Official Title: Multicenter, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group Two Year Study to Evaluate the Effect of Subcutaneous RO4909832 on Cognition and Function in Prodromal Alzheimer's Disease With Option for up to an Additional Two Years of Treatment and an Open-Label Extension With Active Study Treatment

NCT Number: NCT01224106

Document Date: Statistical Analysis Plan Version 1: 18-November-2014

Statistical Analysis Plan Version H: 13-November-2020

STATISTICAL ANALYSIS PLAN

TITLE: **MULTICENTER, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-GROUP TWO YEAR STUDY TO EVALUATE THE EFFECT OF SUBCUTANEOUS RO4909832 ON COGNITION AND FUNCTION IN PRODROMAL ALZHEIMER'S DISEASE WITH OPTION FOR AN ADDITIONAL TWO YEARS OF TREATMENT**

PROTOCOL NUMBER: WN25203

STUDY DRUG: Gantenerumab (RO4909832)

VERSION NUMBER: 1

IND NUMBER: 102266

EUDRACT NUMBER: 2010-019895-66

SPONSOR: F. Hoffmann-La Roche Ltd

PLAN PREPARED BY: [redacted], Ph.D.

DATE FINAL: See electronic date stamp below.

STATISTICAL ANALYSIS PLAN APPROVAL

Name	Reason for Signing	Date and Time (UTC)
[redacted]	Company Signatory	18-Nov-2014 15:07:51

TABLE OF CONTENTS

1. BACKGROUND 5

2. STUDY DESIGN 5

2.1 Protocol Synopsis 5

2.2 Outcome Measures 5

2.2.1 Primary Efficacy Outcome Measure 5

2.2.2 Secondary Efficacy Outcome Measures 6

2.2.2.1 Alzheimer's Disease Assessment Scale–Cognition 6

2.2.2.2 Mini Mental State Exam Total Score 6

2.2.2.3 Cambridge Neuropsychological Test Automated Battery, Selected Subset Scores 6

2.2.2.4 Clinical Dementia Rating–Global Score 7

2.2.2.5 Functional Activities Questionnaire 8

2.2.2.6 Neuropsychiatric Inventory–Questionnaire 8

2.2.2.7 Dementia Assessment 8

2.2.2.8 Free and Cued Selective Reminding Test–Immediate Recall 8

2.2.2.9 Magnetic Resonance Imaging Volumetric Measures 8

2.2.3 Exploratory Efficacy Outcome Measures 8

2.2.4 Pharmacokinetic Outcome Measures 8

2.2.5 Safety Outcome Measures 9

2.3 Determination of Sample Size 9

2.3.1 Blinded Sample Size Assessment 10

2.4 Analysis Timing 10

3. STUDY CONDUCT 10

3.1 Randomization 10

3.2 Data Monitoring 11

4. STATISTICAL METHODS 11

4.1 Analysis Populations 11

4.1.1 Intent-to-Treat Population 11

4.1.2 ECG-Evaluable Population 11

4.1.3 Safety Population 12
4.2 Analysis of Study Conduct 12
4.2.1 Disposition of Subjects 12
4.3 Analysis of Treatment Group Comparability 12
4.3.1 Demographic Data and Baseline Characteristics 12
4.3.2 Medical Events 13
4.4 Efficacy Analysis 13
4.4.1 Primary Efficacy Endpoint 13
4.4.2 Statistical Models and Methods of Primary Outcome Analysis 13
4.4.3 Secondary Efficacy Endpoints 15
4.4.3.1 Key Secondary Efficacy Endpoint 15
4.4.3.2 Cognition 15
4.4.3.3 Global 16
4.4.3.4 Functioning 16
4.4.3.5 Neuropsychiatric Functioning 16
4.4.3.6 MRI Volumetric Measures 16
4.4.4 Statistical Models and Methods of Secondary Efficacy Analyses 16
4.4.5 Sensitivity Analyses 17
4.4.6 Subgroup Analyses 17
4.5 Pharmacokinetic and Pharmacodynamic Analyses 18
4.6 Safety Analyses 18
4.6.1 Exposure of Study Medication 19
4.6.2 Adverse Events 19
4.6.3 MRI Safety Findings and Injection-Site Reactions 20
4.6.4 Laboratory Data 21
4.6.5 Vital Signs 21
4.6.6 Electrocardiograms 21
4.6.7 Deaths 22
4.7 Missing Data 23
4.8 Interim Analyses 23
5. REFERENCES 25

LIST OF APPENDICES

Appendix 1 Protocol Synopsis 26
Appendix 2 Schedules of Assessments 30
Appendix 3 Alzheimer's Disease Assessment Scale–Cognition 35

1. BACKGROUND

This document describes the statistical methodology to be applied to the data that will be reported in the Clinical Study Report (CSR) for Study WN25203. Study WN25203 was originally designed as a proof-of-efficacy trial with safety data review to be performed by an internal safety monitoring committee. On the basis of feedback from European Medicines Agency/Committee for Medicinal Products for Human Use and U.S. Food and Drug Administration, the study has been modified to an adequate and well-controlled study to support the marketing registration of gantenerumab. The efficacy and safety endpoints that will be the basis for treatment comparison of 225 mg and 105 mg gantenerumab with placebo will be fully defined in this document along with the subject populations to be included in the analyses.

Although it is anticipated that there will be reporting of study data in a variety of contexts beyond that of the CSR, such as to 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.

2. STUDY DESIGN

2.1 PROTOCOL SYNOPSIS

The Protocol Synopsis is in Appendix 1. For additional details, see the Schedules of Assessments in Appendix 2.

2.2 OUTCOME MEASURES

2.2.1 Primary Efficacy Outcome Measure

The primary outcome measure is the change from baseline in the Clinical Dementia Rating–Sum of Boxes (CDR-SOB) score at Week 104.

The CDR is obtained through semi-structured interviews of subjects and informants. The cognitive functioning is rated in six domains: memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care. Each domain is rated on a 5-point scale of functioning as follows: 0, no impairment; 0.5, questionable impairment; 1, mild impairment; 2, moderate impairment; and 3, severe impairment (personal care is scored on a 4-point scale, without a 0.5 rating available).

The CDR-SOB score is obtained by summing the scores for each of the domain box; thus, the CDR-SOB score for a subject has a range of 0–18.

2.2.2 Secondary Efficacy Outcome Measures

The change from baseline at Week 104 will be assessed for the following secondary outcome measures.

2.2.2.1 Alzheimer's Disease Assessment Scale–Cognition

The Alzheimer's Disease Assessment Scale-cognition (ADAS-Cog) with 11-item total score and its modified version with 13-item total score are used in the study. The item scores are derived per description in Appendix 3. The 13-item total score is calculated using the sum of all the 13-item scores, whereas the 11-item total score is calculated using the sum of all-item scores except Item 4 (delayed word recall task) and Item 13 (number cancellation). 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, whereas the 11-item ADAS-Cog total score has a range of 0–70. Higher scores indicate greater pathology.

2.2.2.2 Mini Mental State Exam Total Score

The Mini Mental State Exam (MMSE) comprises 11 items to assess a subject'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.

2.2.2.3 Cambridge Neuropsychological Test Automated Battery, Selected Subset Scores

The Cambridge Neuropsychological Test Automated Battery (CANTAB) is a computerized assessment battery that consists of tests for neuropsychological function. CANTAB is evaluated by Cambridge Cognition Ltd independent of this study. The following tests were selected for this study and will be summarized descriptively:

- Simple reaction time (SRT): the duration between the onset of the stimulus and the time when a subject released the button in correct, assessed trials in which the stimulus could appear in one location only
- Simple movement time (SMT): the length of time to touch the stimulus after the button has been released, calculated in correct, assessed trials in which the stimulus could appear in one location only
- Five-choice reaction time (RTI): the duration between the onset of the stimulus and the release of the button, calculated in correct, assessed trials in which the stimulus could appear in any one of five locations
- Delayed match to sample (DMS) latency (the subject's speed of response): the numbers of correct patterns selected and statistical analysis measuring the probability of an error after a correct or incorrect response
- Spatial working memory (SWM): the ability to remember the location in which something is perceived and in addition, the ability to recall a series of visited locations

- Pattern recognition memory (PRM) immediate: a test of visual pattern recognition memory in a two-choice forced discrimination paradigm
- PRM delayed: a test of visual pattern recognition memory in a two-choice forced discrimination paradigm
- Rapid visual information processing (RVP): a signal-detection measure of sensitivity to the target, regardless of response tendency (expected range is 0.000–1.000 [bad to good])
- Paired associates learning (PAL): the errors made by the subject, the number of trials required to locate the pattern(s) correctly, memory scores, and stages completed

2.2.2.4 Clinical Dementia Rating–Global Score

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 ("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 scores are recorded on the electronic Case Report Form (eCRF) by investigators and are the basis of analyses of CDR global scores in this study.

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 are given 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.

2.2.2.5 Functional Activities Questionnaire

The Functional Activities Questionnaire (FAQ) is an informant-based assessment that presents a forced choice of four levels of functioning for ten activities of daily living. Total scores for FAQ will be summarized.

2.2.2.6 Neuropsychiatric Inventory–Questionnaire

The Neuropsychiatric Inventory–Questionnaire (NPI-Q) is an informant-based instrument in which 10 behavioral domains are evaluated on a 3-point scale (mild, moderate, or severe). The caregiver's distress portion of the scale will not be used in this study.

2.2.2.7 Dementia Assessment

Time to dementia is defined as the time interval between the first treatment date and the date that subject is assessed as having Alzheimer-type dementia by investigators.

2.2.2.8 Free and Cued Selective Reminding Test–Immediate Recall

The version of the Free and Cued Selective Reminding Test–Immediate Recall (FCSRT-IR) uses written words and immediate recall after each card with four words completed. The FCSRT-IR is a measure of memory under conditions that control attention and cognitive processing in order to obtain an assessment of memory not confounded by normal age-related changes in cognition. The performance of the FCSRT-IR has been associated with nonclinical and early dementia in several longitudinal epidemiologic studies.

2.2.2.9 Magnetic Resonance Imaging Volumetric Measures

The change in hippocampal volume from baseline will be calculated. The change in whole-brain volume from baseline and the change in ventricular volume from baseline will be assessed. Change from baseline in other volumetric measures such as cortical thickness may be assessed. The assessment details are available in the magnetic resonance imaging (MRI) Charter.

2.2.3 Exploratory Efficacy Outcome Measures

The following parameters will be recorded and analyzed:

- Change from baseline in plasma amyloid-beta $(A\beta)_{1-40}$ and $A\beta_{1-42}$ levels
- Change from baseline in cerebrospinal fluid (CSF) T-tau, P-tau, and $A\beta_{1-42}$ levels

2.2.4 Pharmacokinetic Outcome Measures

The following pharmacokinetic (PK) outcome measures will be assessed:

- Peak plasma concentration at steady state (C_{max})
- Time to peak concentration
- Trough plasma concentration at steady state
- Area under the time-plasma concentration curve at steady state ($AUC_{0-\tau}$)

Other parameters derived from the PK model will be regarded as secondary parameters. Potential covariates for the model will be explored as appropriate.

2.2.5 Safety Outcome Measures

MRI safety findings such as amyloid-related imaging abnormalities (ARIA) of the microhemorrhage (ARIA-H) and effusive or edematous (ARIA-E) and any other findings are:

- Adverse events (AEs)
- Physical and neurologic examination findings
- Geriatric Depression Scale
- Columbia-Suicide Severity Rating Scale

2.3 DETERMINATION OF SAMPLE SIZE

A total of 770 subjects (257 for each gantenerumab dose group and 256 for the placebo group) were planned to be randomized in a 2:1 ratio to receive either gantenerumab (high or low dose, depending on Apolipoprotein E [ApoE] genotype groups) or placebo. The power calculation was based on simulations of the mixed-effects model repeated measures (MMRM) analysis planned for analysis of the primary outcome variable. The following assumptions were included:

- 80% power in demonstrating an effect size of 0.35 (225 mg of gantenerumab vs. placebo)
- Overall Type I error at the 0.05 level
- Four post-baseline assessment visits at Weeks 24, 52, 76, and 104
- Overall dropout rate of 30% at Week 104 for both the placebo and active treatment groups, incremental rates over the 104-week period
- Effect sizes of 0.35 (225 mg of gantenerumab vs. placebo) and 0.25 (105 mg of gantenerumab vs. placebo) at Week 104, with the assumption of increasing magnitude of the treatment difference over the 104-week period
- A correlation structure with stronger correlations for assessments that are adjacent in time and weaker between assessments that are further apart
- Approximately 5% reductions in sample size with hippocampal volume at baseline as a covariate in the analysis

In order to protect the overall Type I error at the 0.05 significance level, a hierarchical testing procedure will be used in which the first hypothesis to be tested will be the comparison of the 225-mg gantenerumab group with the corresponding placebo group.

The second test, which compares the 105-mg gantenerumab group to the placebo group, will be carried out only if the first test is statistically significant at the two-sided 0.05 level.

2.3.1 Blinded Sample Size Assessment

A blinded assessment of the variance of the absolute change from baseline in CDR-SOB at Months 6, 12, 18, and 24 will be conducted in order to assess whether the assumed SDs of this variable at these timepoints are appropriate. The sample size and power considerations for the study were developed on the basis of simulations of an MMRM analysis with use of absolute change from baseline in CDR-SOB values at these scheduled timepoints. The assumed SD values of these CDR-SOB values used in these simulations (estimated on the basis of the Alzheimer Disease Neuroimaging Initiative-1 database) are 0.96 at Month 6, 1.20 at Month 12, 1.44 at Month 18, and 2.00 at Month 24.

A blinded sample size assessment was conducted approximately 6 months prior to the completion of recruitment of the planned sample size of 770 subjects. At that time, the observed SD of the absolute change from baseline in CDR-SOB at each timepoint was calculated. This calculation was done on a completely blinded basis, with no results generated by treatment group. The observed SDs at these timepoints did not differ substantially from the assumptions described above. As a consequence, no sample size adjustment was made to ensure at least 80% power of the study.

2.4 ANALYSIS TIMING

The final analysis will be performed after all randomized subjects have been treated for at least 2 years.

Safety data and selected efficacy data are being reviewed approximately once every 3–4 months by the iDMC.

An administrative interim analysis was conducted in June 2013 to assess whether preparations for an additional adequate and well-controlled study should begin. For this interim analysis, the review included unblinded results for positron emission tomography (PET) scan, safety, anti-drug antibody (ADA), biomarkers, and pharmacokinetics. Details of this administrative interim analysis are documented in a separate interim Statistical Analysis Plan.

A 2-level futility interim analysis will be performed when approximately 50% of the sample size is scheduled to complete the Week 104 visit. Further details of the futility analysis are provided in Section 4.8.

3. STUDY CONDUCT

3.1 RANDOMIZATION

Randomization was performed centrally with use of an interactive system (interactive voice/Web response system). The randomization ratio was 2 active (gantenerumab 105 mg, gantenerumab 225 mg) to 1 placebo. Within the 2ε4 genotype, all subjects randomized to gantenerumab are receiving 105 mg. A dynamic randomization scheme

was applied. The algorithm for dynamic patient allocation to treatment was based on minimization with biased coin assignment (Pocock and Simon 1975). The randomization was stratified by PET substudy participation, ApoE4 genotype (0ε4 vs. 1ε4 vs. 2ε4), and region (Europe vs. rest of world).

3.2 DATA MONITORING

An external iDMC was created to review the safety and selected efficacy data once every 3 months, starting in mid-July 2012. The iDMC Charter describes the roles of the iDMC and the frequency and conduct of the planned review meetings in detail.

4. STATISTICAL METHODS

Details about the statistical models for the various types of endpoints (continuous, categorical, and time-to-event endpoints) are provided in Section 4.4.2.

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

4.1 ANALYSIS POPULATIONS

4.1.1 Intent-to-Treat Population

The intent-to-treat (ITT) population will include all subjects who have received any dose of study treatment and had at least one post-baseline assessment of CDR-SOB. The ITT analysis will be performed by randomized treatment.

The ITT population will be the primary population for all analyses of primary and secondary outcome variables as well as for the imaging and biomarker and clinical efficacy variables.

4.1.2 ECG-Evaluable Population

The ECG-evaluable analysis population will consist of subjects who meet all of the following criteria:

- Received at least one dose or a partial dose of gantenerumab or placebo
- Have at least one interpretable predose ECG measurement prior to dosing
- Have at least one interpretable postdose ECG measurement on any of the post-baseline timepoints

All ECG analyses will be based on this population.

4.1.3 Safety Population

The safety population will consist of all subjects who have received at least one dose of study drug, regardless of whether a subject withdrew prematurely or not. All safety data will be analyzed according to study drug actually received.

4.2 ANALYSIS OF STUDY CONDUCT

4.2.1 Disposition of Subjects

The number of subjects randomized, the duration of the study, and the number and percentage of subjects who prematurely withdraw from the study, including the reasons for discontinuation, will be summarized.

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 ITT population and discussed in the CSR. The comparisons will be made on patient demographic and baseline variables, including the following:

- Sex
- Age
- Race
- Years of education
- Female reproductive status
- Weight
- Height
- Body mass index
- ApoE status
- FcγR genotype
- FCSRT-IR, total recall at baseline
- MMSE, total score at baseline
- ADAS-Cog total score (sum of 11-item scores) at baseline
- ADAS-Cog total score (sum of 13-item scores) at baseline
- CDR-SOB at baseline
- CDR-Memory at baseline

All comparisons will be based on descriptive statistics.

4.3.2 Medical Events

All listings and summaries will include only the preferred terms (PTs) or superclass terms (SCTs).

Medical events (such as AEs, treatments, procedures, diseases) will be assigned as previous or concomitant on or after a reference date (such as index event date, randomization date, start date of study drug) on the basis of start and end dates. The reference date must always be a complete and valid date.

If the medical event does not have complete or valid start and end dates, the following rules will be applied:

- If the medical event start date is 1) completely missing, 2) has only the year specified and this is the same as the year of the reference date, or 3) has only the month/year specified and these are the same 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 1) completely missing, 2) has only the year specified and this is the same as the year of the reference date, or 3) has only the month/year specified and these are the same 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).

With use of these data handling rules, when needed, each event will be classified as one of the following:

- Previous event: A medical event with an end date (applying the rules above if necessary) strictly before the reference date will be considered as previous to the reference date.
- Concomitant event: A medical event with a start date (applying the rules above if necessary) before the reference date and end date (imputed as above if necessary) on or after the reference date will be considered as present concomitant at the reference date.

4.4 EFFICACY ANALYSIS

4.4.1 Primary Efficacy Endpoint

The primary efficacy endpoint is the change from baseline in the CDR-SOB score at Week 104.

4.4.2 Statistical Models and Methods of Primary Outcome Analysis

The primary inferential statistics will be applied to the pairwise comparisons of each of the two gantenerumab dose groups with the placebo group. To protect the overall Type I error at the 0.05 significance level, a hierarchical testing procedure will be used in

which the first hypothesis to be tested will be the comparison of the 225-mg gantenerumab group with the matching subgroup of placebo patients with ApoE 0ε4 and 1ε4 genotypes. The second test, which compares the 105-mg gantenerumab group with the entire placebo group, will be performed only if the first test is statistically significant at the two-sided 0.05 level.

For the assessment of differences between each gantenerumab group and placebo with regard to the mean change from baseline in the CDR-SOB score at Week 104, an MMRM analysis, which incorporates data up to 104 weeks of treatment, will be used to utilize all the data recorded over time with consideration of the variance-covariance matrix of the repeated measures. This method allows a general unstructured variance-covariance matrix and enables the inclusion of data from subjects with incomplete data across scheduled timepoints.

Independent MMRM models will be fitted for the comparisons of the two gantenerumab dose group with the placebo group. For each comparison, the following null (H_0) and alternative (H_a) hypotheses will be tested to compare the mean change values between the gantenerumab group and placebo group:

H_0: $MEAN_{RO} = MEAN_{placebo}$ versus

H_a: $MEAN_{RO} \neq MEAN_{placebo}$

for which the $MEAN_{RO}$ and $MEAN_{placebo}$ refer to the mean change of gantenerumab and placebo from baseline to Week 104, respectively.

The model will include the change from baseline in the CDR-SOB score as the dependent variable. The effects in the model will include independent variables of the fixed, categorical effects of treatment, assessment weeks relative to the first dose of study drug (i.e., time), treatment-by-time interaction, and ApoE status (carrier vs. non-carrier), along with baseline CDR-SOB and hippocampal volume at baseline. Hereby, the ApoE status "carrier" comprises all patients with ApoE status 1ε4 or 2ε4, and patients with ApoE status 0ε4 are non-carriers. Time will be treated as the repeated variable within a subject. Subject, treatment, and time will be treated as class variables. An unstructured variance-covariance structure will be applied to model the within-subject errors. The model will be fitted using the restricted maximum likelihood method. Denominator degrees of freedom will be estimated using Satterthwaite's approximation. A treatment-by-time interaction contrast will be constructed to estimate the difference between the respective gantenerumab dose group and placebo in the mean change from baseline to Week 104. On the basis of the above analysis, p-values, least square means, SEs, and the 95% CIs of treatment difference will be reported for each dose of gantenerumab relative to placebo.

4.4.3 Secondary Efficacy Endpoints

In light of the yet-to-be defined considerations arising from continuing scientific and regulatory evolution that could impact Alzheimer's disease drug development activities in the near future, the planned secondary endpoints specified in this Statistical Analysis Plan might need to be adjusted to accommodate these considerations as they emerge. Additionally, presentation and analysis of the biomarker data from this trial are planned and are expected to be critical to the interpretation of the overall dataset. The plans for evaluation of these biomarkers will be addressed in the future as scientific understanding on the appropriate biomarkers to be used continues to evolve.

4.4.3.1 Key Secondary Efficacy Endpoint

Ongoing discussions within the Sponsor's organization and between the Sponsor and Regulatory Authorities are being conducted to establish the optimal endpoint and analytic process to implement for the key secondary efficacy endpoint for Study WN25203 (hereby "key secondary efficacy endpoint" is used to refer to the first secondary endpoint in the hierarchical testing and the results of which the Sponsor would propose to present in the prescribing information).

The Sponsor is exploring Dementia Assessment (time to dementia), as performed by investigators every 6 months, as a potential key secondary efficacy endpoint. Emerging composites to examine clinical decline are also being evaluated as a potential key secondary efficacy endpoint. The Sponsor is currently conducting the analyses on the performance of the Dementia Assessment (and potential emerging composites) outside of this study to determine its clinical characteristics (such as median time to event, number of events, etc.) compared with other measures of clinical assessment before final proposal to Regulatory authorities for inclusion as the key secondary efficacy endpoint in Study WN25203.

4.4.3.2 Cognition

The change from baseline at Week 104 in the following will be assessed:

- ADAS-Cog total score (13 items and 11 items)
- MMSE total score
- CANTAB with selected subtest scores:
    - SRT
    - SMT
    - Five-choice RTI
    - DMS latency
    - SWM
    - PRM immediate
    - PRM delayed
    - RVP

PAL

- FCSRT-IR

4.4.3.3 Global

- CDR global score

4.4.3.4 Functioning

- The change from baseline in FAQ total score

4.4.3.5 Neuropsychiatric Functioning

- The change from baseline in the total and domain scores of the NPI-Q

4.4.3.6 MRI Volumetric Measures

- The change from baseline in hippocampal volume
- The change from baseline in whole brain volume
- The change from baseline in ventricular volume
- The change from baseline in other volumetric measures such as cortical thickness, if adequate

4.4.4 Statistical Models and Methods of Secondary Efficacy Analyses

Continuous Variables

The secondary efficacy variables of the continuous type with multiple post-baseline scheduled assessments will be analyzed using the same methods described above for the primary outcome variable: MMRM including the covariates of baseline CDR-SOB and hippocampal volume at baseline.

Time-to-Event Variables

Time to event in the gantenerumab group and the placebo group will be compared using a two-sided log-rank test stratified by ApoE status (carrier vs. non-carrier). The proportion of patients with the event given the treatment group and time will be estimated using Kaplan-Meier methodology.

In order to estimate the overall hazard ratio, a proportional hazards model will be assumed with respect to time-to-event variables. Under the following model, the hazard function for the jth subject is expressed as:

$$\lambda_j(t) = \lambda_0(t)\exp(z_j\beta)$$

where $\lambda_0(t)$ is the baseline hazard function and z_j is the vector of explanatory variables for the jth subject who has received treatment. For time-to-event variables, the explanatory variables included in the regression model will be treatment, ApoE genotype status (carrier vs. non-carrier), baseline CDR-SOB, and hippocampal volume at baseline. The treatment effect is then the hazard ratio for gantenerumab versus placebo, which is assumed to be constant over time.

Subjects will be considered to be at risk for an event until the database cutoff for the primary analysis is reached. Subjects who are lost to follow-up or who withdraw consent (while event-free) will be censored at the time that they are last known to be event-free.

Point estimates and two-sided 95% CIs will be determined for treatment effects and will be based on the estimate (and SD) for the treatment effect term in a Cox proportional hazard model.

Hierarchical Testing

Only if both tests for the primary efficacy endpoint result in statistically significant outcomes will a statistical test for the key secondary efficacy endpoint be conducted in a confirmatory manner. Hereby, the overall Type I error will be controlled at the 0.05 level.

4.4.5 Sensitivity Analyses

If statistically significant treatment effects are demonstrated in the analysis of the primary outcome measure, the robustness of the results will be evaluated with respect to their potential dependence on the assumptions made and definitions of the composite endpoint components and subject populations.

- Analysis of the primary outcome measure, excluding hippocampal volume as a covariate

  The MMRM model excluding the hippocampal volume at baseline covariate will also be performed to compare each gantenerumab dose group with the corresponding placebo group to assess the effect of this factor on the primary outcome measure.

- Analysis of the primary outcome measure excluding the ApoE4 genotype carrier status as a covariate

  The MMRM model excluding the ApoE4 genotype will also be performed to compare each gantenerumab dose group with the corresponding placebo group to assess the effect of this factor on the primary outcome measure.

Pattern mixture models and multiple imputation methods, as described in Section 4 of the National Research Council report, will be performed as additional sensitivity analyses.

4.4.6 Subgroup Analyses

The primary variable (i.e., the change from baseline in the CDR-SOB score at Week 104) and ADAS-Cog will be summarized within subgroups defined by the following baseline characteristics. Inferential statistics may be applied in case the difference is clinical relevant.

- Sex
- Age
- Race

- ApoE status
- FcγR genotype
- Region (North American [America and Canada] vs. others)

Descriptive statistics for baseline-demographics and disease-characteristics variables, as well as for the primary outcome variable, will also be generated for the subgroups of subjects randomized prior to and after the CSF Aβ inclusion-criterion cutoff was modified and the expansion of the study sample size with an effective date of 14 March 2012, submission date of Protocol Study WN25203C for the IND, 29 March 2012 (S-0020). The subgroup analyses will be used to evaluate the consistency of the data between the two study populations over time.

4.5 PHARMACOKINETIC AND PHARMACODYNAMIC ANALYSES

Nonlinear mixed-effects modeling (using the software NONMEM®) will be used to analyze the dose concentration-time data of gantenerumab. Information from other studies in humans may be incorporated to establish the PK model. The selection of parameters and the derivation of individual measures of exposure, such as $AUC_{0-\tau}$, C_{max}, and steady-state concentration at the end of a dosing interval (C_{trough}) will depend on the final PK model used for this analysis. All PK parameters will be presented in listings and descriptive summary statistics, including the arithmetic mean (for $AUC_{0-\tau}$, C_{max}, and C_{trough}), median, range, SD, and coefficient of variation. Results of this analysis will be reported separately.

In this study, PET using the amyloid tracer AV-45 will be used as a pharmacodynamic (PD) measurement of potential effects of gantenerumab on amyloid load in the brains of subjects with early (prodromal) Alzheimer's disease. The primary PET variable, percent change from screening in cortical composite standard uptake value ratio (SUVR), will be summarized using descriptive statistics and analyzed using the MMRM method. The effects in the model will include independent variables of the fixed, categorical effects of treatment, assessment time, and treatment-by-time interaction. Time will be treated as the repeated variable within a subject. Subject, treatment, and time will be treated as class variables. An unstructured variance-covariance structure will be applied to model the within-subject errors. Similar analyses may be performed for absolute change and percent change from screening in regional SUVR values.

4.6 SAFETY ANALYSES

Descriptive statistics will be used to analyze all safety data. Wherever possible, Roche's safety reporting system will be used to present safety evaluations. The outputs defined in this section will be produced using the safety analysis population, unless otherwise specified.

Selected safety outcome measures will also be reported by ApoE genotype.

For all safety analyses, Day 1 will be defined as the start date of double-blind randomized treatment. For all safety data (laboratory findings, ECG results, vital sign measurements), the baseline value will be the last available value before the start of dosing.

4.6.1 Exposure of Study Medication

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

- Duration of treatment (defined as the time from the date of first treatment to date of last treatment)
- Cumulative dose
- Dose intensity (%) (defined as total amount of study treatment administrated relative to total amount of study treatment assigned per protocol)

4.6.2 Adverse Events

AEs will be coded using the MedDRA, Version 15.0. For each treatment group, the frequency of each AE preferred term will be defined as the number of subjects experiencing at least one occurrence of the event. The incidence rate will be calculated as the frequency count divided by the total number of subjects in the population and treatment group specified.

Each table will also present the overall number of subjects experiencing at least one AE and the total number of AEs reported. Subject listings will contain PTs and comments for each event. In summary tables, AEs will be sorted by body system and then by PT (both in decreasing order of overall incidence).

All AEs will be summarized separately for the following two study periods and include the specified AEs:

- Treatment period includes 1) the AEs for which the onset date is on or after the start of dosing and on or prior to 4 weeks after the last dose day or 2) the AEs for which the onset date is prior to the first dosing day and the end date is on or after the first dose day or is unresolved.
- Follow-up period includes 1) the AEs for which the onset date is 28 days or more after the last dose day or 2) the AEs for which onset date is during the treatment period and the end date is after the last dose day or is unresolved. When the last dose date is not available (e.g., for subjects who are still in the study), all the AEs reported will be considered as occurring during the “treatment period.”

The following safety information will be summarized:

- AEs
- Serious AEs
- Study drug–related AEs

- Study drug–related serious AEs
- AEs leading to discontinuation of study treatment
- AEs leading to dose reduction
- AEs leading to dose held

In addition, summaries of AEs will be provided by age, race, and sex.

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

- Multiple occurrences of the same AE in 1 subject will be counted only once.
- For administrative analysis and iDMC (due to potentially incomplete coding of the terms), missing PTs will be presented as the investigator's terms, preceded by ″NO_PT_Found″ in all outputs. Missing 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 treatment emergent, 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 considered treatment emergent.
- If the end date is missing and the onset date is on or after the first dosing date, then the event will be considered treatment emergent.

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

- An AEs will be included in the summary table of AEs leading to study drug discontinuation if the "study drug adjustment" tick box on the Adverse Event CRF is checked "discontinued."
- In the summary table of AEs by intensity, if a subject has more than one occurrence of an event, the event with the most severe intensity will be counted. If the intensity of an AE is missing, then the AE will be included only in the total number of events column and not in the count of subjects with the event by intensity.
- In the summary table of AEs by relationship to study drug, if a subject has more than one occurrence of an event, the most closely related event will be counted. If the relationship of an AE is missing, then the AE will be included only in the total number of events column and not in the count of subjects with the event by relationship.

4.6.3 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 ApoE4 genotype.

Both, ARIA and injection-site reactions will be also captured as AEs, but respective listings and reports will be summarized in a separate subchapter.

4.6.4 Laboratory Data

Laboratory data will be summarized for each assessment with use of descriptive statistics (e.g., mean, SD, and median). The proportion of subjects who develop marked abnormalities during study treatment on the basis of National Cancer Institute Common Terminology Criteria for Adverse Events, Version 4.0 will be summarized.

4.6.5 Vital Signs

Vital sign assessments include systolic blood pressure, diastolic blood pressure, and pulse measured throughout the study. Summaries of vital sign measurements at each assessment timepoint using descriptive statistics will be presented for each group with respect to both the actual values and changes from baseline.

Categorical analysis will include number of subjects exceeding a systolic blood pressure of 160−179 and $\geq$180 or a diastolic blood pressure of 95−109 and $\geq$110 mm Hg or with blood pressure of $<$90 mm Hg systolic and $<$60 mm Hg diastolic. The number of subjects with changes in systolic and/or diastolic blood pressure by 5 to $<$10, 10 to $<$15, and $\geq$15 mm Hg will also be provided.

With respect to heart rate (HR), the number of subjects with an HR of $>$100 bpm or $<$60 bpm and the number of subjects with a change in HR by 5 to $<$10, 10 to $<$15 and $\geq$15 bpm will be provided.

4.6.6 Electrocardiograms

An average of the three values of QT interval (QT), QTcF (Fridericia's correction), and HR from each ECG triplicate reading will be used as a single observation at the timepoint for all analyses except the T-wave and U-wave morphology.

Duration of the QT interval is inversely correlated with the corresponding HR so it has to be corrected relative to HR to allow a reliable interpretation of changes in the QTc interval independent of the changes in the HR. The QT intervals will be corrected for the HR with use of Bazett's (QTcB) and QTcF formulae 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 cubed root of the RR interval (in seconds).

Descriptive statistics will be generated for the actual values and the change from baseline for the two visits at which ECGs are scheduled for all subjects (baseline and end of treatment).

The following parameters will be summarized:

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

Additional categorical analyses will include the number and percentage of subjects at each timepoint whose ECG recordings meet any of the following criteria:

- Absolute QT/QTc values: >450 to <480 ms, ≥ 480 to <500 ms, and ≥ 500 ms
- Change from baseline in QTc interval: 20 to <30 ms, ≥ 30 to <60 ms, and ≥ 60 ms
- PR changes from baseline are $\geq 50\%$ if absolute baseline value was ≤ 200 ms and are $\geq 25\%$ if absolute baseline value was >200 ms.
- QRS changes from baseline are $\geq 50\%$ if absolute baseline value was ≤ 100 ms and $\geq 25\%$ if absolute baseline value was >100 ms.
- New incidence of abnormal U waves (a morphology abnormality in any lead will be scored as the result with use of three individual ECG tracings from each triplicate as three observations)
- New incidence of abnormal T waves (a morphology abnormality in any lead will be scored as the result with use of three individual ECG tracings from each triplicate as three observations)
- Number and percentage of subjects with abnormal ECG findings overall

4.6.7 Deaths

Deaths will be categorized by study phase (between start date of study drug and 30 days after end date of study drug or >30 days after end date of study drug) and summarized.

4.7 MISSING DATA

For the efficacy rating scales with multiple items, the overall total and domain scores (i.e., "summed scores") will be set to missing for the primary analysis when at least one item score that contributes to the "summed scores" is missing. In addition, an exploratory analysis will be performed with imputed missing "summed score" (in case there are $>10\%$ of subjects who have a missing "summed score" for a given test at a given assessment time) with use of various imputation rules as long as $<25\%$ of the items involved in the given summed score are missing. When $\geq 25\%$ of item scores are missing, the summed score will remain as missing. Imputation rules will include the following simple algorithm:

> Total score from the non-missing items in a domain (or total) multiplied by (maximum possible scores for items involved in the domain [or total] divided by maximum possible scores for items with non-missing score)

A more complex method may be explored, such as the two-way imputation (based on a linear model) method, proposed by Bernaards and Sijtsma (2000). The authors impute item j of subject i by $PMi + IMj - OM$, for which PMi is the subject's mean across all of his or her available item responses, calculated for each subject.

- IMj is the mean of item j across all available scores, calculated for each item
- OM is the overall mean across all subjects and items.

4.8 INTERIM ANALYSES

In addition to the regularly scheduled safety monitoring reviews by the iDMC as described in Section 10.1.2 of the Protocol, one administrative interim analysis was conducted in June 2013. For this interim analysis, unblinded summary tables showing results for selected safety parameters (such as MRI findings) and selected PET PD parameters collected in the PET substudy were generated. Unblinded summary tables showing results for PK and ADA data were also generated.

The results of the administrative interim analysis were generated by the same independent data coordinating center (iDCC) that is used for the quarterly safety reviews. The interim analysis results were reviewed by both the iDMC and the Sponsor's Data Review Board (DRB). The details of the composition of the DRB and the data outputs to be reviewed are described in the iDMC Charter and a separate interim Statistical Analysis Plan.

The iDMC recommendation on the basis of the results of the administrative interim analysis was used by the DRB to determine whether to initiate activities for additional clinical trials, and no efficacy data was provided to DRB for review. The Type I error rate has not been adjusted for this interim analysis because the study team has remained blinded to the data. The results of the interim analysis have not been used to inform the future conduct of this trial and therefore there is no risk to introduce operational bias.

A 2-level futility interim analysis will be performed when approximately 50% of the sample size is scheduled to complete the Week 104 visit. This analysis will be performed by the iDCC that generates the unblinded results to be reviewed by the iDMC on a regular basis. For the 2-level futility analysis, the iDMC will review the unblinded results for the primary and secondary efficacy endpoints (absolute change from baseline at Week 104) as well as emerging safety data. The specifics of the futility analysis will be documented prior to its conduct in an Interim Statistical Analysis Plan.

No adjustment for multiple comparisons will be made to the α level for this analysis because the decision rules for the futility analysis will not allow for the opportunity to stop the study early for overwhelming efficacy.

5. REFERENCES

Bernaards CA, Sijtsma K. Influence of imputation and EM methods on factor analysis when item nonresponse in questionnaire data is nonignorable. Multivariate Behav Res 2000;35:321–64.

Pocock SJ, Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. Biometrics 1975;31:103–15.

Appendix 1
Protocol Synopsis

SYNOPSIS OF PROTOCOL NUMBER WN25203F

TITLE	Multicenter, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group Two Year Study to Evaluate the Effect of Subcutaneous RO4909832 on Cognition and Function in Prodromal Alzheimer's Disease with Option for an Additional Two Years of Treatment		
SPONSOR	Roche	Clinical Phase	3
INDICATION	Prodromal Alzheimer's Disease		
OBJECTIVES			
Primary	To evaluate the effect of gantenerumab versus placebo on the change in the Clinical Dementia Rating scale Sum of Boxes (CDR-SOB), a global measure of cognition and functional ability.		
Secondary			
Cognition, global, and functioning	To evaluate the effect of gantenerumab versus placebo on cognition (assessed with the A*lzheimer Disease Assessment Scale-Cognition [*ADAS-Cog*]*, *Mini Mental State Exam [*MMSE*]*, *Cambridge Neuropsychological Test Automated Battery [*CANTAB*]*, and the *Free and Cued Selective Reminding Test-Immediate Recall [*FCSRT-IR*]*) on global measures (based on the CDR global score), and on functioning (assessed with the *Functional Activities Questionnaire [*FAQ*]*), on neuropsychiatric functioning (assessed with the *Neuropsychiatric Inventory Questionnaire [*NPI-Q*]*) and on time to onset of dementia.		
Safety	To assess the safety and tolerability of gantenerumab assessed by magnetic resonance imaging (MRI), physical and neurological examinations, vital signs, blood and urine safety tests, ECGs, Columbia-Suicide Severity Rating Scale (C-SSRS), Geriatric Depression Scale (GDS), and adverse event monitoring.		
MRI volumetry	To evaluate the effect of gantenerumab versus placebo on hippocampal volume, whole brain volume, ventricular enlargement, and possibly other volumetric measures of the brain.		
CSF Biomarkers	To evaluate the effect of *gantenerumab versus placebo on levels of cerebrospinal fluid (CSF) biomarkers of Alzheimer's disease (AD) pathology (i.e., total tau [T-tau], phospho-tau [p-tau], and amyloid-beta ($A\beta_{1-42}$]).*		
Pharmacokinetics/ *Anti-Drug Antibodies*	To determine the relationship of plasma and CSF concentrations of gantenerumab on other responses. To assess incidence of anti-gantenerumab antibodies, and if relevant, evaluate its effect on the pharmacokinetic (PK), pharmacodynamic (PD), efficacy. and safety parameters.		

Appendix 1
Protocol Synopsis (cont.)

Exploratory *Plasma Biomarkers*	To evaluate the effect of treatment on peripheral biomarkers of amyloid deposition and/or clearance, neurodegeneration, inflammation, and other markers known to be involved in the pathogenesis of AD.
Clinical Genotyping	*To evaluate the effect of gene encoding apoliprotein E (ApoE) ε4 genotype and fragment crystalizable gamma receptor (FcγR) genotype on PK/PD, efficacy, and safety parameters*
TRIAL DESIGN	Multicenter, randomized, double-blind, placebo-controlled, parallel-group
NUMBER OF SUBJECTS	Approximately a total of 770; 514 on gantenerumab and 256 on placebo with an option to increase to no more than 960 based on a blinded evaluation of the observed variability in the absolute change from baseline in CDR-SOB at Week 76 done at the time of the administrative interim analysis.
TARGET POPULATION	Subjects who meet memory and biomarker criteria which indicate a strong likelihood that they are in the prodromal phase of Alzheimer's Disease
LENGTH OF STUDY	Part 1: Approximately *3* years and *2* months which includes 8 weeks for screening, *2 years (*104 weeks*)* of treatment, and 1 *year (52* weeks*)* of follow-up. Part 2: additional 2 years which includes *1 year (52* weeks*)* of follow-up at the end of Year 4 rather than Year 2.
END OF STUDY	The end of the study will be considered to be the date of the last visit (including the last scheduled follow-up visit according to study protocol) of the last subject in the study.
INVESTIGATIONAL MEDICAL PRODUCT(S) DOSE/ ROUTE/ REGIMEN	Gantenerumab (105 mg or 225 mg for the 0ε4 and 1ε4 ApoE genotypes or 105 mg for the 2ε4 ApoE genotype patients) by subcutaneous (SC) injection, every 4 weeks
COMPARATOR DOSE/ ROUTE/ REGIMEN	Placebo, by SC injection, every 4 weeks
ASSESSMENTS OF:	
- EFFICACY	CDR SOB, ADAS-Cog, MMSE, CANTAB, FAQ, *FCSRT-IR,* NPI-Q, onset of dementia
- SAFETY	MRI, physical and neurological examinations, vital signs, blood and urine safety tests, C-SSRS, *GDS,* ECGs and adverse event monitoring
- PHARMACOKINETICS	Gantenerumab levels in plasma and CSF
- BIOMARKERS	Peripheral and CSF biomarkers of amyloid deposition and/or clearance (including but not restricted to CSF and plasma $A\beta_{1-40}$ and $A\beta_{1-42}$), markers of neurodegeneration (including CSF T-tau and P-tau), inflammatory and other markers thought to be involved in the pathogenesis of AD.

Appendix 1
Protocol Synopsis (cont.)

- CLINICAL GENOTYPING SAMPLES	ApoE and FcγR
- *ROCHE CLINICAL REPOSITORY SPECIMENS*	Roche Clinical Repository DNA (blood) and non-DNA (plasma and RNA) specimens will be taken from consenting subjects The RCR sampling is optional.

PROCEDURES (summary):

Screening and baseline: After written informed consent is obtained, screening assessments, which include general health and cognitive testing, brain MRI, and lumbar puncture for CSF collection, must be completed within 8 weeks followed by up to 1 week for baseline assessments. (See schedule of assessments.)

Treatment: In this study, all subjects will receive 26 treatments (drug administrations) with a 4-week interval between each dose. On each dosing day, gantenerumab or matching placebo will be administered as [redacted] SC injections in the abdominal area. Safety and efficacy evaluations will be performed according to the schedule of assessments.

Treatment Extension: Subjects who complete the entire 2 years of treatment and the Week 104 follow-up visit may continue to receive the same treatment they were receiving at the Week 100 visit. The double-blind treatment can resume at the Week 104 visit after all assessments have been completed for that visit. If this amendment was not yet approved at the time of the subject's Week 104 visit, the subject can resume treatment at a subsequent visit; however, assessments as indicated in the Schedule of Assessments for Week 104Ext must be completed prior to administration of study medication.

Follow-up: Patients will be followed for *1 year (52* weeks*)* after the last dose of study treatment.

Futility Analysis: A futility interim analysis will be performed approximately when 50% of the sample size is scheduled to complete the Week 104 visit.

STATISTICAL ANALYSES:

All efficacy and safety data will be evaluated in the following manner:

- All gantenerumab dose groups combined versus placebo for overall assessments regardless of the ApoE genotype. All placebo patients will be pooled.
- Each of the two gantenerumab dose groups (i.e., 105 mg vs. 225 mg) versus placebo. The primary inferential statistics will be applied to the pairwise comparison of each of the two gantenerumab dose groups to the corresponding placebo group. The corresponding placebo group for the comparison to the gantenerumab 105 mg group is the entire placebo group. The corresponding placebo group for the comparison to the gantenerumab 225 mg group is the subgroup of placebo patients with ApoE 0E4 and 1E4 types (as there are no patients with ApoE 2E4 randomized to the

gantenerumab 225 mg arm). In order to adjust for the multiplicity in the pairwise comparisons of each gantenerumab dose to placebo, a hierarchical testing procedure will be used.

Each of the ApoE genotypes will also be assessed separately in comparison to the corresponding placebo group.

The efficacy variables of continuous type will be analyzed using a mixed-effects model repeated measures technique incorporating the data up to 104 weeks of treatment will be used to utilize all the data collected over time with consideration of the variance-covariance matrix of the repeated measures. Treatment difference between gantenerumab and placebo will be estimated for the change from baseline values at Week 104.

All safety variables will be summarized for each assessment time (including follow-up) using descriptive statistics.

Appendix 2
Schedules of Assessments

Table 1 Schedule of Assessments – First Year

Visit by week #	Screen	Baseline	1	4	8	12	16	20	24	28	32	36	40	44	48	52	53
Dose Number	**(8 wks)**	**1**		**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	
***Demographics*[12]**	X																
Medical/surgical history	X																
Physical Exam	X															X	
Neurological Exam	X															X	
Modified Hachinski	X																
Clinical Genotyping	X																
RCR Plasma and RNA	X[6]																X
RCR DNA	X[6]																
Vital Signs	X	P	X	P	P	P	P	P	P	P	P	P	P	P	P	P	X
ECG (triple if underlined)	X	P	X		P			P			P			P			X
Blood and Urine Labs	X[13]	P	X		X			X			X			X			X[13]
Urine Pregnancy Test[5]	X	P		P	P	P	P	P	P	P	P	P	P	P	P	P	
PK plasma		P	X		P			P						P			X
ADA samples		P	X		P			P						P			
Plasma Biomarkers		P						P						P			
CDR	X	P-1 wk							P							P	
FAQ, NPI-Q, GDS	X	P-1 wk							P							P	
MMSE	X	P-1 wk				P			P			P				P	
Dementia Assessment	X								P							P	
ADAS-Cog	X	P-1 wk				P			P			P				P	
FCSRT-IR	X[4]	P-1 wk				P						P					
CANTAB	X	P-1 wk							P							P	
C-SSRS		P-1wk		X	X	X	X	X	X	X	X	X	X	X	X	X	
LP/CSF, plasma PK and matching serum sample[3]	X[2]															10/20d[2]	
RCR CSF	X															10/20d[2]	
***MRI*[9]**	X[2]			10/20d	*	10/20d	*	10/20d	*		10/20d	*		10/20d	*		

Appendix 2
Schedules of Assessments (cont.)

Table 2 Schedule of Assessments – Second Year

Visit by week #	56	60	64	68	72	76	80	84	88	92	96	100	101	FU1/104[7]	FU2/112	*FU3/152*	Unsch.[1]	FU1/DO1	FU2/DO2	*FU3/DO3*
Dose Number	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**								
Physical Exam															X	*X*	X		X	*X*
Neurological Exam															X	*X*	X		X	*X*
RCR Plasma and RNA														X[6]				X[6]		
Vital Signs	P	P	P	P	P	P	P	P	P	P	P	P	X	X	X	*X*	X	X	X	*X*
ECG (triple if underlined)				P				P				P	X		X		X		X	
Blood and Urine Labs				X			X					X	X[13]		X	*X*	X	X[13]	X	*X*
Urine Pregnancy Test[5]	P	P	P	P	P	P	P	P	P	P	P	P			X		X		X	
PK plasma				P								P	X		X		X		X	
ADA samples				P								P			X		X		X	
Plasma Biomarkers				P								P			X		X		X	
CDR			P			P			P					X		*X*	X	X		*X*
FAQ, NPI-Q, GDS						P								X		*X*[14]	X	X		*X*[14]
MMSE			P			P			P					X			X	X		
Dementia Assessment						P								X			X	X		
ADAS-Cog			P			P			P					X			X	X		
FCSRT-IR			P						P					X			X	X		
CANTAB						P								X			X	X		
C-SSRS	X	X	X	X	X	X	X	X	X	X	X	X		X	X		X	X	X	
LP/CSF, plasma PK and matching serum sample[3]														10/20d[2]			X	10/20d[2]		
RCR CSF														10/20d[2]				10/20d[2]		
MRI[9]		10/20d	*				10/20d	*						10/20d[2]		X[8]	X	*10/20d*[2]		X[8]

P indicates Prior to Injections: P-1 wk is within 1 week prior the 1st dose; P is prior to the injections on the day of dosing.
ECG readings underlined above to be done in triplicate for subjects enrolled prior to protocol amendment D. Subjects enrolled after approval of amendment D will have triplicate ECG done only at Screen. All ECGs are sent for central reading.
DO1, DO2 *and DO3* are drop-out visits; FU1, FU2*, and FU3* are 4, 12*, and 52* weeks after last dose *of study treatment.*
LP/CSF and MRI: the times are days after the visit (e.g. 10/20d is 10-20 days. Attempt to schedule MRI closer to 10 days than 20 days as the result must be received prior to the next dose).
RCR CSF: after aliquoting the required samples, any remaining CSF fluid from consenting subjects will be saved for RCR sample.
Visit window: +/- 7 days for dosing days; +/- 3 days for visits that are 1 week post-dose (see Section 5.1.3).
* *The MRI result as determined by the central reader must be known before a subsequent dose can be given.*

Appendix 2
Schedules of Assessments (cont.)

Table 2 Schedule of Assessments – Second Year (Cont.)

[1] Any procedure or assessment marked may be done at an unscheduled visit as required

[2] LP performed in the morning (8 am−12 pm); MRI should not be done during the 3 days following an LP. The *LPs at* screening and *Week 104 (FU1/104, FU1/DO 1 or 104 Ext)* are mandatory. The LPs at 1 year and during the extension part of the study are optional. *The final LP in the extension (FU1/208 or ext208/DO1) is highly encouraged.*

[3] *LP/*CSF for biomarkers and PK assays, *plasma PK and* matching serum *samples:* at screen, Wk 52 *Wk FU1/104 (or FU1/DO 1 or Wk 104 Ext), Wk 156* and *Wk FU1/208 (or ExtFU1/DO 1).* See Section 5.2.14.2.

[4] FCSRT-IR performed up to 1 month prior to screening can be used to satisfy inclusion criteria

[5] Women who are of childbearing potential must have a urine pregnancy test done at the site prior to each dose.

[6] RCR *plasma and RNA* samples should be taken at the same time as the blood samples taken during the LP visit.

[7] Subjects continuing *directly* into *the* Extension, should follow Week 104 Ext schedule in Table 3. FU2/112 *and FU3/152* visit in Table 2 will not be performed.

[8] The final MRI should be done approximately 4 weeks before the final follow-up visit to allow review before the final visit.

[9.] *"Reset" MRI schedule if restarted on half dose during Part 1 can be found in Appendix 3.*

[10.] *"Reset" MRI schedule if entering the extension already on a "reset" MRI schedule can be found in Appendix 4.*

[11.] *"Reset" MRI schedule if restarted on half dose during Part 2 can be found in Appendix 5.*

[12.] At screening, demographic data collected will include years of education.

[13] *During the screening period, Week 53, the Week 101 or DO1 and, Week 204, or Ext DO1 only, HbA1C, folic acid and, B12, T4, free T4, and thyroid stimulating hormone levels will also be assessed.*

[14] *At FU, DO3, Ext FU3 or Ext DO3, only FAQ will be administered.*

Appendix 2
Schedules of Assessments (cont.)

Table 3 Schedule of Assessments – Extension Year 1

Visit by week #	104 Ext	108	112	116	120	124	128	132	136	140	144	148	152
Dose Number	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**
Physical Exam			P										
Neurological Exam			P										
RCR Plasma and RNA	*X*[6]												
Vital Signs	P	P	P	P	P	P	P	P	P	P	P	P	P
ECG			P							P			
Blood and Urine Labs			*X*							*X*			
Urine Pregnancy Test[5]	P	P	P	P	P	P	P	P	P	P	P	P	P
PK plasma			P							P			
ADA samples			P							P			
Plasma Biomarkers													
CDR	P			P			P			P			
FAQ, NPI-Q, GDS	P						P						
MMSE	P			P			P			P			
Dementia Assessment	P						P						
ADAS-Cog	P			P			P			P			
FCSRT-IR	P			P						P			
CANTAB	P												
C-SSRS	*X*	*X*	*X*	*X*	*X*	*X*	*X*	*X*	*X*	*X*	*X*	*X*	*X*
LP/CSF, plasma PK and matching serum sample[3]	10/20d[2]												
RCR CSF	10/20d[2]												
MRI[10, 11]	10/20d	*					10/20d	*					10/20d

Also refer to footnotes Table 2

Appendix 2
Schedules of Assessments (cont.)

Table 4 Schedule of Assessments – Extension Year 2

Visit by week #	156	160	164	168	172	176	180	184	188	192	196	200	204	FU1/208	FU2/216	FU3/256	Unsch.[1]	ExtFU1/ DO1	ExtFU2/ DO2	ExtFU3/ DO3
Dose Number	40	41	42	43	44	45	46	47	48	49	50	51	52							
Physical Exam	P												P		X	X	X		X	X
Neurological Exam	P												P		X	X	X		X	X
RCR Plasma and RNA	X[6]													X[6]				X[6]		
Vital Signs	P	P	P	P	P	P	P	P	P	P	P	P	P	X	X	X	X	X	X	X
ECG				P						P			P		X		X		X	
Blood and Urine Labs				X						X			X[13]		X	X	X	X[13]	X	X
Urine Pregnancy Test[5]	P	P	P	P	P	P	P	P	P	P	P	P	P		X		X		X	
PK plasma				P						P			P		X		X		X	
ADA samples				P						P			P		X		X		X	
Plasma Biomarkers															X		X		X	
CDR	P			P			P			P				X		X	X	X		X
FAQ, NPI-Q, GDS	P						P							X		X[14]	X	X		X[14]
MMSE	P			P			P			P				X			X	X		
Dementia Assessment	P						P							X			X	X		
ADAS-Cog	P			P			P			P				X			X	X		
FCSRT-IR				P						P				X			X	X		
C-SSRS	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	
LP/CSF, plasma PK and matching serum sample[3]	10/20d[2]													10/20d[2]			X	10/20d[2]		
RCR CSF	10/20d[2]													10/20d[2]				10/20d[2]		
MRI[10,11]	*					10/20d	*							10/20d[2]		X[8]	X	10/20d[2]		X[8]

Also refer to footnotes in Table 2.

Appendix 3
Alzheimer's Disease Assessment Scale–Cognition

The Alzheimer's Disease Assessment Scale (ADAS)-cognition, also referred to as ADAS-Cog, with 11-item total score and its modified version with 13-item total score are used in the study. The 13-item total score is calculated using the sum of all the 13-item scores shown below, whereas the 11-item total score is calculated using the sum of all-item scores except Item 4 (delayed word recall task) and Item 13 (number cancellation).

The item scores 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 Case Report Form (eCRF) for the three trials. The item score has a range of 0–10. If the "total not recalled" in any of the three trials is missing, then the item score will be set to missing.

- Item 2: following commands

  The item score is the number of incorrect responses on commands (i.e., no) for Questions 2a–e on the eCRF. The score has a range of 0–5.

- Item 3: constructional praxis

  The item score is available from Question 3e on the eCRF. The score has a range of 0–5.

- Item 4: delayed word-recall task

  The item score is "total not recalled" on the eCRF. The score has a range of 0–10.

- Item 5: naming objects and fingers

  The item score is derived on the basis of "total incorrect" recorded on the eCRF and determined as follows:

  0 = 0–2 incorrect

  1 = 3–5 incorrect

  2 = 6–8 incorrect

  3 = 9–11 incorrect

  4 = 12–14 incorrect

  5 = 15–17 incorrect

Appendix 3
Alzheimer's Disease Assessment Scale–Cognition (cont.)

- Item 6: ideational praxis

  The item score is equal to the "total incorrect" recorded on the eCRF for this item. The score has a range of 0–5.

- Item 7: orientation

  The item score is equal to the "total incorrect" recorded on eCRF for this item. The score has a range of 0–8.

- Item 8: word recognition

  The item score is equal to the "total incorrect" recorded on eCRF if "total incorrect" is ≤ 12. If the "total incorrect" is > 12, the item score is 12. The item score has a range of 0–12.

- Item 9: remembering test instructions

  The item score is derived on the basis of replies recorded on the eCRF and determined as follows: 0 = none, 1 = very mild, 2 = mild, 3 = moderate, 4 = moderately severe, and 5 = severe.

- Item 10: comprehension

  The item score is derived on the basis of replies recorded on the eCRF and determined as follows: 0 = none, 1 = very mild, 2 = mild, 3 = moderate, 4 = moderately severe, and 5 = severe.

- Item 11: word-finding difficulty

  The item score is derived on the basis of replies recorded on the eCRF and determined as follows: 0 = none, 1 = very mild, 2 = mild, 3 = moderate, 4 = moderately severe, and 5 = severe.

- Item 12: spoken language ability

  The item score is derived on the basis of replies recorded on the eCRF and determined as follows: 0 = none, 1 = very mild, 2 = mild, 3 = moderate, 4 = moderately severe, and 5 = severe.

Appendix 3
Alzheimer's Disease Assessment Scale–Cognition (cont.)

- Item 13: number cancellation

  The item score has a range of 0–5 with a maximum score of 5, given if the item is not performed for cognitive reasons. The item score is derived from the cancellation score, which is equal to the number of target hits (minus) the number of errors (minus) the number of times reminded of task, and is determined as shown below:

  0 = cancellation score is > 23

  1 = cancellation score is 18–22

  2 = cancellation score is 13–17

  3 = cancellation score is 9–12

  4 = cancellation score is 5–8

  5 = cancellation score is 0–4

  This 45-second version of scoring rule is the scoring rule that Alzheimer Disease Neuroimaging Initiative is using and is based on the proposal of the ADAS Instrument Committee.

STATISTICAL ANALYSIS PLAN

TITLE: **MULTICENTER, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-GROUP TWO YEAR STUDY TO EVALUATE THE EFFECT OF SUBCUTANEOUS RO4909832 ON COGNITION AND FUNCTION IN PRODROMAL ALZHEIMER'S DISEASE WITH OPTION FOR UP TO AN ADDITIONAL TWO YEARS OF TREATMENT AND AN OPEN-LABEL EXTENSION WITH ACTIVE STUDY TREATMENT; PART III: OPEN-LABEL EXTENSION FOR PARTICIPATING PATIENTS**

PROTOCOL NUMBER:	WN25203
STUDY DRUG:	Gantenerumab (RO4909832)
VERSION NUMBER:	H
IND NUMBER:	102,266
EUDRACT NUMBER:	2010-019895-66
SPONSOR:	Hoffmann-La Roche Ltd
PLAN PREPARED BY:	[redacted]
DATE FINAL:	See electronic date stamp below.

STATISTICAL ANALYSIS PLAN APPROVAL

Date and Time(UTC)	Reason for Signing	Name
13-Nov-2020 10:42:15	Company Signatory	[redacted]

TABLE OF CONTENTS

1. BACKGROUND 6

2. STUDY DESIGN 6

2.1 Protocol Synopsis 7

2.2 Outcome Measures 7

2.2.1 Efficacy Outcome Measures 7

2.2.1.1 Clinical Dementia Rating 8

2.2.1.2 Functional Activities Questionnaire 8

2.2.1.3 Mini Mental State Exam 8

2.2.1.4 Dementia Assessment 9

2.2.1.5 Alzheimer Disease Assessment Scale-Cognition 9

2.2.1.6 Free and Cued Selective Reminding Test – Immediate Recall 9

2.2.2 Safety Outcome Measures 9

2.2.3 Pharmacokinetic Outcome Measures 10

2.2.4 Biomarker Outcome Measures 10

2.3 Determination of Sample Size 10

2.4 Analysis Timing 10

3. STUDY CONDUCT 10

3.1 Randomization 10

3.2 Independent Imaging Review Facility 11

3.3 Data Monitoring 11

4. STATISTICAL METHODS 11

4.1 Analysis Populations 11

4.1.1 Safety Population 11

4.1.2 Safety Population with Post OLE Baseline MRI 12

4.2 Analyses of Study Conduct 12

4.3 Analyses of Treatment Group 12

4.3.1 Demographic Data and Baseline Characteristics 12

4.3.2 Previous and Concomitant Medications and Medical History 13

4.4 Efficacy Analyses 13

4.4.1 Efficacy Endpoint 14
4.5 Pharmacokinetic and Pharmacodynamic Analyses 14
4.5.1 PK and ADA Analysis 14
4.5.2 PD and Exploratory Biomarker Analyses 14
4.5.3 Magnetic Resonance Imaging Volumetric Measures 15
4.6 Safety Analyses 15
4.6.1 Exposure to Study Medication 15
4.6.2 Adverse Events 15
4.6.3 MRI Safety Findings and Injection-Site Reactions 16
4.6.4 Laboratory Data 17
4.6.5 Vital Signs and Electrocardiograms 17
4.6.6 Deaths 18
4.6.7 Subgroup Analyses 18
4.7 Interim Analyses 18
4.8 Analyses Related to COVID-19 18

LIST OF APPENDICES

Appendix 1 Protocol Synopsis 19
Appendix 2 Schedule of Assessments 25
Appendix 3 Alzheimer's Disease Assessment Scale - Cognition 32

GLOSSARY OF ABBREVIATIONS

AD	Alzheimer's disease
ADAS-Cog	Alzheimer Disease Assessment Scale - Cognition
AE	adverse event
AESI	adverse event of special interest
APOE	apolipoprotein E
ARIA	amyloid-related imaging abnormality
ARIA-E	amyloid-related imaging abnormality – edema
ARIA-H	amyloid-related imaging abnormality – hemosiderin
AUC_{0-T}	area under the plasma concentration-time curve
CCOD	clinical cutoff date
CDR	Clinical Dementia Rating
CDR-SOB	Clinical Dementia Rating – Sum of Boxes
C_{max}	maximum plasma concentration
C_{min}	minimum plasma concentration
CSF	cerebrospinal fluid
CSR	clinical study report
C-SSRS	Columbia−Suicide Severity Rating Scale
DO	dropout
ECG	electrocardiogram
eCRF	electronic Case Report Form
FAQ	Functional Activities Questionnaire
FCSRT-IR	Free and Cued Selective Reminding Test – Immediate Recall
FU	follow-up
iDMC	independent Data Monitoring Committee
IMC	Internal Monitoring Committee
LEAP	Learning Embeddings for Atlas Propagation
MedDRA	Medical Dictionary for Regulatory Activities
MMSE	Mini Mental State Exam
MRI-C	MRI Review Committee
OLE	open-label extension
PD	pharmacodynamic
PET	positron emission tomography
PK	pharmacokinetic
PT	preferred term

Q4W	every 4 weeks
QTcB	QT corrected using Bazett's formula
QTcF	QT corrected using Fridericia's formula
SAE	serious adverse event
SAP	statistical analysis plan
SC	subcutaneous
SE	safety-evaluable
SOC	system organ class
SUVR	standard uptake value ratio

1. BACKGROUND

This document describes the planned statistical analyses that will be reported in the final Clinical Study Report (CSR) for open-label extension (OLE) treatment (Part 3) of Study WN25203.

Study WN25203 was originally designed as a double-blind placebo-controlled proof-of-efficacy trial with safety data review to be performed by an independent Data Monitoring Committee (iDMC). As stated in Protocol Version H, the study has 3 parts:

- In Part 1 (double-blind), participants were randomized to gantenerumab or placebo, with 26 planned treatments every 4 weeks (Q4W).
- Participants who completed Part 1 of the study through the Week 104 visit could enter the optional (blinded) extension part of the study (Part 2) and continue receiving the same treatment for up to 2 additional years. However, dosing was suspended on 19 December 2014 after an interim analysis showed that the study was unlikely to meet the primary endpoint. Details are available in the published CSR (January 2016, with clinical cutoff date (CCOD) 22 June 2015).
- After a treatment-free follow-up period (approximately 12-18 months based on the approval of Protocol Version G), participants had the option to enter an open-label extension (Part 3) and receive higher doses of gantenerumab treatment.

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 and will be covered in the CSR content map.

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

2. STUDY DESIGN

Study WN25203 is a Phase III, multicenter, randomized, double-blind, placebo-controlled, parallel-group study, followed by an open-label extension with active study drug treatment for participants with prodromal Alzheimer's disease (AD). A total of 799 participants were randomized and 797 participants received gantenerumab or placebo. In Part 1, there were 266 participants in the placebo group, 271 participants in the gantenerumab 105 mg group, and 260 participants in the gantenerumab 225 mg group. Participants could enter Part 2 to receive treatment for up to 2 years of extension. After the completion of the double-blind treatment in Part 1 or Part 2, participants had the option to enter Part 3, the OLE, after completing a treatment-free follow-up period of approximately 12−18 months.

Overall,154 participants that had originally enrolled in Study WN25203 continued in Part 3 (OLE).

In the OLE, gantenerumab was administered subcutaneously (SC) Q4W with the following up-titration regimens to reach the target dose of 1200 mg gantenerumab.:

- Up-titration scheme I: dosing in Apolipoprotein E (APOE) ε4 carriers (1ε4 and 2ε4 genotypes) previously on placebo, in any participant previously on 105 mg gantenerumab, and in any participant previously on 225 mg gantenerumab meeting criteria for dose reduction due to amyloid-related imaging abnormalities (ARIA):
  - Doses 1, 2, and 3: 105 mg SC
  - Doses 4, 5, and 6: 225 mg SC
  - Doses 7 and 8: 450 mg SC
  - Doses 9 and 10: 900 mg SC
  - Dose 11 onwards: 1200 mg SC

- Up-titration scheme II: dosing in APOE ε4 non-carriers (0ε4 genotype) and in APOE ε4 carriers receiving 225 mg gantenerumab until the end of double-blind treatment (1ε4 genotype):
  - Doses 1 and 2: 225 mg SC
  - Doses 3 and 4: 450 mg SC
  - Doses 5 and 6: 900 mg SC
  - Dose 7 onwards: 1200 mg SC

2.1 PROTOCOL SYNOPSIS

The Protocol Synopsis and Schedule of Assessments can be found in Appendix 1 and Appendix 2 respectively.

2.2 OUTCOME MEASURES

The references for the efficacy assessment and procedures can be found in protocol Sections 5.2 and 11.

2.2.1 Efficacy Outcome Measures

Efficacy outcome assessments collected in the OLE include Clinical Dementia Rating scale (CDR), Functional Activities Questionnaire (FAQ), Mini Mental State Exam (MMSE), Dementia Assessment, Alzheimer Disease Assessment Scale-Cognition (ADAS-Cog), and Free and Cued Selective Reminding Test – Immediate Recall (FCSRT-IR). Definition and schedules of these measurements can be found in WN25203 protocol Section 5.2 and in the Schedule of Assessments in Appendix 2. Results of these efficacy analyses will be interpreted in a descriptive and exploratory fashion.

2.2.1.1 Clinical Dementia Rating

The CDR global score is calculated on the basis of 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 scores are recorded on the electronic Case Report Form (eCRF) by investigators and are the basis of analyses of CDR global scores in this study.

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 are given 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 $\geqslant$1.
- If M = 0, then CDR = 0; if there are impairments ($\geqslant$ 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 $\geqslant$ 1, then CDR cannot be 0; in such a circumstance, CDR= 0.5 when the majority of secondary categories are 0.

2.2.1.2 Functional Activities Questionnaire

The FAQ is an informant-based assessment that presents a forced choice of four levels of functioning for ten activities of daily living. Total scores for FAQ will be summarized.

2.2.1.3 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.

2.2.1.4 Dementia Assessment

Time to dementia is defined as the time interval between the first treatment date and the date that participant is assessed as having Alzheimer-type dementia by investigators.

2.2.1.5 Alzheimer Disease Assessment Scale-Cognition

The ADAS-Cog with an 11-item total score and its modified version with a 13-item total score are used in the study. The item scores are derived per description in Appendix 3. The 13-item total score is calculated using the sum of all the 13-item scores, whereas the 11-item total score is calculated using the sum of all item scores except Item 4 (delayed word recall task) and Item 13 (number cancellation). 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, whereas the 11-item ADAS-Cog total score has a range of 0−70. Higher scores indicate greater pathology.

2.2.1.6 Free and Cued Selective Reminding Test – Immediate Recall

The FCSRT-IR uses written words and immediate recall after each card with four words completed. The FCSRT-IR is a measure of memory under conditions that control attention and cognitive processing in order to obtain an assessment of memory not confounded by normal age-related changes in cognition. The performance on the FCSRT-IR has been associated with nonclinical and early dementia in several longitudinal epidemiologic studies.

2.2.2 Safety Outcome Measures

The following safety outcome measures will be assessed:

- Incidence and nature of MRI safety findings: amyloid-related imaging abnormality – edema (ARIA-E) and amyloid-related imaging abnormality – hemosiderin (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 baseline over time and incidence of abnormal laboratory values and abnormal laboratory values reported as adverse events
- Mean changes in electrocardiogram (ECG) assessments from baseline over time and incidence of abnormal ECG assessments
- Incidence of anti-gantenerumab antibodies
- 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)

2.2.3 Pharmacokinetic Outcome Measures

The following pharmacokinetic (PK) outcome measures will be assessed:

- Plasma concentrations at each PK sample time point

A population PK model previously developed using nonlinear mixed-effects modeling (Roche Report No. 1105289) will be used to derive individual measures of exposure at steady state such as:

- Maximum plasma concentration (C_{max})
- Trough plasma concentration (C_{min})
- The area under the plasma concentration-time curve ($AUC_{0-\tau}$)

2.2.4 Biomarker Outcome Measures

Cerebrospinal fluid (CSF) samples for biomarker analysis were not collected in the OLE. No exploratory biomarker analyses will be reported.

2.3 DETERMINATION OF SAMPLE SIZE

All individuals who had participated in Part 1 or Part 2 prior to the futility analysis and had at least one follow-up (FU)/dropout (DO) visit after the double-blind treatment (placebo or 105 mg or 225 mg gantenerumab SC Q4W) were offered the opportunity to participate in the OLE. A total of 154 participants enrolled in the OLE. The determination of sample size for the double-blind section of the study can be found in the protocol Section 8.3.

2.4 ANALYSIS TIMING

The analysis covered in this SAP will include all data collected from participants enrolled in the OLE. Participants started transitioning from the double blind to the OLE in February 2016, receiving gantenerumab based on up-titration scheme I or up-titration scheme II up to OLE Week 236. Data from all visits in the OLE until the end of the study will be reported. In addition, safety data collected between the end of the double-blind and the start of OLE will be included.

The start of the OLE is defined as the date when participants received the first dose of gantenerumab in the OLE. The end of the study will be considered to be the date of the last visit (including the last scheduled follow-up visit according to study protocol) of the last participant in the study.

3. STUDY CONDUCT

3.1 RANDOMIZATION

Participants enrolled in the OLE were not randomized. Depending on the APOE ε4 carrier status and/or gantenerumab received previously in Part 1 and Part 2, OLE participants were assigned to one of two up-titration dosing schemes. Details can be found in Section 2.

3.2 INDEPENDENT IMAGING REVIEW FACILITY

All MRI scans were read and reported by the central reader in alignment with the processes from Parts 1 and 2 of the trial. 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. During Parts 1 and 2, the MRI-C reviewed MRI findings as described in protocol Sections 5.1.3 and 5.1.4.3, respectively. In Part 3, the committee reviewed relevant MRI findings as defined in the MRI-C Charter.

3.3 DATA MONITORING

An Internal Monitoring Committee (IMC) was set up in the OLE to review the accumulating safety and efficacy data in an open-label fashion on a quarterly basis until December 2018. Thereafter, the IMC reviewed the accumulating safety and efficacy data every 6 months, supplemented by ad-hoc cross-functional review of the data on a per need basis.

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.

Statistical analyses will primarily focus on safety aspects during and after the up-titration phase in OLE. The efficacy analyses, including cognitive/functional measures, will be analyzed and interpreted in an exploratory fashion.

The start of the OLE is defined as the date when participants 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 open-label extension. Per protocol, participants eligible for OLE should be confirmed during the OLE pre-baseline and MRI should be performed within 8 weeks, with results available before the first dosing visit in the OLE.

4.1 ANALYSIS POPULATIONS

4.1.1 Safety Population

The safety-evaluable (SE) population includes participants who received any dose of gantenerumab during OLE. Analysis using this population will be performed by previous randomization.

4.1.2 Safety Population with Post OLE Baseline MRI

The participants in the SE population which had at least one post-OLE baseline MRI.

4.2 ANALYSES OF STUDY CONDUCT

The following will be summarized using descriptive statistics:

- Number of participants enrolled, and the number and percentage of participants who prematurely withdrew from the study (including the reasons for discontinuation and the distribution of these discontinuations by visit)
- Incidence of protocol deviations – overall and by four main categories (inclusion criteria, exclusion criteria, medication and procedural)

4.3 ANALYSES OF TREATMENT GROUP

4.3.1 Demographic Data and Baseline Characteristics

Baseline characteristics will be analyzed descriptively for the SE population by previous randomization. The OLE baseline is defined as the last available assessment up to and including the first day of study drug intake in OLE.

Treated patients have received at least one dose of study drug. The patient demographic and screening/OLE baseline variables include the following:

- Sex
- Age at screening and at OLE baseline
- Race
- Years of education
- Weight
- Height
- Body mass index
- Region
- APOE $\varepsilon 4$ carrier status (yes vs no)
- APOE genotype ($0\varepsilon 4$, $1\varepsilon 4$, or $2\varepsilon 4$)
- ARIA-E during double blind period (yes vs no)
- Cumulative ARIA-H at OLE baseline
- FAQ, total score at OLE baseline
- FCSRT-IR, total score at OLE baseline
- MMSE, total score at OLE baseline
- ADAS-cog (11-item score, 13-items score) at OLE baseline
- CDR-SOB at OLE baseline
- CDR-Global score at OLE baseline

4.3.2 Previous and Concomitant Medications and Medical History

The medical history, previous medication and concomitant medication data collected at the double blind baseline will be summarized by previous randomization for the participants enrolled in the OLE.

A patient's medical history is defined as a record of all medical events (for example: procedures and diseases) that start prior to the date of the first dose of study treatment at the start of double blind. 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 at the start of double blind is classified as previous-concomitant, while any medication starting after the first dose of study treatment at the start of double blind is classified as concomitant.

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).

4.4 EFFICACY ANALYSES

All efficacy analyses will be descriptive and interpreted in a strictly exploratory fashion. The baseline visit includes all assessments performed on OLE Day 1 or before, and the first post-baseline visit begins on study Day 2.

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.

Descriptive statistics will be summarized by previous randomization.

4.4.1 Efficacy Endpoint

Efficacy outcomes of CDR, FAQ, MMSE, ADAS-Cog, and FCSRT-IR at OLE baseline, Week 24/52/76/104/128/156/FU were assessed and collected. For each of these endpoints, the total scores, change from baseline and percentage change from baseline will be summarized by visit.

4.5 PHARMACOKINETIC AND PHARMACODYNAMIC ANALYSES

4.5.1 PK and ADA Analysis

A population PK model previously developed using nonlinear mixed-effects modeling (Roche Report No. 1105289) will be used to derive individual measures of exposure including C_{max}, C_{min}, and AUC_{0-T} at steady state. These parameters will be descriptively summarized including arithmetic mean, median, standard deviation, range, and CV.

Plasma concentration data for gantenerumab will be tabulated and summarized.

Descriptive summary statistics will include the arithmetic mean, median, range, SD, and coefficient of variation, as appropriate.

Data on anti-gantenerumab antibodies will be summarized.

4.5.2 PD and Exploratory Biomarker Analyses

The positron emission tomography (PET) substudy was initiated in the double-blind period and only included 112 participants. Florbetapir ^{18}F (Amyvid ^{TM}) was used for participants in the amyloid-PET substudy of brain amyloid imaging as a pharmacodynamic (PD) measurement of potential effects of gantenerumab on amyloid load in the brains of subjects with early (prodromal) Alzheimer's disease. When participants continued into the OLE at centers already involved in the PET substudy, participants at those centers were offered to take part in the PET substudy. Additional centers could be included and could offer OLE participants to enroll in the PET substudy.

A prespecified standard uptake value ratio (SUVR) identical to the method report for SR double-blinded analyses will be used. The conversion is a linear transformation which is specified using the mean cerebellar gray reference region:

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

Changes in cerebellar gray-cortical composite SUVR from OLE baseline in amyloid load will be summarized using descriptive statistics.

4.5.3 Magnetic Resonance Imaging Volumetric Measures

The volumetric data for particiants in the OLE were measured by the Learning Embeddings for Atlas Propagation (LEAP) technology, which computed the cortical thickness. The data received were the percentage change from double-blind screening. The percentage changes from screening in hippocampal volume, whole brain volume, ventricular volume, and cortical thickness will be summarized by previous randomization using descriptive statistics. Available time points include OLE baseline, OLE Week 48, OLE Week 100, and OLE Week 152/FU/DO.

4.6 SAFETY ANALYSES

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

For all safety analyses, Day 1 is defined as the start date of gantenerumab treatment in the OLE (Part 3). All analyses, except those for ARIA, will be summarized by previous randomization. For ARIA analyses, outputs will be summarized by up-titration scheme and/or carrier status.

4.6.1 Exposure to Study Medication

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

- Treatment duration (years)
- Total number of administrations
- Total cumulative gantenerumab dose (mg)
- Duration on 105mg / 225mg / 450mg / 900mg / 1200mg gantenerumab dosing steps
- Total patient years on gantenerumab: time from first dose in OLE until CCOD or withdrawal from treatment/study (whichever occurs first)

4.6.2 Adverse Events

AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version that is current at the time of the analysis (Version 22.1). For each treatment group, the frequency of each AE preferred term (PT) will be defined as the number of participants experiencing at least one occurrence of the event. As participants will have had different amounts of time in the study, the incidence rate will be calculated as the occurrences divided by the total number of patient-years exposure for the participants in the population and treatment group specified. Each table will also present the overall number of participants experiencing at least one AE and the total number of AEs reported. Participant listings will contain PTs and comments for each event. In summary tables, AEs will be sorted by system organ class (SOC) (in decreasing order of overall incidence), then by PT (in decreasing order of overall incidence).

The following safety information will be summarized:

- AEs
- SAEs
- Study drug-related AEs
- Study drug-related SAEs
- AEs leading to discontinuation of study treatment
- AEs leading to dose reduction
- AEs events leading to no further up-titration
- AEs leading to dose interruption
- AEs of special interest (AESIs)

For scientific interpretation, it is important to separate ARIA-H and ARIA-E events in listings and tables.Using MedDRA standards, ARIA-E is mapped to the following PTs: “AMYLOID RELATED IMAGING ABNORMALITY-OEDEMA/EFFUSION” or “VASOGENIC CEREBRAL OEDEMA”. ARIA-H is mapped to “CEREBELLAR MICROHAEMORRHAGE" or "CEREBRAL HAEMOSIDERIN DEPOSITION" or "CEREBRAL MICROHAEMORRHAGE" or "AMYLOID RELATED IMAGING ABNORMALITY-MICROHAEMORRHAGES AND HAEMOSIDERIN DEPOSITS”

AESIs for this study are as follows:

- An elevated ALT or AST in combination with either an elevated bilirubin or clinical jaundice
- Suspected transmission of an infectious agent by the study drug

AESIs are discussed in more detail in protocol Section 7.1.1.4. AESIs in this study will be captured by recording all case details on the AE eCRF.

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

4.6.3 MRI Safety Findings and Injection-Site Reactions

[redacted]

[redacted]

Injection-site reactions will be listed and summarized by previous randomization and by up-titration scheme.

4.6.4 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 participants 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.

4.6.5 Vital Signs and Electrocardiograms

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

ECGs are collected throughout the study. 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
- 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).

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.

4.6.6 Deaths

Deaths will be listed.

4.6.7 Subgroup Analyses

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

4.7 INTERIM ANALYSES

No efficacy interim analyses were planned in the OLE.

4.8 ANALYSES RELATED TO COVID-19

Listings of AEs associated with COVID-19 and major protocol deviations related to COVID-19 will be prepared.

Appendix 1
Protocol Synopsis

SYNOPSIS OF PROTOCOL NUMBER WN25203H

TITLE	Multicenter, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group Two Year Study to Evaluate the Effect of Subcutaneous RO4909832 on Cognition and Function in Prodromal Alzheimer's Disease with Option for up to an Additional Two Years of Treatment and an Open-Label Extension with Active Study Treatment		
SPONSOR	Roche	Clinical Phase	3
INDICATION	Prodromal Alzheimer's Disease		
OBJECTIVES OF DOUBLE-BLIND PART OF THE STUDY Primary	To evaluate the effect of gantenerumab versus placebo on the change in the Clinical Dementia Rating scale Sum of Boxes (CDR-SOB), a global measure of cognition and functional ability.		
Secondary Cognition, global, and functioning	To evaluate the effect of gantenerumab versus placebo on cognition (assessed with the Alzheimer Disease Assessment Scale-Cognition [ADAS-Cog], Mini Mental State Exam [MMSE], Cambridge Neuropsychological Test Automated Battery [CANTAB], and the Free and Cued Selective Reminding Test-Immediate Recall [FCSRT-IR]) on global measures (based on the CDR global score), and on functioning (assessed with the Functional Activities Questionnaire [FAQ]), on neuropsychiatric functioning (assessed with the Neuropsychiatric Inventory Questionnaire [NPI-Q]) and on time to onset of dementia.		
Safety	To assess the safety and tolerability of gantenerumab assessed by magnetic resonance imaging (MRI), physical and neurological examinations, vital signs, blood and urine safety tests, ECGs, Columbia-Suicide Severity Rating Scale (C-SSRS), Geriatric Depression Scale (GDS), and adverse event monitoring.		
MRI volumetry	To evaluate the effect of gantenerumab versus placebo on hippocampal volume, whole brain volume, ventricular enlargement, and possibly other volumetric measures of the brain.		
CSF Biomarkers	To evaluate the effect of gantenerumab versus placebo on levels of cerebrospinal fluid (CSF) biomarkers of Alzheimer's disease (AD) pathology (i.e., total tau [T-tau], phospho-tau [p-tau], and amyloid-beta ($A\beta_{1-42}$]).		
Pharmacokinetics/ Anti-Drug Antibodies	To determine the relationship of plasma and CSF concentrations of gantenerumab on other responses. To assess incidence of anti-gantenerumab antibodies, and if relevant, evaluate its effect on the pharmacokinetic (PK), pharmacodynamic (PD), efficacy. and safety parameters.		

Exploratory Plasma Biomarkers	To evaluate the effect of treatment on peripheral biomarkers of amyloid deposition and/or clearance, neurodegeneration, inflammation, and other markers known to be involved in the pathogenesis of AD.
Clinical Genotyping	To evaluate the effect of gene encoding apoliprotein E (ApoE) ε4 genotype and fragment crystalizable gamma receptor (FcγR) genotype on PK/PD, efficacy, and safety parameters
OBJECTIVES OF THE OPEN-LABEL EXTENSION - PART 3	
Primary: Safety:	To assess short-term and long-term safety and tolerability of gantenerumab (RO4909832) given at doses up to 1200 mg subcutaneous (SC) every 4 weeks (Q4W) by MRI, physical and neurological examinations, vital signs, blood safety tests, ECGs, C-SSRS, and adverse event monitoring.
Secondary:	
Brain MRI Volumetry	To evaluate the effect of gantenerumab over time on hippocampal volume, whole brain volume, ventricular enlargement, and possibly other volumetric measures of the brain compared to baseline and to the start of Part 3
Brain Amyloid Load by Positron Emission Tomography (PET) Imaging	To assess changes in amyloid load over time using Florbetapir ^{8}F (AV-45; Amyvid™) compared to screening (if applicable) and to the start of Part 3. PET scans performed within 12 months prior to the first study drug administration in Part 3 will be used for the baseline of Part 3. This objective will be evaluated in a subset of consenting subjects participating in the PET substudy.
Pharmacokinetics	To determine the relationship of plasma concentrations of gantenerumab on other responses
Anti-Drug Antibodies	To assess the development of anti-gantenerumab antibodies and, if detected, explore their association with the PK, PD, efficacy, and safety parameters of gantenerumab.
Cognition, Global, Functioning, and Dementia Assessment	To evaluate the effect of gantenerumab over time as assessed with the CDR-SOB, the ADAS-Cog, the MMSE, the FCSRT-IR, the CDR global score, the FAQ compared to baseline and to the start of Part 3, and to determine the presence of and time to onset of dementia.
TRIAL DESIGN	Multicenter, randomized, double-blind, placebo-controlled, parallel-group study, followed by open-label extension with active study drug treatment

NUMBER OF SUBJECTS	Approximately a total of 770 subjects; 514 on gantenerumab and 256 on placebo. The sample size was confirmed following a planned blinded evaluation of the variance in the change in the CDR-SOB at Week 76. The study enrolled 799 subjects, and 797 subjects were dosed with double-blind treatment: 266 subjects in the placebo group, 271 subjects in the gantenerumab 105 mg group, and 260 subjects in the gantenerumab 225 mg group. The number of subjects *enrolled in Part 3 is 154*.
TARGET POPULATION	Subjects who meet memory and biomarker criteria which indicate a strong likelihood that they are in the prodromal phase of Alzheimer's Disease
LENGTH OF STUDY	**Double-blind treatment period:** Part 1: Approximately 2 years and 2 months which includes 8 weeks for screening and 2 years (104 weeks) of treatment. Part 2: Up to additional 2 years of treatment. Dosing was suspended on 19 December 2014. As a result duration of the double-blind treatment period varied depending on the number of study weeks that subjects completed. **Treatment-free follow-up period:** Twelve weeks of follow up after completion of the double-blind treatment in Part 1 or Part 2, with an optional follow-up visit approximately 1 year (52 weeks) after the end of the double-blind treatment. Approximate duration of the study for subjects completing both Part 1 and Part 2 with the maximum 52-week follow-up and not continuing in Part 3 will be 5 years and 2 months. For the last subjects enrolled, approximate duration of the study will be 2 years and 2 months (year 1 of Part 1 and 1 year of treatment-free follow-up). **Open-label extension** Part 3: Up to 3 years of active study drug treatment (with a 4-week follow-up visit) after treatment-free follow-up period. For subjects continuing in Part 3, treatment-free follow-up period will last approximately between 12 and 18 months (depending on the approval of study protocol WN25203G by local Competent Authorities). *All subjects who complete 3 years of open-label treatment will be given the option of extending open-label treatment until July 2020 and may receive up to 21 additional administrations every 4 weeks during Years 4 and 5.*
END OF STUDY	The end of the study will be considered to be the date of the last visit (including the last scheduled follow-up visit according to study protocol) of the last subject in the study.

INVESTIGATIONAL MEDICAL PRODUCT(S) DOSE/ROUTE/REGIMEN	Part 1 and Part 2: Gantenerumab (105 or 225 mg for the 0ε4 and 1ε4 ApoE genotypes or 105 mg for the 2ε4 ApoE genotype subjects) by SC injection Q4W Part 3 (open-label extension): Gantenerumab will be administered SC Q4W to the subjects using up-titration regimens dependent on APOE ε4 genotype (and on the dose received during the double-blind treatment period) that will allow subjects to reach the target dose of 1200 mg gantenerumab within 10 and 6 months for titration schedules below. Titration schedule 1: Dosing in APOE ε4 carriers (1ε4 ApoE and 2ε4 ApoE genotypes who were in the placebo and 105 mg gantenerumab groups, and in the 225 mg gantenerumab group meeting criteria for dose reduction due to amyloid-related imaging abnormalities (ARIAs) during double-blind treatment period): • Doses 1, 2, and 3: 105 mg • Doses 4, 5, and 6: 225 mg • Doses 7 and 8: 450 mg • Doses 9 and 10: 900 mg • Dose 11 onwards: 1200 mg Titration schedule 2: Dosing in subjects who are APOE ε4 non-carriers (0ε4 genotype) and in APOE ε4 carriers receiving 225 mg gantenerumab until the end of double-blind treatment (1ε4 genotype): • Doses 1 and 2: 225 mg • Doses 3 and 4: 450 mg • Doses 5 and 6: 900 mg • Dose 7 onwards: 1200 mg Gantenerumab doses of 105 and 225 mg will be administered SC [redacted] The 1200-mg gantenerumab dose will be administered SC [redacted].
COMPARATOR DOSE/ROUTE/REGIMEN	Placebo, by SC injection, Q4W in Part 1 and Part 2.
PART 3 ASSESSMENTS OF:	
- EFFICACY	CDR-SOB, ADAS-Cog, MMSE, FAQ, FCSRT-IR, CDR global, onset of dementia, *during the initial 3 years of Part 3. After which only MMSE will be administered.*
- SAFETY	MRI, physical and neurological examinations, vital signs, laboratory assessments (blood safety tests, urine pregnancy [for women of childbearing potential]), Columbia-Suicide Severity Rating Scale (C-SSRS), ECGs, and adverse event monitoring
- PHARMACOKINETICS	Gantenerumab levels in plasma
- BIOMARKERS	MRI and PET, to assess changes in brain volumetry and brain amyloid load over time *during the initial 3 years of Part 3*, respectively.

PROCEDURES (summary):

Screening and baseline: After written informed consent is obtained, screening assessments, which include general health and cognitive testing, brain MRI, and lumbar puncture for cerebrospinal fluid (CSF) collection, and must be completed within 8 weeks followed by up to 1 week for baseline assessments. (See schedule of assessments in protocol)

Treatment (Part 1): In this study, all subjects will receive 26 treatments (drug administrations) with a 4-week interval between each dose. On each dosing day, gantenerumab or matching placebo will be administered as [redacted] SC injections in the abdominal area. Safety and efficacy evaluations will be performed according to the schedule of assessments.

Treatment Extension (Part 2): Subjects who complete the entire 2 years of treatment and the Week 104 follow-up visit may continue to receive the same treatment that they were receiving at the Week 100 visit. The double-blind treatment can resume at the Week 104 visit after all assessments have been completed for that visit.

Follow-up: Subjects will be followed for 12 weeks, with an optional follow-up visit approximately 1 year (52 weeks) after the last dose of study treatment in Part 1 or Part 2 (double-blind treatment phase). The Week 52 visit is not required for subjects entering Part 3.

Futility Analysis: A futility interim analysis was performed when approximately 50% of the sample size completed the Week 104 visit. As the study was declared futile, with a low likelihood of meeting the prespecified primary outcome measure with the doses studied (105 and 225 mg), dosing was suspended on 19 December 2014. This analysis did not reveal any new safety signal.

Open-Label Extension (Part 3): Additional analyses of efficacy and PD results (CSF biomarkers of neurodegeneration and amyloid PET standardized uptake value ratio [SUVR]) indicated that higher doses of gantenerumab may have clinically relevant effect on cognition and function. Doses were modeled based on the results of the PET substudy (brain amyloid PET SUVR) and on the supportive results of the Phase 1b (PRIME) study, investigating the effects of aducanumab on a total of 165 patients with prodromal or mild AD. Aducanumab is a monoclonal antibody with similar properties and mode of action as gantenerumab. The results showed a dose- and time-dependent reduction of amyloid plaques in the brain that was associated with dose-dependent clinical effect, which reached statistical significant clinical effect with the highest aducanumab dose tested. Extrapolation of the data indicated that approximately 20% reduction in brain amyloid detected by PET may be needed to achieve clinical effect, compared to overall approximately 5% achieved with 225 mg gantenerumab over 2 years.

All subjects who were eligible for and enrolled into protocol WN25203 and had at least one follow-up/dropout visit after the double-blind treatment will be offered the

opportunity to participate in Part 3. Subjects who meet any of the following criteria will be exempted:

- Prematurely discontinued from Part 1 and Part 2 for safety reasons (e.g., MRI findings meeting criteria for treatment discontinuation in Part 1 and Part 2)
- [redacted]
- Received another investigational medication after the end of double-blind treatment
- Participation in Part 3 deemed inappropriate by investigator or Sponsor

During Part 3 study, subjects will receive gantenerumab Q4W that will be uptitrated starting from 105 or 225 mg based on the ApoE genotype to up to 1200 mg to reach the doses expected to give clinically meaningful effect in a less than 1 year. The titration schemes were designed to optimally manage ARIA events in ApoE e4 carriers and non-carriers by a PK-PD model using PK and safety results from this study and from similarly acting anti-amyloid antibodies (aducanumab and bapineuzumab). In addition, all subjects entering Part 3 at centers already involved in a WN25203 PET substudy of brain amyloid imaging during Part 1 and Part 2 will be offered to take part in the PET substudy, independent of their previous participation during the double-blind treatment. Subjects will be followed until the end of treatment phase (4 weeks after the last dose).

STATISTICAL ANALYSES IN THE OPEN-LABEL EXTENSION:

Statistical analyses will primarily focus on safety aspects during and after the uptitration phase in Part 3 of the trial. In addition, it is planned that efficacy analyses will be conducted on cognitive and functional measures as well as biomarkers. It is planned that efficacies will be calculated using both, the original, drug-naïve baseline as well as the baseline that gets available before start of *Part* 3. Results of these analyses will be interpreted in an exploratory fashion.

Appendix 2
Schedule of Assessments

Table 1 Schedule of Assessments – Open-Label Extension Year 1 - for ApoE ε4 carriers previously on: 1) Placebo, 2) 105 mg Gantenerumab and 3) 225 mg Gantenerumab meeting Criteria for Dose Reduction due to ARIAs

Visit by week #	OLE Pre-baseline (8 weeks)	Day 1 (OLE Baseline)	4	8	12	16	20	24	28	32	33	36	40	44	48	52
Dose Number		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**		**10**	**11**	**12**	**13**	**14**
Dose (mg)		**105**	**105**	**105**	**225**	**225**	**225**	**450**	**450**	**900**		**900**	**1200**	**1200**	**1200**	**1200**
Physical Exam		P						P								P
Neurological Exam		P						P								P
Vital Signs		P	P	P	P	P	P	P	P	P		P	P	P	P	P
ECG		P														P
Blood Labs		P														X
Urine Pregnancy Test[4]		P	P	P	P	P	P	P	P	P		P	P	P	P	P
PK plasma		P								P	X	P		P		
ADA samples		P								P		P				
CDR		P-1wk						P								P
FAQ		P-1wk						P								P
MMSE		P-1wk						P								P
Dementia Assessment		P-1wk						P								P
ADAS-Cog		P-1wk						P								P
FCSRT-IR		P-1wk						P								P
C-SSRS		P-1wk						X								X
MRI	X			X[2]	*		X[2]	*	X[2]	*		X[2]	*		X[2]	*

Refer to **Table 4** footnotes.

Table 2 Schedule of Assessments – Open-Label Extension Year 1- for: 1) ApoE ε4 non-carriers and 2) ApoE ε4 Carriers receiving 225 mg Gantenerumab until the End of Double-blind Treatment

Visit by week #	OLE Pre-baseline (8 weeks)	Day 1 (OLE Baseline)	4	8	12	16	17	20	24	28	32	36	40	44	48	52
Dose Number		1	**2**	**3**	**4**	**5**		**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Dose (mg)		**225**	**225**	**450**	**450**	**900**		**900**	**1200**	**1200**	**1200**	**1200**	**1200**	**1200**	**1200**	**1200**
Physical Exam		P							P							P
Neurological Exam		P							P							P
Vital Signs		P	P	P	P	P		P	P	P	P	P	P	P	P	P
ECG		P														P
Blood Labs		P														X
Urine Pregnancy Test[4]		P	P	P	P	P		P	P	P	P	P	P	P	P	P
PK plasma		P				P	X	P		P						
ADA samples		P				**P**				P						
CDR		P-1wk							P							P
FAQ		P-1wk							P							P
MMSE		P-1wk							P							P
Dementia Assessment		P-1wk							P							P
ADAS-Cog		P-1wk							P							P
FCSRT-IR		P-1wk							P							P
C-SSRS		P-1wk							X							X
MRI	X		X[2]	*	X[2]	*		X[2]	*		X[2]	*			X[2]	*

Refer to **Table 4** footnotes.

Table 3 Schedule of Assessments – Open-Label Extension Year 2 for all Participants

Visit by week #	56	60	64	68	72	76	80	84	88	92	96	100	101	104
Dose Number	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**		**27**
Physical Exam														P
Neurological Exam														P
Vital Signs	P	P	P	P	P	P	P	P	P	P	P	P		P
ECG														P
Blood Labs														X
Urine Pregnancy Test[4]	P	P	P	P	P	P	P	P	P	P	P	P		P
PK plasma			P									P	X	P
ADA samples			P											P
CDR						P								P
FAQ						P								P
MMSE						P								P
Dementia Assessment						P								P
ADAS-Cog						P								P
FCSRT-IR						P								P
C-SSRS						X								X
MRI					X[2]	*						X[2]		*

Refer to **Table 4** footnotes.

Table 4 Schedule of Assessments – Open-Label Extension Year 3 for all Participants

Visit by week #	108	112	116	120	124	128	132	136	140	144	148	152	OLE156/ FU/DO [5]	Unsch.[1]
Dose Number	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**		
Physical Exam													X	X
Neurological Exam													X	X
Vital Signs	P	P	P	P	P	P	P	P	P	P	P	P	X	X
ECG													X	X
Blood Labs													X	**X**
Urine Pregnancy Test[4]	P	P	P	P	P	P	P	P	P	P	P	P	X	X
PK plasma								P					X	X
ADA samples								P					X	X
CDR						P							X	X
FAQ						P							X	X
MMSE						P							X	X
Dementia Assessment						P							X	X
ADAS-Cog						P							X	X
FCSRT-IR						P							X	X
C-SSRS						X							X	X
MRI					X2	*						**X2,3**	*	X

Table 4 Schedule of Assessments – Open-Label Extension Year 3 for all Participants (cont.)

Note: Visit window: +/- 7 days for dosing days; +/- 3 days for visits that are 1 week post-dose.

OLE = Open-Label Extension; P = prior to study drug administration: P-1 wk is within 1 week prior the first dose.

Eligibility for Part 3 should be confirmed during OLE pre-baseline, and MRI should be performed within 8 weeks and the results available before the first dosing visit in the OLE. Additional assessments can be performed as per the investigator's judgment.

MRI scans must be performed within a maximum of 20 days *(ideally 10–20 days)* after dose administration and results made available and reviewed before the next scheduled dose.

* The MRI result as determined by the central reader must be known before a subsequent dose can be given.

[1] *Any procedure or assessment may be performed at an unscheduled visit as required.*

[2] Subjects will be asked if they experience CNS adverse events up to 1 week before each MRI is performed.

[3] The results of the final MRI should be available before the final follow-up visit to allow review before the final visit *(after the last dose for patients not continuing in Part 3 extension after 3 years)*.

[4] Women who are of childbearing potential must have a urine pregnancy test done at the site prior to each dose.

[5] Follow-up/drop-out visit to be performed 4 weeks after the last dose during the initial 152 weeks of treatment.

Table 5 Schedule of Assessments – Open-Label Extension Years 4 and 5 for Participants Continuing Part 3 until July 2020

Visit by week #	*156*	*160–176*	*180*	*184–204*	*208*	*212–Onwards (Up to 236)*	*OLE/FU/DO* [5]	*Unsch.* [1]
Dose Number	*40*	*41–45*	*46*	*47–52*	*53*	*54–Onwards (Up to 60)*		
Physical Exam	P				P		X	X
Neurological Exam	P				P		X	X
Vital Signs	P	P	P	P	P	P	X	X
ECG	P				P		X	X
Blood Labs	P				P		X	X
Urine Pregnancy Test[4]	P	P	P	P	P	P	X	X
PK plasma	P				P		X	X
ADA samples	P				P		X	X
CDR	P							
FAQ	P							
MMSE	P		P		P		X	X
Dementia Assessment	P							
ADAS-Cog	P							
FCSRT-IR	P							
C-SSRS	P		P		P		X	X
MRI	*	L[2, 3]	*	L[2, 3]	*	L[2, 3]	*	X

Table 5 Schedule of Assessments – Open-Label Extension Years 4 and 5 for Participants Continuing Part 3 until July 2020 (cont.)

Note: Visit window: +/- 7 days for dosing days.

L =an MRI after the Last dose within the indicated interval; OLE=Open-Label Extension; P =prior to study drug administration and prior to the injections on the day of dosing.

The schedule of assessment outlines the maximum number of week visits and doses until July 2020 and serves as guidance. The duration of Part 3 extension will vary for each subject dependent on when the subject begins the OLE.

MRI scans are scheduled at approximately 6-month intervals and is mandatory after the last study drug administration. The time between the previous scan and the final study scan should not exceed 8 months.

An MRI must be performed within 20 days (ideally 10–20 days) after dose administration and results made available and reviewed before the next scheduled dose.

** The MRI result as determined by the central reader must be known before a subsequent dosing visit and should be available before the OLE follow-up/drop-out visit to allow review before the final visit.*

[1] *Any procedure or assessment may be performed at an unscheduled visit as required.*

[2] *Subjects will be asked if they experience CNS adverse events up to 1 week before each MRI is performed.*

[3] *The results of the final MRI (after the last dose) should be available before the final follow-up visit to allow review before the final visit.*

[4] *Women who are of childbearing potential must have a urine pregnancy test done at the site prior to each dose.*

[5] *OLE follow-up/drop-out visit to be performed 4 weeks after the last dose in Years 4or 5 of Part 3.*

Appendix 3
Alzheimer's Disease Assessment Scale - Cognition

The Alzheimer's Disease Assessment Scale (ADAS)-cognition, also referred to as ADAS-Cog, with 11-item total score and its modified version with 13-item total score are used in the study. The 13-item total score is calculated using the sum of all the 13-item scores shown below, whereas the 11-item total score is calculated using the sum of all-item scores except Item 4 (delayed word recall task) and Item 13 (number cancellation).

The item scores 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 Case Report Form (eCRF) for the three trials. The item score has a range of 0–10. If the "total not recalled" in any of the three trials is missing, then the item score will be set to missing.

- Item 2: following commands

  The item score is the number of incorrect responses on commands (i.e., no) for Questions 2a–e on the eCRF. The score has a range of 0–5.

- Item 3: constructional praxis

  The item score is available from Question 3e on the eCRF. The score has a range of 0–5.

- Item 4: delayed word-recall task

  The item score is "total not recalled" on the eCRF. The score has a range of 0–10.

- Item 5: naming objects and fingers

  The item score is derived on the basis of "total incorrect" recorded on the eCRF and determined as follows:

  0 = 0–2 incorrect

  1 = 3–5 incorrect

  2 = 6–8 incorrect

  3 = 9–11 incorrect

  4 = 12–14 incorrect

  5 = 15–17 incorrect

- Item 6: ideational praxis

  The item score is equal to the "total incorrect" recorded on the eCRF for this item. The score has a range of 0–5.

- Item 7: orientation

  The item score is equal to the “total incorrect” recorded on eCRF for this item. The score has a range of 0–8.

- Item 8: word recognition

  The item score is equal to the “total incorrect” recorded on eCRF if “total incorrect” is ≤ 12. If the “total incorrect” is > 12, the item score is 12. The item score has a range of 0–12.

- Item 9: remembering test instructions

  The item score is derived on the basis of replies recorded on the eCRF and determined as follows: 0 = none, 1 = very mild, 2 = mild, 3 = moderate, 4 = moderately severe, and 5 = severe.

- Item 10: comprehension

  The item score is derived on the basis of replies recorded on the eCRF and determined as follows: 0 = none, 1 = very mild, 2 = mild, 3 = moderate, 4 = moderately severe, and 5 = severe.

- Item 11: word-finding difficulty

  The item score is derived on the basis of replies recorded on the eCRF and determined as follows: 0 = none, 1 = very mild, 2 = mild, 3 = moderate, 4 = moderately severe, and 5 = severe.

- Item 12: spoken language ability

  The item score is derived on the basis of replies recorded on the eCRF and determined as follows: 0 = none, 1 = very mild, 2 = mild, 3 = moderate, 4 = moderately severe, and 5 = severe.

- Item 13: number cancellation

  The item score has a range of 0–5 with a maximum score of 5, given if the item is not performed for cognitive reasons. The item score is derived from the cancellation score, which is equal to the number of target hits (minus) the number of errors (minus) the number of times reminded of task, and is determined as shown below:

  0 = cancellation score is > 23

  1 = cancellation score is 18–22

  2 = cancellation score is 13–17

  3 = cancellation score is 9–12

  4 = cancellation score is 5–8

  5 = cancellation score is 0–4

  This 45-second version of scoring rule is the scoring rule that Alzheimer Disease Neuroimaging Initiative is using and is based on the proposal of the ADAS Instrument Committee.