population with AD, and 24 centiloid is consistent with the diagnostic amyloid positivity threshold (Klunk et al. 2015; Navitsky et al. 2018).

The equation for conversion of florbetapir SUVR to centiloid is specified using the mean cerebellar gray reference region (Klein et al. 2019):

$$^{FBP}CL = 184.118 \times {}^{FBP}SUVR - 233.718$$

^{FBP}CL = florbetapir centiloid; $^{FBP}SUVR$ = florbetapir standard uptake value ratio.

For the SUVR and Centiloid measures, the value and change from the OLE baseline will be summarized by visit and by previous randomization, and overall. A mixed model for repeated measurement (MMRM) may also be used if applicable.

5.6.3 Magnetic Resonance Imaging Volumetric Measures

The ventricular volume and the percentage change from the DB screening in hippocampal volume, whole brain volume and cortical grey matter volume were assessed at OLE Week 52, OLE Week 104 and OLE Week 152. These volumetric measures will be summarized by previous randomization and overall using descriptive statistics and analyzed by mixed model of repeated measurement if applicable.

- Ventricular volume is a validated automated measurement of the volume of the lateral ventricles. The method is based on anatomic non-linear image matching and labeling with a minimum deformation template library (Fonov et al. 2010, 2011) and local patch-based label fusion (Coupe et al. 2011).
- The percentage change in hippocampal volume measurement relies on a locally developed longitudinal image-processing pipeline that builds a subject-specific MRI average template using all available time points.
- The percentage change in whole brain volume is measured based on Jacobian integration using locally developed JacobianAtrophy software (Nakamura et al. 2013). A T1w image with the same image acquisition parameters must exist for the two timepoints over which the percent whole brain volume change is to be calculated.
- The percentage change in cortical grey matter volume is calculated within the cortical grey matter volume mask using the JacobianAtrophy software described above for the percent brain volume change

See the MRI charter for details of the MRI volumetric measures.

5.7 SAFETY ANALYSES

Descriptive statistics will be used to analyze all safety data in the SE population, unless otherwise specified. All analyses will be summarized by previous randomization and overall, unless otherwise specified.

5.7.1 Exposure to Study Drug

Exposure to study drug information will be descriptively summarized by previous randomization and overall as follows:

- Treatment duration (in weeks and years)
- Total number of administrations
- Total cumulative dose (mg)
- Duration on 225mg /300mg / 450mg/ 600mg / 900mg / 1200mg gantenerumab dosing steps (in days)
- Number of administrations on 225mg /300mg / 450mg/ 600mg / 900mg / 1200mg gantenerumab dosing steps
- Total cumulative dose on 225mg /300mg / 450mg/ 600mg / 900mg / 1200mg gantenerumab dosing steps (mg).

5.7.2 Adverse Events

AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version that is current at the time of the analysis (MedDRA Version 24.0). For each group, the frequency of each AE preferred term (PT) will be defined as the number of patients experiencing at least one occurrence of the event. Each summary table will present the overall number of patients experiencing at least one AE and the total number of reported AEs. Furthermore, the AEs will be sorted by system organ class (SOC) in decreasing order of overall incidence, then by PT in decreasing order of overall incidence. Patient listings will contain PTs and comments for each AE.

The following safety information during the treatment period in the OLE (from the start of OLE until 4 weeks after the last dose in the OLE) will be summarized and listed by previous randomization and overall:

- AEs
- SAEs
- Study drug-related AEs
- Study drug-related SAEs
- AEs leading to discontinuation of study treatment
- AEs leading to dose reduction
- AEs leading to dose interruption
- AEs of special interest (AESIs)

The following safety information will be summarized and listed by previous randomization and overall separately:

- AEs starting at or after the first dose of treatment in the DB part but before 4 weeks after the last dose in the OLE
- AEs starting at or after the first dose of treatment in the DB part and unresolved before the first dose of study treatment in the OLE
- SAEs starting at or after the first dose of treatment in the DB part but before 4 weeks after the last dose in the OLE
- SAEs starting at or after the first dose of treatment in the DB part and unresolved before the first dose of study treatment in the OLE

AEs during the follow-up period (from 4 weeks after the last dose in the OLE until the CCOD) will be summarized and listed by previous randomization and overall.

For scientific interpretation, it is important to separate ARIA-H and ARIA-E events and findings in listings and tables.

Using MedDRA standards,

- ARIA-E events is mapped to the following PTs:

  “AMYLOID RELATED IMAGING ABNORMALITY-OEDEMA/EFFUSION” or “VASOGENIC CEREBRAL OEDEMA”

- ARIA-H events is mapped to the following PTs:

  “CEREBELLAR MICROHAEMORRHAGE" or "CEREBRAL HAEMOSIDERIN DEPOSITION" or "CEREBRAL MICROHAEMORRHAGE" or "AMYLOID RELATED IMAGING ABNORMALITY-MICROHAEMORRHAGES AND HAEMOSIDERIN DEPOSITS.

AESIs are discussed in more detail in the study protocol Section 5.2.3. AESIs in this study will be captured by recording all case details on the AE eCRF.

In the tables of study drug related AEs or SAEs, if a patient has more than one occurrence of an event, the most closely related event will be counted. If the relationship of an AE to the study drug is missing, then the AE will be included only in the total number of events column, but not in the count of patients with the related event.

The AE and SAEs incidence rates per patient-years in the OLE will be calculated as the total number of occurrences divided by the total duration of treatment in the OLE in patient-years for each specified population or group. The duration of treatment in the OLE is from the start of OLE until 4 weeks after last dose in the OLE.

5.7.2.1 Administration Related Reactions

Signs and symptoms associated with administration related reactions will be listed and summarized by previous randomization and overall. The number of patients that

experienced administration related reactions will be summarized by administrated dose level at each visit.

5.7.2.2 CNS Symptoms Temporally Associated with ARIA-E MRI Findings

In this report, symptomatic ARIA-E is defined as onset or worsening of CNS symptom(s) that is/are temporally associated with ARIA-E MRI findings. A programmatic approach to capture symptomatic ARIA-E will be adopted in the analysis, since CNS symptoms have not been consistently captured in the CRF AE form. For programmatic retrieval, CNS symptoms temporally associated with ARIA-E MRI findings were defined as all AEs from the Nervous System Disorders and Psychiatric Disorders SOCs that had onset between 4 weeks prior to ARIA-E onset through ARIA-E resolution, and not falling under ARIA-associated PTs.

The list of ARIA-associated PTs include:

- Amyloid related imaging abnormalities
- Amyloid related imaging abnormality-microhaemorrhages and haemosiderin deposits
- Amyloid related imaging abnormality-oedema/effusion
- Cerebral amyloid angiopathy
- Brain oedema
- Vasogenic cerebral oedema
- Brain stem microhaemorrhage
- Cerebellar microhaemorrhage
- Cerebral microhaemorrhage
- Cerebral haemosiderin deposition
- Magnetic resonance imaging head abnormal
- Superficial siderosis of central nervous system

The CNS symptoms temporally associated with ARIA-E MRI findings (programmatic retrieval) will be listed and summarized by previous randomization and APOE ε4 genotype, and overall. A listing of AE reported as Symptoms of ARIA-E (eCRF) will be provided as well.

5.7.3 Magnetic Resonance Imaging Safety Findings

The MRI safety findings will be analyzed only on the safety population with post-OLE baseline MRI. The incidence and severity of ARIA-E and ARIA-H will be summarized by APOE ε4 genotype and up-titration and overall at each dose level. Recurrence of ARIA-E will be summarized by previous randomization and overall. Other ARIA-E characteristics will be summarized by previous randomization and APOE ε4 genotype and overall, and by up-titration scheme and overall. Same for other ARIA-H

characteristics as well. ARIA-E with CNS symptoms will be summarized by previous randomization and APOE ε4 genotype and overall.

5.7.4 Laboratory Data

Laboratory data will be summarized by previous randomization and overall for each assessment visit using descriptive statistics of absolute values, change from baseline values, and percentage change from baseline. The proportion of patients who develop marked abnormalities during study treatment, on the basis of the National Cancer Institute Common Terminology Criteria for Adverse Events Version 4.0 (NCI CTCAE v4.0) and the International Standard for Handling and Reporting of Laboratory Data (COG 2.2), will be summarized.

5.7.5 Vital Signs

Vital signs assessments include systolic blood pressure, diastolic blood pressure, and pulse measured throughout the study. Summaries of vital sign measurements by previous randomization and overall at each assessment visit using descriptive statistics will be generated for the actual values, the change from baseline, and the percentage change from baseline.

5.7.6 Electrocardiograms

ECGs were collected throughout the study. Descriptive statistics will be generated by previous randomization and overall at each assessment visit for the actual values, the change from baseline and the percentage change from baseline for the following parameters:

- Heart rate
- PQ (PR) interval
- QRS interval
- QT interval
- RR interval
- QT corrected using Bazett's formula (QTcB) and QT corrected using Fridericia's formula (QTcF).

Duration of the QT interval is inversely correlated with the corresponding heart rate, so it should be corrected relative to heart rate to allow a reliable interpretation of changes in the QTc interval, independent of the changes in the heart rate. The QT intervals will be corrected for the heart rate using Bazett's and Fridericia's as follows:

- The QTcB value will be calculated as the QT interval (in milliseconds) divided by the squared root of the RR interval (in seconds).
- The QTcF value will be calculated as the QT interval (in milliseconds) divided by the cube root of the RR interval (in seconds).

QT/QTc outliers will be categorized as recommended in the E14 guidance. In some cases, multiple ECGs were performed in error. In these cases, the mean values across the ECGs should be used.

Abnormal ECG measurements will be summarized by previous randomization and overall at each assessment visit.

5.7.7 Columbia−Suicide Severity Rating Scale

The C-SSRS (http://www.cssrs.columbia.edu) is an assessment tool used to assess the lifetime suicidality of a patient (C-SSRS at baseline) as well as any new instances of suicidality (C-SSRS since last visit). The structured interview prompts recollection of suicidal ideation, including the intensity of the ideation, behavior, and attempts with actual/potential lethality. During the OLE, the C-SSRS will be collected every 6 months. Additional C-SSRS may be collected as deemed necessary by the Principal Investigator.

Suicidal ideation, suicidal behavior, and self-injurious behavior without suicidal intent will be summarized by previous randomization and overall at each assessment visit.

5.7.8 Deaths

Deaths will be listed.

5.7.9 Analyses of Subgroups of Interest

Selected safety analyses will be performed for patients with post-OLE baseline MRI.

5.8 INTERIM ANALYSES

No efficacy interim analyses were planned in the OLE.

5.9 ANALYSES RELATED TO CORONA VIRUS DISEASE-19

Listings of AEs associated with corona virus disease 2019 (COVID-19) and major protocol deviations related to COVID-19 will be prepared.

6. SUPPORTING DOCUMENTATION

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

7. REFERENCES

Report 1105289-Population Pharmacokinetic analysis of gantenerumab with varying degrees of high mannose content. October 2020.

Collins DL, and Pruessner JC. Towards accurate, automatic segmentation of the hippocampus and amygdala from MRI by augmenting ANIMAL with a template library and label fusion. NeuroImage 2010; 52(4):1355-1366.

Coupé P, Manjón JV and Fonov V, et al. Patchbased segmentation using expert priors: application to hippocampus and ventricle segmentation. Neuroimage. 2011 Jan 15;54(2):940-54.

Fonov VS, Narayanan S and Arnold DL, et. al. Lateral ventricle segmentation based on fusion of expert priors in AD, ISMRM 2010 Stockholm 1-7 May 2010.

Fonov VS, Evans AC and Botteron K, et. al. Unbiased average age-appropriate atlases for pediatric studies. NeuroImage 2011;54:313-327.

Nakamura K, Guizard N and Fonov VS, et.al. Pairwise Jacobian Integration Method to Measure Grey Matter Atrophy in Multiple Sclerosis. International Society for Magnetic Resonance in Medicine Workshop on Multiple Sclerosis as a Whole-Brain Disease, London, England, United Kingdom. 2013.

Klein G, Delmar P, Voyle N, et al. Gantenerumab reduces amyloid-β plaques in patients with prodromal to moderate Alzheimer's disease: A PET substudy interim analysis. Alzheimers Res Ther 2019; 11:101.

Klunk WE, Koeppe RA and Price JC, et. al. The Centiloid Project: standardizing quantitative amyloid plaque estimation by PET. Alzheimers Dement. 2015 Jan;11(1):1-15.

Navitsky M, Joshi AD and Kennedy I, et. al. Standardization of amyloid quantitation with florbetapir standardized uptake value ratios to the Centiloid scale. Alzheimers Dement. 2018 Dec;14(12):1565-1571.