# CENTER FOR DRUG EVALUATION AND RESEARCH

***APPLICATION NUMBER:***

# 761269Orig1s000

# STATISTICAL REVIEW(S)

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

# STATISTICAL REVIEW AND EVALUATION

## CLINICAL STUDIES

**NDA/BLA #:** 761269 (0000)

**Drug Name:** BAN2401 Lecanemab
**Indication(s):** Alzheimer's
**Applicant:** Eisai
**Date(s):** September 27, 2021

**Review Priority:**

**Biometrics Division:** I
**Statistical Reviewer:** Tristan Massie, Ph.D.
**Concurring Reviewers:** Kun Jin, Ph.D., Team Leader
James (Hsien-Ming) Hung, Ph.D., Division Director
**Medical Division:** Division of Neurology I
**Clinical Team:** Kevin Krudys, Ph.D.
Ranjit Mani, M.D.
**Project Manager:** Andrew Papanastasiou

**Keywords:**
Biomarker Accelerated Approval

Table of Contents

# LIST OF TABLES

# LIST OF FIGURES

APPEARS THIS WAY ON ORIGINAL

# 1 EXECUTIVE SUMMARY

The applicant submitted Study 201 to evaluate lecanemab for treatment of Alzheimer's disease. In this phase 2 study, the applicant's primary endpoint was an Alzheimer's disease composite, ADCOMS. Study 201 had a Bayesian adaptive randomization design with multiple interim analyses. Randomization ratios were adapted to focus on the best performing dose based on the interim results.

Randomization of APOE4 carriers to the 10 mg/kg bi-weekly dose group, the best performing dose, was stopped in the middle of the study at the request of European regulators due to increased risk of Amyloid Related Imaging Abnormalities (ARIA) for carriers at this dose. This led to a large imbalance in APOE4 carrier status between the lecanemab 10 mg/kg bi-weekly and placebo groups (70% and 29% APOE4 non-carriers for 10 mg/kg bi-weekly and placebo, respectively).

As prespecified by the applicant, the trial was to be considered a success if either 1) early success was declared at an interim analysis or 2) the trial continued to completion and was declared a success at the 12 months Bayesian analysis. If the trial continued to completion, the trial was to be considered a success if there was at least an 80% probability that the best performing dose achieved a clinical threshold of a 25% difference from placebo on ADCOMS change from Baseline at Month 12.

The trial did not meet the success criteria at the interim analyses or at the final analysis. The final posterior probability indicates that there is a 58.5% probability of achieving a difference between placebo and lecanemab greater than or equal to the clinical threshold. This posterior probability does not meet the prespecified success probability of 80%.

The applicant specified several key secondary endpoints and several other secondary endpoints. Due to the failure of the primary endpoint and the large number of secondary endpoints, all secondary endpoints should be considered exploratory. Among the key secondary endpoints, the applicant focused on the effect of lecanemab compared to placebo on brain amyloid pathophysiology at 18 months of treatment in subjects with Early Alzheimer's Disease as measured by amyloid positron emission tomography (PET). Eight hundred and fifty-four patients were randomized in Study 201, 315 (36.8%) of whom participated voluntarily in the PET sub-study. The mixed model for repeated measures analyses of change from Baseline in brain amyloid levels as measured by PET SUVr normalized to whole cerebellum mask at 18 months estimated a reduction in percent change of 21.5 for 10 mg/kg biweekly compared to placebo (nominal p-value of <0.001).

The impact of amyloid reduction on the clinical outcome is uncertain. The Spearman correlation at the patient level is 0.128 between change in ADCOMS and change in Amyloid PET SUVR at Month 18. This correlation is small. Furthermore, there is no apparent treatment effect on the

clinical endpoint in APOE4 non-carriers despite the comparable amyloid reductions to APOE4 carriers (see Figure 6 and Table 14).

# 2 INTRODUCTION

## 2.1 Overview

The IND for this drug development of lecanemab, also referred to as BAN2401 during development, is IND 105081. There is one completed phase 2 study, study 201. This study had a Bayesian response adaptive randomization (see details in section 3.2.1.1). The primary endpoint was the ADCOMS (Alzheimer's Disease Composite) at 12 months. The effect(s) on Amyloid imaging was a secondary endpoint.

While the primary Bayesian analysis at 12 months on ADCOMS did not meet the primary outcome, additional conventional analyses are outlined below, and nominal p-values are included to provide context around the observed results. These analyses do not account for multiple comparisons.

**Table 1. Double Blind Phase 2 Placebo Controlled Study Characteristics**

| Study Name | Phase and Design | Treatment Period | Follow-up Period | # of Subjects per Arm | Study Population |
|---|---|---|---|---|---|
| 201 | Bayesian Adaptive Randomization Design. Randomization ratio could be changed at numerous interim analyses. | 18 months | 18 months | ITT/PET<br>Pl 245/ 99<br>2.5 bw 52/ 28<br>5 mth 51 /28<br>5 bw 92/ 27<br>10mth 253/ 89<br>10 bw 161 /44 | Alzheimer's |

Note: bw=bi-weekly and mth=monthly

Note that a phase 3 Placebo-Controlled, Double-Blind, Parallel-Group, 18-Month study, BAN2401-G000-301, is ongoing.

The primary objective of this study (301) is:

☒ To evaluate the efficacy of BAN2401 in subjects with Early Alzheimer's Disease (EAD) by determining the superiority of BAN2401 compared with placebo on the change from baseline in the CDR-SB at 18 months of treatment

Key Secondary Objectives of study 301 are as follows.

☒ To determine that BAN2401 is superior to placebo in reducing brain amyloid levels as measured by amyloid PET at 18 months of treatment in subjects with EAD

☒ To evaluate the efficacy of BAN2401 in subjects with EAD by determining the superiority of BAN2401 compared with placebo on the change from baseline in the ADCOMS at 18 months of treatment

☒ To evaluate the efficacy of BAN2401 in subjects with EAD by determining the superiority of BAN2401 compared with placebo on the change from baseline in the ADAS-cog14 at 18 months of treatment

The ongoing BAN2401-G000-301 (Study 301) is an 18 month treatment (Core Study), multicenter, double-blind, placebo controlled, parallel-group study in subjects with EAD (MCI due to AD with intermediate likelihood/Prodromal AD or mild AD dementia) with confirmed amyloid pathology indicated by either positive amyloid load confirmed by amyloid PET assessment or CSF assessment of t-tau/ $A\beta$ [1-42]. A total of 1,566 subjects will be randomized in the Core Study across 2 treatment groups, (placebo and BAN2401 10 mg/kg, biweekly) according to a fixed 1:1 (placebo:BAN2401) schedule. Randomization will occur across 2

clinical subgroups (MCI due to AD/prodromal AD or mild AD dementia), and will be reasonably balanced, such that not less than approximately 50% of total number of subjects will be in the MCI due to AD clinical subgroup. Subjects will be stratified according to clinical subgroup; presence or absence of ongoing approved AD treatment (eg, AChEIs, memantine, or both); ApoE4 status (ie, ApoE4 carriers or non-carriers); and geographical region. Treatment in the Core Study will be for 18 months. Clinical assessments (MMSE, CDRSB, ADAS-cog14) will be conducted every 3 months. At the end of the Core Study, subjects will have the option of entering the Extension Phase. An Extension Phase will be available for subjects who complete the full 18 months of treatment in the Core Study and will continue for up to 2 years, or until commercial availability of BAN2401, or until a positive risk-benefit assessment in this indication is not demonstrated. Longitudinal amyloid PET will be conducted at 3, 6, 12, and 18 months of treatment in a subgroup of consenting subjects (imaging substudy) from any participating country to demonstrate target engagement and to assess amyloid clearance for BAN2401. Longitudinal CSF assessments will be collected at 12 and 18 months of treatment for soluble biomarker analysis (eg, A$\beta$ [1-42], neurogranin, NFL, t-tau, and p-tau) in consenting subjects in any participating country to assess effects on indicators of disease pathology.

### 2.2 Data Sources

The amyloid imaging data set for the completed phase 2 study, 201, is located as follows: \\CDSESUB1\evsprod\BLA761269\0002\m5\datasets\ban2401-g000-201\analysis\adam\datasets\adca.xpt. The primary endpoint \\CDSESUB1\evsprod\BLA761269\0002\m5\datasets\ban2401-g000-201\analysis\adam\datasets\adcss.xpt and the link to the analysis data set for CDR-SB is \\CDSESUB1\evsprod\BLA761269\0002\m5\datasets\ban2401-g000-201\analysis\adam\datasets\adcdr.xpt.

## 3 STATISTICAL EVALUATION

### 3.1 Data and Analysis Quality

The submitted data and analysis quality appear adequate.

### 3.2 Evaluation of Efficacy

**BAN2401-G000-201**
**Study Protocol Title: A Placebo-Controlled, Double–Blind, Parallel-Group, Bayesian Adaptive Randomization Design and Dose Regimen–Finding Study, with an Open-Label Extension Phase, to Evaluate Safety, Tolerability and Efficacy of BAN2401 in Subjects With Early Alzheimer's Disease(EAD)**

**Start Date: 20 Dec 2012 (date of first subject's signed informed consent)**
**End Date: 19 Jul 2018 (21-month data; Core Study)**

## 3.2.1.1 Study Design and Endpoints

### Bayesian Response Adaptive Randomization (RAR) method of Enrolling Subjects

There was an initial burn-in period of 196 subjects with 2:1:1:1:1:1 randomization ratio to placebo control and each of the 5 active dose arms (BAN2401 2.5 mg/kg biweekly, 5 mg/kg monthly, 5 mg/kg biweekly, 10 mg/kg monthly, and 10 mg/kg biweekly), respectively. After this initial burn-in, RAR began. Adaptive randomization probabilities were updated when 196, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, and 750 subjects were enrolled. The randomization probability for each of the 5 active doses was weighted according to the variance components. The randomization probability to the control was designed to match the randomization probability to the most likely ED90 dose, i.e., dose regimen with at least 90% of the maximum effect. The RAR allowed more subjects to be allocated to the most likely ED90 dose and placebo control and therefore to increase the study power. The RAR allocated randomization into the dose groups likely to demonstrate efficacy (10 mg/kg monthly and biweekly) as early as the 350th subject randomized, and these doses remained the most likely doses to demonstrate efficacy throughout the remainder of the study.

However, before the interim analysis of 350 subjects, European Health Authorities (through the VHP) restricted randomization around ApoE4 carrier status, which meant subjects confirmed as ApoE4 carriers (hetero- or homozygous) were not to be randomized to the 10 mg/kg biweekly dose going forward. As a consequence, the RAR algorithm was revised. After each interim analysis (starting with 350 subjects enrolled), the randomization probability vector was split between ApoE4 carrier and non-carrier strata to ensure no ApoE4 carriers were enrolled on the 10 mg/kg biweekly dose. At the same time, the revised RAR would also preserve the overall randomization probabilities.

Primary efficacy of BAN2401 was assessed, by comparing to placebo, the change from Baseline at 12 months on ADCOMS, a composite clinical score representing a new approach to the analysis of selected items (12 total) from 3 fully validated and well-established clinical tools, including CDR (all 6 items), ADAS–Cog14 (4 items), and MMSE (2 items).

## 3.2.1.2 Statistical Methodologies

### Study 201 Statistical Analysis Plan

### Primary Objectives

1. To evaluate the efficacy of BAN2401 compared to placebo by establishing the ED90 (as defined in the protocol) for BAN2401 on the derived ADCOMS at 12 months of

treatment in subjects with Early Alzheimer's Disease (EAD), defined as mild cognitive impairment (MCI) due to Alzheimer's disease (AD) – intermediate likelihood or mild Alzheimer's disease dementia

2. To assess the safety and tolerability of 3 doses and 2 dose regimens of BAN2401 in subjects with EAD

**Key Secondary Objectives**

1. To evaluate the effects of BAN2401 compared to placebo on brain amyloid pathophysiology at 18 months of treatment in subjects with EAD as measured by amyloid positron emission tomography (PET)

2. To evaluate the efficacy of BAN2401 compared to placebo on the ADCOMS at 18 months of treatment in subjects with EAD

3. To evaluate the efficacy of BAN2401 compared to placebo on the Clinical Dementia Rating – Sum of Boxes (CDR-SB) at 18 months of treatment in subjects with EAD

4. To evaluate the efficacy of BAN2401 compared to placebo on Alzheimer's Disease Assessment Scale – Cognitive Subscale (ADAS-Cog) at 18 months in subjects with EAD

5. To evaluate the effects of BAN2401 compared to placebo at 18 months on clinical status separately within subjects with MCI and mild AD dementia for the following assessments: ADCOMS, CDR-SB, and ADAS-Cog

6. To evaluate the effects of BAN2401 compared to placebo on cerebrospinal fluid (CSF) biomarkers (Aβ[1-42], t-tau, and p-tau) at 18 months of treatment in subjects with EAD

7. To evaluate the effects of BAN2401 compared to placebo on total hippocampal volume using volumetric magnetic resonance imaging (vMRI) at 18 months of treatment in subjects with EAD

**Secondary Objectives**

1. To evaluate the effects of BAN2401 compared to placebo on brain amyloid pathophysiology at 12 months of treatment in subjects with EAD as measured by amyloid PET

2. To evaluate the effects of BAN2401 compared to placebo at 12 months on clinical status in subjects with EAD for the following assessments: ADCOMS, CDR-SB, and ADAS-Cog

3. To evaluate the effects of BAN2401 compared to placebo at 12 months on clinical status separately within subjects with MCI and mild AD dementia for the following assessments: ADCOMS, CDR-SB, and ADAS-Cog

4. To evaluate the effects of BAN2401 compared to placebo on cerebrospinal fluid (CSF) biomarkers (Aβ[1-42], t-tau, and p-tau) at 12 months of treatment in subjects with EAD

5. To evaluate the effects of BAN2401 compared to placebo on total hippocampal atrophy as measured by vMRI at 6 and 12 months, and left and right hippocampus, whole brain and total ventricular volume as measured by vMRI at 6, 12, and 18 months of treatment in subjects with EAD

## Endpoints and Analyses

### vMRI

The null hypothesis is that there is no difference between active dose and placebo. There are 5 null hypotheses corresponding to 5 active dose regimens. The alternative hypothesis is that at least 1 null hypothesis is false (1-sided). The null hypotheses were to be tested using Dunnett-Hsu method with 1-sided alpha of 0.05. The statistical power was estimated through simulation for a moderate dose-response assumption that the percent reduction in change from baseline of total hippocampal volume compared to placebo would be 15%, 20%, 25%, 15%, and 20% corresponding to the 5 dose regimens (2.5, 5, and 10 mg/kg biweekly, 5 and 10 mg/kg 4-week interval), respectively. The estimated power for actual planned study sample size is 76.4% at 12 months and 69.4% at 18 months, assuming an attrition rate of 20% at 12 months and an exponential dropout model. Under a stronger dose-response assumption (i.e, the percent reduction in change from baseline of total hippocampal volume compared to placebo is twice as much as that in the moderate dose-response assumption), the estimated power is at least 99% at both 12 and 18 months.

### AMYLOID PET (IMAGING SUBGROUP)

The null and alternative hypotheses and statistical test are the same as that for vMRI. To estimate the sample size for amyloid PET imaging subgroup at 18 months, a common standard deviation of change from baseline was estimated as 0.4 and a reasonable mean difference between treatment and placebo was estimated as 0.25 (SUVR). The corresponding standard deviation and mean difference at 12 months were estimated as 0.27 and 0.17, respectively. Assuming a moderate dose-response that the best dose regimen would achieve a difference of 0.25 (SUVR), and that the 2 middle and 2 low dose regimens would achieve a difference of 0.2 and 0.15 (SUVR), respectively, it would require a total of 230 subjects from all 6 arms at 18 months to achieve an 80% power. This sample size took into account possible unequal numbers of subjects per arm. Under the same attrition rate assumption as that for vMRI, a total sample size of 306 is required at 12 months in order to target 230 subjects at 18 months. The estimated power with 306 subjects at 12 months is approximately 85%. Because amyloid PET imaging is required for every subject at baseline, the post baseline amyloid PET in the imaging subgroup was planned for the first 306 subjects who consented and were still on treatment at 12 months.

**Primary Endpoint**

The maximum ADCOMS derived composite score is achieved when each item is assigned the maximum score. This maximum composite score is 1.97. The range of this new composite score is therefore between 0 and 1.97.

The primary endpoint is the change from baseline to 12 months in the derived ADCOMS. The primary efficacy analysis is based on Bayesian statistics. The dose-response of the primary endpoint is modeled with a two-dimensional first-order normal dynamic linear model (NDLM), where Normal and Inverse-Gamma priors are used. The primary efficacy analysis was to calculate the Bayesian posterior probability that the dose identified is the most likely ED90 dose that achieves the clinically significant difference (CSD) compared to the placebo arm. At each interim analysis and the final analysis, three Bayesian probabilities were to be summarized for each active dose: the probability of being the ED90 dose, the probability of being statistically superior to placebo, and the probability of being statistically superior to placebo by the CSD. The study was to be considered a success in the final analysis (full randomization and all subjects reach 12 months of treatment) if there is at least an 80% probability that the ED90 achieves the clinically significant difference from placebo.

**Censoring**

For the clinical efficacy data (i.e. derived ADCOMS, MMSE, CDR-SB, ADAS-cog and FAQ)

 Subjects were to be censored at the time of initiation of new AChEIs or memantine treatment regimens if they were not on AChEIs or memantine at randomization, and were to be censored at the time of dose adjustment of AChEIs or memantine if they were already on stable treatment with AChEIs or memantine at randomization. The value of the primary endpoint for censored subjects and censored assessments was to be imputed using data up to the censoring time and the Bayesian imputation method.

## Interim Analyses

An unblinded independent Interim Monitoring Committee (IMC) was to provide oversight to ensure that the response adaptive randomization process and interim analyses perform as expected. An independent data analysis group was to perform all of the interim analyses and was to provide the results to the IMC.

The first interim analysis was to be conducted when the first 196 subjects had been randomized, again when 250 subjects were randomized, and again after each additional 50 subjects until 800 subjects were randomized. If the study reaches 800 randomized subjects, only these initial 800 subjects randomized were to contribute to additional interim analyses that were to be conducted every 3 months for a total of 9 months (i.e., 3 additional interim analyses). Each interim analysis was to be conducted on clean data for the derived ADCOMS. The Bayesian analysis at 12 months and 18 months were to be performed when all of the enrolled subjects (i.e., including any

subjects who randomized after the initial 800) had been on study treatment for 12 months and 18 months, respectively, and their clinical efficacy measures evaluated.

**Stopping Rules For Efficacy (Early Success)**

The trial was to be considered a success if either 1) early success is declared at an interim analysis or 2) the trial continues to completion and is declared a success at the 12 months Bayesian analysis. Since adaptive randomization decisions were to be based on the 12 month treatment endpoint, a decision for early success was to require that some subjects complete 12 months of treatment, and the early success decision could not be made prior to the study randomizing 350 subjects. Interim monitoring for early success was to begin at the 350 subject interim analysis. If there is greater than a 95% probability that the ED90 achieves a clinically significant difference from placebo during the accrual period, the trial was to stop randomization and was to be declared an early success. If the trial continues to completion, the trial was to be considered a success, if there is at least an 80% probability that the ED90 achieves the clinically significant difference from placebo.

**Stopping Rules For Futility**

The study meets statistical futility criteria if the probability that the ED90 is better than the placebo by the CSD is less than 0.05 (with ≤300 subjects randomized) or 0.075 (with ≥350 subjects randomized) at 12 months of treatment. If the study meets statistical futility criteria, the applicant was to make the final decision pertaining to study futility after reviewing recommendations by the DSMB. Based on Applicant's final decision, if the study continues to randomize subjects, it was to be considered a failed study.

**Conventional Analyses**

Statistical methods for the final conventional analyses were to use a mixed effects model with repeated measures on the change from baseline in the derived ADCOMS at 12 months. The model was to include baseline ADCOMS as covariate, with treatment group, visit, randomization stratification variables (i.e., clinical subgroup [MCI due to AD, Mild Alzheimer's Disease Dementia], the presence or absence of ongoing AD treatment [ AChEIs or memantine or both], *APOE4* status [positive, negative]), region, treatment group-by-visit interaction as fixed effects. An unstructured covariance matrix was to be employed to model the covariance of within subject effects. These analyses were to be performed on the full analysis set as well as the per protocol set. The analysis was only to be performed at early success or at the end of 12 months assessment using all randomized subjects in the Full Analysis Set.

There was to be no conventional analysis at each interim analysis before early success is claimed and as such the applicant felt that there was to be no multiplicity adjustment needed to account for the interim analyses.

In case of early success, the conventional analyses were to be performed using the data collected prior to the interim analysis as well as at the end of 12 months assessment, and the analysis at

the end of 12 months assessment was to be considered as sensitivity analysis to the analysis at early success.

*Reviewer's Comment:*

*The reviewer does not agree that no multiplicity adjustment is needed since Bayesian stopping criteria without some kind of multiplicity adjustment may inflate frequentist type I error as the applicant's simulations showed.*

**Determination of Sample Size**

The sample size and design characteristics to test the hypothesis under the primary objective based on the primary endpoint, the ADCOMS at 12 months, were determined by means of simulations. Extensive trial simulations have shown that a total of 800 subjects will be sufficient to demonstrate that the most likely ED90 dose achieves a clinically significant difference (CSD) from placebo (of .03 on ADCOMS) with a probability of at least 95% in the interim analyses and greater than 80% in the analysis at 12 months if the trial does not stop for early success. With 800 subjects, the average probability of study success is 80% across a wide range of dose response scenarios, and there is at least 80% probability of study success for the dose response scenarios where the treatment is considered to show a clinically significant difference from placebo as defined in the Protocol. Simulations have also shown that if there is no efficacy at all for any dose, then the probability of falsely claiming superiority to placebo is no more than 10% assuming a 20% dropout rate.
For each of the 6 treatment groups (5 active dose regimens and placebo) the final number of subjects per group will differ depending on the observed interim treatment responses.
The simulation plan is described in the appendix section of this document which presents further details.

*Reviewer's Comment:*

*During the IND stage the applicant acknowledged that the simulated type I error was 10-12% one-sided, exceeding the usual 0.025 one-sided level due to the Bayesian adaptive dose selection and many interim analyses without any multiplicity corrections. Due to this limitation the statistical reviewers stated that the study was to be considered exploratory.*

**Key Secondary Endpoint Analyses**

Conventional analysis was to be performed using all available data at the time of success (if applicable), and at 18 months of treatment, regardless of success. Conventional analysis was to be based on the MMRM model including baseline as covariate, with treatment group, visit, randomization stratification variables (i.e., clinical subgroup [MCI due to AD, Mild Alzheimer's Disease Dementia], the presence or absence of ongoing AD treatment [AChEIs or memantine or both], *APOE4* status [positive, negative]), region, treatment group-by-visit interaction as fixed effects. If there is an imbalance across treatment groups based on randomization stratification variables, an additional MMRM model was to be run without the randomization stratification variables. In addition, an ANCOVA model was also to be used with treatment as a factor and baseline value as a covariate. These statistical models were to be used to compare the combined 10 mg dose group (including bi-weekly and monthly regimens) with placebo for the key secondary endpoints. The rationale for combining the two 10 mg dose regimens was to account for the loss of subjects positive for APOE4 in the 10 mg/kg biweekly dose group and the inability to randomize APOE4 carriers to the 10 mg/kg biweekly group following a regulatory request by European Health Authorities in July 2014.

In addition, analysis was to be performed for the following treatment comparisons:
 Combining 2 high doses (10 mg bi-weekly + 10 mg monthly), 2 middle doses (5 mg biweekly

+ 5 mg monthly), and the 2.5 mg bi-weekly dose, resulting in 3 treatment comparisons with placebo (for dose response)
- ED90 dose of BAN2401 versus placebo, where the ED90 dose of BAN2401 was to be established by Bayesian analysis at 18 months as described in the next paragraph
- By treatment regimen (10 mg bi-weekly, 10 mg monthly, 5 mg bi-weekly, 5 mg monthly, 2.5 mg bi-weekly, placebo)
- Combining the top 3 doses (10 mg bi-weekly, 10 mg monthly, and 5 mg bi-weekly) versus placebo

The key Secondary endpoints (ADCOMS, CDR-SB, and ADAS-Cog) were also to be analyzed separately within subjects with MCI and mild AD dementia. Change from baseline in the ADCOMS at 18 months was to be analyzed, as a sensitivity analysis to the conventional analysis, using the same Bayesian methodology as that for analysis of change from baseline in the ADCOMS at 12 months, but using the full 18 months of efficacy data with the model projecting to 18 months of treatment. The Bayesian analysis of change from baseline in the ADCOMS at 18 months was to be positive if the analysis results in an ED90 dose with at least an 80% probability of being better than the placebo by the CSD. The ED90 dose was to be identified based on change from baseline in the ADCOMS at the 18 month Bayesian analysis and was to be used to compare ED90 dose of BAN2401 and placebo in the above sensitivity analyses for all key secondary endpoint analyses.

Specifically, the following Bayesian analyses were to be performed at 18 months:

- Fit the 2-dimensional NDLM Bayesian model defined in Appendix with 5 doses across 2 schedules of administration for the 18-month endpoint. This includes imputation of missing values from the longitudinal model.
- Compare placebo versus the combined 10mg dose group. This includes imputation of missing values from the longitudinal model.
- Fit a one-dimensional Bayesian model with 3 doses: 2.5mg dose group, combined 10mg dose group, and combined 5mg dose group. This was to be a one-dimensional NDLM model with imputation of missing values from the longitudinal model.
- Fit the European Prevention of Alzheimer's Dementia (EPAD) disease progression model for selected doses versus the placebo control.

In this study, subjects consented to PET sub-study have been using two different tracers for PET scan: Florbetapir and Flutemetamol. Only less than 10 subjects have used Flutemetamol and the rest have used Florbetapir. It is possible that different tracers may yield different results and data using different tracers may not be directly comparable, therefore only subjects using Florbetapir were to be included in summary tables and other statistical analysis based on PET data. Subjects using Flutemetamol were only to be listed and were not to be included in summary tables or other statistical analysis based on PET data.

**Other Secondary Endpoint Analyses**

The same analysis described above for key secondary endpoints was to be performed for other secondary endpoints.

For vMRI and amyloid PET endpoints, the adjusted p-value based on Dunnett-Hsu method with 1-sided alpha of 0.05 was also to be provided in addition to the p-value corresponding to pairwise comparison.

**Exploratory Analyses for Biomarkers**

Scatter plots including the Spearman correlation coefficient were to be produced for change from baseline in vMRI values (total hippocampal volume, right and left hippocampal, whole brain, and total ventricular volumes) and change from baseline in amyloid PET values (SUVR of global cortical average with reference region of subcortical white matter, whole cerebellum, whole cerebellum mask, whole cerebellum adjusted by subcortical white matter, cerebellar grey matter, and composite reference consisting of cerebellar cortex, pons, subcortical white matter and cerebellar white matter) at 12 and 18 months.

Change from baseline at 12 and 18 months in PET SUVR, volumetric MRI and CSF biomarkers was also to be analyzed by baseline PET SUVr subgroups (≥ median and < median). The median here is the median baseline value among all subjects.

Change from baseline in brain amyloid levels as measured by amyloid PET SUVR of global cortical average with other reference regions at 12 and 18 months was also to be explored. Some other PET data parameters may also have been analyzed. The amyloid PET SUVR of global cortical average with whole cerebellum as reference region and subcortical white matter as longitudinal adjustment factor can be derived as follows (Landau et al, 2015):

1) For each visit, the global cortical average (*Ctx*) is first normalized to the whole cerebellum (*WhCereb*) at the same visit:
Baseline:
$Ctx^*_{BL} = Ctx_{BL} / WhCereb_{BL}$

At post baseline visit $v_i$:

$Ctx^*_{V_i} = Ctx_{V_i} / WhCereb_{V_i}$

2) For each visit, calculate the ratio below for subcortical white matter (*WM*):
At post baseline visit $v_i$ :
$WhCereb^*_{V_i} = WhCereb_{V_i} / WM_{V_i}$

3) The normalized global cortical average in 1) is then scaled by the ratio of subcortical white matter in 2).
Baseline: $SUVR_{BL} = Ctx^*_{BL} / (WM^*_{BL} / WM^*_{BL})$
At post baseline visit $v_i$ : $^* SUVR_{V_i} = Ctx^*_{V_i} / (WM^*_{V_i} / WM^*_{BL})$

4) Finally, the change from baseline at post baseline visit $v_i$ is:
$\Delta SUVR = SUVR_{V_i} - SUVR_{BL}$
After some simple algebra, it can be shown that the $SUVR_{V_i}$ in 3) can be written as follows:

Baseline:
$SUVR_{BL} = Ctx_{BL} / WM_{BLi}$ $(WM_{BL}/WhCereb_{BL})$
At post baseline visit $v_i$ :

$SUVR_{Vi} = Ctx_{Vi} / WM_{Vi}$ $(WM_{BL}/WhCereb_{BL})$

Therefore, this SUVR is essentially the SUVR of global cortical average with subcortical white matter as reference region multiplied by the ratio of subcortical white matter and whole cerebellum at baseline.

**Exploratory Analyses to Investigate the Relationship between Clinical Endpoints and Biomarkers**

Clinical endpoints to be investigated include ADCOMS and CDR-SB. Biomarkers to be investigated include CSF (Aβ1-42, t-tau and p-tau), vMRI (total hippocampal volume, right and left hippocampal, whole brain, and total ventricular volumes), and PET values (SUVR of global cortical average versus the various reference regions).

The relationship between baseline biomarker values and change from baseline in clinical endpoints at 12 or 18 months were to be investigated as follows:

 An MMRM will be fitted with dependent variable of change from baseline in clinical endpoints at 12 months and covariates of baseline clinical endpoints, treatment group, visit, randomization stratification variables (i.e., clinical subgroup [MCI due to AD, Mild Alzheimer's Disease Dementia], the presence or absence of ongoing AD treatment [AChEIs or memantine or both], APOE ε4 carrier status [positive, negative]), region, treatment group-by-visit interaction, and baseline CSF (as continuous variables). The model was to be repeated using baseline vMRI and baseline amyloid PET values to replace baseline CSF, respectively.
 The above analyses were also to be performed for change from baseline in clinical endpoints at 18 months and baseline values of CSF, vMRI and amyloid PET values.
 The same analyses were to be done with baseline CSF, baseline vMRI and baseline amyloid PET values being categorized into two groups: ≥ median, < median. The median here is the median baseline value among all subjects.

The relationship between change from baseline in biomarker values and change from baseline in clinical endpoints at 12 or 18 months was to be investigated as follows:

 A scatter plot by treatment group was to be made between change from baseline in clinical endpoints at 12 months and change from baseline in each of the biomarkers at 12 months, respectively: CSF, vMRI and amyloid PET values. A Spearman correlation coefficient with its p-value was also to be displayed in each scatterplot.
 The same analyses was also to be performed for change from baseline in clinical endpoints, CSF, vMRI and amyloid PET values at 18 months.
 A longitudinal plot of mean change in clinical endpoints over time (baseline, 12 months, 18 months) by treatment group and subgroups defined by change from baseline in CSF, vMRI and PET values (≥ median, < median) was also to be produced. Summary statistics were also to be provided.

The relationship between change from baseline in clinical endpoints and change from baseline in biomarkers at 18 months may also be evaluated using an ANCOVA model within treatment group as appropriate. Changes from baseline in clinical endpoints at 18 months were to be the response variables and either continuous or categorical change of biomarkers were to be independent variables. The ANCOVA model was also to include baseline value of clinical endpoint as a covariate, randomization stratification variables as factors, and other terms as appropriate.

Proportion of subjects who have become amyloid negative during study was also to be summarized by treatment groups. Summary of clinical endpoints was to be presented for those subjects who remain amyloid positive and subjects who change from amyloid positive to amyloid negative by visual read at either 12 months or 18 months of treatment. Time to conversion from amyloid positive to amyloid negative may also be analyzed.

Additional analyses were also to be performed whereby PET SUVR values are converted to the Centiloid scale (Klunk, et al., 2015).

Only subjects that have data in both clinical endpoints and each individual biomarker at specified timepoints were to be included in these respective analyses.

## 3.2.1.3 Patient Disposition, Demographic and Baseline Characteristics

**Patient Disposition**

A higher percentage of subjects in the BAN2401 treatment groups (234 [38.4%] subjects) discontinued the study (i.e, discontinued the 18-month treatment period and the 3 month Follow-Up Period), than in the placebo group (68 [27.8%] subjects) [

**Table 2**]. In particular, this was associated with a higher percentage of subjects who discontinued from the study at higher doses of BAN2401 (98 [38.7%] and 74 [46.0%] subjects in the 10 mg/kg monthly and biweekly groups, respectively).

A greater percentage of subjects discontinued due to Adverse Events (AE) was observed at higher doses of BAN2401 (23 [9.1%] and 12 [7.5%] subjects in 10 mg/kg monthly and biweekly, respectively) as compared to placebo (10 [4.1%] subjects). The AE driving the observed higher discontinuation rate was ARIA–E (a total of 48 cases, 25 and 16 cases in 10 mg/kg monthly and biweekly, respectively). In addition, 25 subjects who were ApoE4 carriers and who were on 10 mg/kg biweekly for less than 6 months were discontinued in accordance with VHP request. This implementation of the VHP request was the driver for the observed higher discontinuation rate due to “Other” in ApoE4 carriers (21 [43.8%] subject), compared to ApoE4 non-carriers (10 [8.8%] subjects) on 10 mg/kg biweekly dose.

**Table 2. Overall Subject Disposition in Study 201 -Randomized Set**

| | | BAN2401 | | | | | |
|---|---|---|---|---|---|---|---|
| | Placebo | 2.5 mg/kg Biweekly | 5 mg/kg Monthly | 5 mg/kg Biweekly | 10 mg/kg Monthly | 10 mg/kg Biweekly | Total |
| **Overall** | | | | | | | |
| Randomized, n | 247 | 52 | 51 | 92 | 253 | 161 | 609 |
| Not treated, n | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| Treated, n | 245 | 52 | 51 | 92 | 253 | 161 | 609 |
| Completed the study, n (%)[a] | 177 (72.2) | 35 (67.3) | 37 (72.5) | 61 (66.3) | 155 (61.3) | 87 (54.0) | 375 (61.6) |
| Discontinued from the study, n (%)[a] | 68 (27.8) | 17 (32.7) | 14 (27.5) | 31 (33.7) | 98 (38.7) | 74 (46.0) | 234 (38.4) |
| Primary reason for discontinuation[b] | | | | | | | |
| Adverse event[c] | 10 (4.1) | 4 (7.7) | 2 (3.9) | 5 (5.4) | 23 (9.1) | 12 (7.5) | 46 (7.6) |
| Lost to follow-up | 7 (2.9) | 0 | 1 (2.0) | 2 (2.2) | 4 (1.6) | 3 (1.9) | 10 (1.6) |
| Subject choice | 15 (6.1) | 5 (9.6) | 2 (3.9) | 7 (7.6) | 14 (5.5) | 8 (5.0) | 36 (5.9) |
| Withdrawal of consent | 23 (9.4) | 1 (1.9) | 5 (9.8) | 13 (14.1) | 37 (14.6) | 20 (12.4) | 76 (12.5) |
| Other | 13 (5.3) | 7 (13.5) | 4 (7.8) | 4 (4.3) | 20 (7.9) | 31 (19.3) | 66 (10.8) |

Note: Table copied from page 143 of applicant's study report

Table 3 shows the numbers of patients for each randomized group within each of the analysis sets, e.g., PD analysis set 2 is the relevant analysis set for the PET analyses since participation in the PET study was voluntary.

**Table 3. Numbers of Subjects Randomized to Treatment**

| | | BAN2401 | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Analysis Sets | Placebo (N=247) n (%) | 2.5 mg/kg Biweekly (N=52) n (%) | 5 mg/kg Monthly (N=51) n (%) | 5 mg/kg Biweekly (N=92) n (%) | 10 mg/kg Monthly (N=253) n (%) | 10 mg/kg Biweekly (N=161) n (%) | Total (N=609) n (%) | Combined Total (N=856) n (%) |
| Safety Analysis Set[a] | 245 (99.2) | 52 (100.0) | 51 (100.0) | 92 (100.0) | 253 (100.0) | 161 (100.0) | 609 (100.0) | 854 (99.8) |
| Full Analysis Set[b] | 238 (96.4) | 52 (100.0) | 48 (94.1) | 89 (96.7) | 246 (97.2) | 152 (94.4) | 587 (96.4) | 825 (96.4) |
| PP Analysis Set[c] | 236 (95.5) | 50 (96.2) | 46 (90.2) | 89 (96.7) | 242 (95.7) | 152 (94.4) | 579 (95.1) | 815 (95.2) |
| PK Analysis Set[d] | 0 | 52 (100.0) | 51 (100.0) | 92 (100.0) | 251 (99.2) | 161 (100.0) | 607 (99.7) | 607 (70.9) |
| PD Analysis Set 1[e] | 209 (84.6) | 41 (78.8) | 46 (90.2) | 73 (79.3) | 188 (74.3) | 99 (61.5) | 447 (73.4) | 656 (76.6) |
| PD Analysis Set 2[f] | 99 (40.1) | 28 (53.8) | 28 (54.9) | 27 (29.3) | 89 (35.2) | 44 (27.3) | 216 (35.5) | 315 (36.8) |
| PD Analysis Set 3[g] | 24 (9.7) | 7 (13.5) | 13 (25.5) | 20 (21.7) | 16 (6.3) | 12 (7.5) | 68 (11.2) | 92 (10.7) |

Percentages are based on the number of randomized subjects in the relevant treatment group.
CSF = cerebrospinal fluid, PD = pharmacodynamics, PET = positron emission tomography, PK = pharmacokinetics, PP = Per Protocol, vMRI = volumetric magnetic resonance imaging.
a: The Safety Analysis Set is the group of subjects who received at least 1 dose of study drug and had at least 1 postdose safety assessment.
b: The Full Analysis Set is the group of randomized subjects who received at least 1 dose of study drug and had Baseline and at least 1 postdose primary efficacy measurement.
c: The PP Analysis Set is the subset of subjects in the Full Analysis Set who complied with the protocol.
d: The PK Analysis Set is the group of subjects with at least 1 quantifiable BAN2401 serum concentration with a documented dosing history.
e: The PD Analysis Set 1 is the group of subjects who had sufficient vMRI data to derive at least 1 vMRI parameter.
f: The PD Analysis Set 2 is the group of subjects who had sufficient amyloid PET data to derive at least 1 amyloid PET parameter.
g: The PD Analysis Set 3 is the group of subjects who had sufficient CSF data to derive at least 1 CSF parameter.

## Baseline Demographics

Demographic and Baseline characteristics were consistent with clinical studies in this population, and similar between placebo and BAN2401 treatment groups (Table 4). There was a similar proportion of male (416 [50.4%]) and female (409 [49.6%]) subjects in the Full

Analysis Set overall. However, the BAN2401 groups had more male subjects than the placebo (53.7% vs 42.4%, respectively). The median age in the Full Analysis Set was 72.0 (range: 50 to 90) years for placebo and BAN2401 overall. The majority of subjects were white (746 [90.4%] subjects).

The characteristics of primary disease diagnosis were similar across all treatment groups. The majority of subjects in the Full Analysis Set had a CDR-Global rating of 0.5 (704 [85.3%] subjects), and the remaining subjects had a rating of 1 (121 [14.7%] subjects). The median time since disease diagnosis was 2 (range: 1 to 12) years. The median age at diagnosis was 71 (range: 49 to 90) years. The median time since onset of symptoms was 4 (range: 1 to 17) years. The median age at onset of symptoms was 69 (range: 45 to 89) years. For the clinical diagnosis of EAD confirmed during the prerandomization phase of the study, 529 (64.1%) subjects had a diagnosis of MCI due to AD, and 296 (35.9%) subjects had mild AD dementia. In general, the demographic and Baseline characteristics of these 2 subgroups were similar to the overall population, except that a greater percentage of subjects with mild AD dementia were taking concomitant AD treatment (AChEIs and/or memantine) at Baseline (206 [69.6%] subjects) than subjects with MCI due to AD (241 [45.6%] subjects), and compared to the overall population (447 [54.2%] subjects taking concomitant AChEIs and/or memantine at Baseline).

Regarding ApoE4 genotype, the majority of subjects overall were ApoE4 carriers (589 [71.4%] subjects, of which 453 [54.9%] were heterozygous and 136 [16.5%] were homozygous), compared to ApoE4 non-carriers (236 [28.6%] subjects). However, ApoE4 genotype was imbalanced across treatment groups due to the randomization restrictions resulting from the VHP interaction that meant ApoE4 carriers could no longer be randomized to the 10 mg/kg biweekly dose and any ApoE4 carriers at this dose were discontinued if they had less than 6 months of exposure. Thus, in the BAN2401 10 mg/kg biweekly group there were fewer ApoE4 carriers (30.3% of subjects) as the RAR allocated most of the ApoE4 carriers to the next most efficacious groups, 10 mg/kg monthly (88.6% of subjects who were ApoE4 carriers). Demographic and other Baseline characteristics in the Safety Analysis Set were comparable to the Full Analysis Set.

**Table 4. Study 201 Baseline Demographics in Randomized population**

| Category | | Placebo (N=238) | BAN2401 2.5 mg/kg Biweekly (N=52) | BAN2401 5 mg/kg Monthly (N=48) | BAN2401 5 mg/kg Biweekly (N=89) | BAN2401 10 mg/kg Monthly (N=246) | BAN2401 10 mg/kg Biweekly (N=152) | BAN2401 Total (N=587) | Combined Total (N=825) |
|---|---|---|---|---|---|---|---|---|---|
| Age (year)[a] | n | 238 | 52 | 48 | 89 | 246 | 152 | 587 | 825 |
| | Mean (SD) | 71.11 (8.892) | 70.50 (8.257) | 70.42 (7.514) | 70.64 (7.446) | 71.26 (7.455) | 72.64 (8.777) | 71.39 (7.907) | 71.31 (8.198) |
| | Median | 72.00 | 70.50 | 71.00 | 72.00 | 71.00 | 73.00 | 72.00 | 72.00 |
| | Min, max | 50.0, 89.0 | 50.0, 86.0 | 55.0, 84.0 | 52.0, 87.0 | 53.0, 90.0 | 51.0, 88.0 | 50.0, 90.0 | 50.0, 90.0 |
| Age group, n (%) | <65 years | 55 (23.1) | 11 (21.2) | 9 (18.8) | 20 (22.5) | 44 (17.9) | 27 (17.8) | 111 (18.9) | 166 (20.1) |
| | ≥65 to <80 years | 144 (60.5) | 35 (67.3) | 35 (72.9) | 60 (67.4) | 168 (68.3) | 94 (61.8) | 392 (66.8) | 536 (65.0) |
| | ≥80 years | 39 (16.4) | 6 (11.5) | 4 (8.3) | 9 (10.1) | 34 (13.8) | 31 (20.4) | 84 (14.3) | 123 (14.9) |
| Sex, n (%) | Male | 101 (42.4) | 26 (50.0) | 24 (50.0) | 41 (46.1) | 136 (55.3) | 88 (57.9) | 315 (53.7) | 416 (50.4) |
| | Female | 137 (57.6) | 26 (50.0) | 24 (50.0) | 48 (53.9) | 110 (44.7) | 64 (42.1) | 272 (46.3) | 409 (49.6) |
| Ethnicity, n (%) | Hispanic or Latino | 9 (3.8) | 4 (7.7) | 1 (2.1) | 3 (3.4) | 9 (3.7) | 9 (5.9) | 26 (4.4) | 35 (4.2) |
| | Not Hispanic or Latino | 229 (96.2) | 48 (92.3) | 47 (97.9) | 86 (96.6) | 237 (96.3) | 143 (94.1) | 561 (95.6) | 790 (95.8) |
| Race, n (%) | White | 216 (90.8) | 48 (92.3) | 46 (95.8) | 7 (82.0) | 222 (90.2) | 141 (92.8) | 530 (90.3) | 746 (90.4) |
| | Black or African American | 5 (2.1) | 2 (3.8) | 1 (2.1) | 4 (4.5) | 4 (1.6) | 4 (2.6) | 15 (2.6) | 20 (2.4) |
| | Chinese | 1 (<1.0) | 0 | 0 | 0 | 0 | 0 | 0 | 1 (<1.0) |
| | Japanese | 10 (4.2) | 1 (1.9) | 0 | 6 (6.7) | 12 (4.9) | 5 (3.3) | 24 (4.1) | 34 (4.1) |
| | Other Asian | 5 (2.1) | 1 (1.9) | 1 (2.1) | 3 (3.4) | 5 (2.0) | 2 (1.3) | 12 (2.0) | 17 (2.1) |
| | Other | 1 (<1.0) | 0 | 0 | 3 (3.4) | 3 (1.2) | 0 | 6 (1.0) | 7 (<1.0) |
| Region, n (%) | North America | 195 (81.9) | 47 (90.4) | 41 (85.4) | 70 (78.7) | 215 (87.4) | 135 (88.8) | 508 (86.5) | 703 (85.2) |
| | Western Europe | 28 (11.8) | 4 (7.7) | 6 (12.5) | 7 (7.9) | 15 (6.1) | 10 (6.6) | 42 (7.2) | 70 (8.5) |
| | Asia | 15 (6.3) | 1 (1.9) | 1 (2.1) | 12 (13.5) | 16 (6.5) | 7 (4.6) | 37 (6.3) | 52 (6.3) |

| Category | | Placebo (N=238) | BAN2401 2.5 mg/kg Biweekly (N=52) | BAN2401 5 mg/kg Monthly (N=48) | BAN2401 5 mg/kg Biweekly (N=89) | BAN2401 10 mg/kg Monthly (N=246) | BAN2401 10 mg/kg Biweekly (N=152) | BAN2401 Total (N=587) | Combined Total (N=825) |
|---|---|---|---|---|---|---|---|---|---|
| Height (cm) | n | 238 | 52 | 48 | 89 | 246 | 152 | 587 | 825 |
| | Mean (SD) | 166.92 (10.479) | 169.15 (10.963) | 167.53 (9.412) | 166.67 (8.519) | 168.27 (8.772) | 168.56 (9.804) | 168.12 (9.269) | 167.77 (9.642) |
| | Median | 165.10 | 169.00 | 167.50 | 165.10 | 168.55 | 167.95 | 167.90 | 167.60 |
| | Min, max | 142.2, 197.0 | 149.9, 195.0 | 149.9, 187.0 | 147.3, 188.5 | 144.7, 188.0 | 144.8, 188.5 | 144.7, 195.0 | 142.2, 197.0 |
| Weight (kg) | n | 238 | 52 | 48 | 89 | 246 | 152 | 587 | 825 |
| | Mean (SD) | 71.39 (13.905) | 75.03 (17.214) | 73.22 (16.319) | 68.70 (15.302) | 74.38 (14.476) | 75.01 (14.177) | 73.64 (15.044) | 72.99 (14.752) |
| | Median | 70.40 | 75.75 | 72.45 | 67.00 | 73.55 | 74.45 | 73.30 | 72.70 |
| | Min, max | 29.2, 108.0 | 36.3, 124.7 | 36.7, 118.7 | 35.9, 110.0 | 36.7, 117.7 | 47.3, 123.0 | 35.9, 124.7 | 29.2, 124.7 |
| CDR-Global, n (%) | 0.5 | 200 (84.0) | 44 (84.6) | 40 (83.3) | 77 (86.5) | 210 (85.4) | 133 (87.5) | 504 (85.9) | 704 (85.3) |
| | 1 | 38 (16.0) | 8 (15.4) | 8 (16.7) | 12 (13.5) | 13 (14.6) | 19 (12.5) | 83 (14.1) | 121 (14.7) |
| ApoE4 carrier status, n (%) | Carrier | 169 (71.0) | 38 (73.1) | 37 (77.1) | 81 (91.0) | 218 (88.6) | 46 (30.3) | 420 (71.6) | 589 (71.4) |
| | Heterozygous | 129 (54.2) | 33 (63.5) | 26 (54.2) | 67 (75.3) | 160 (65.0) | 38 (25.0) | 324 (55.2) | 453 (54.9) |
| | Homozygous | 40 (16.8) | 5 (9.6) | 11 (22.9) | 14 (15.7) | 58 (23.6) | 8 (5.3) | 96 (16.4) | 136 (16.5) |
| | Non-carrier | 69 (29.0) | 14 (26.9) | 11 (22.9) | 8 (9.0) | 28 (11.4) | 106 (69.7) | 167 (28.4) | 236 (28.6) |
| Disease stage, n (%) | MCI due to AD | 154 (64.7) | 34 (65.4) | 33 (68.8) | 52 (58.4) | 166 (67.5) | 90 (59.2) | 375 (63.9) | 529 (64.1) |
| | Mild AD | 84 (35.3) | 18 (34.6) | 15 (31.3) | 37 (41.6) | 80 (32.5) | 62 (40.8) | 212 (36.1) | 296 (35.9) |
| AChEIs and/or memantine at Baseline, n (%) | No | 110 (46.2) | 24 (46.2) | 23 (47.9) | 33 (37.1) | 115 (46.7) | 73 (48.0) | 268 (45.7) | 378 (45.8) |
| | Yes | 128 (53.8) | 28 (53.8) | 25 (52.1) | 56 (62.9) | 131 (53.3) | 79 (52.0) | 319 (54.3) | 447 (54.2) |
| Number of years of disease since diagnosis | n | 237 | 52 | 48 | 89 | 245 | 152 | 586 | 823 |
| | Mean (SD) | 2.38 (1.659) | 2.27 (1.705) | 2.08 (1.235) | 2.16 (1.242) | 2.20 (1.551) | 2.22 (1.491) | 2.19 (1.479) | 2.25 (1.534) |
| | Median | 2.00 | 2.00 | 2.00 | 2.00 | 2.00 | 2.00 | 2.00 | 2.00 |
| | Min, max | 1.0, 11.0 | 1.0, 7.0 | 1.0, 6.0 | 1.0, 6.0 | 1.0, 12.0 | 1.0, 9.0 | 1.0, 12.0 | 1.0, 12.0 |

| Category | | Placebo (N=238) | BAN2401 2.5 mg/kg Biweekly (N=52) | BAN2401 5 mg/kg Monthly (N=48) | BAN2401 5 mg/kg Biweekly (N=89) | BAN2401 10 mg/kg Monthly (N=246) | BAN2401 10 mg/kg Biweekly (N=152) | BAN2401 Total (N=587) | Combined Total (N=825) |
|---|---|---|---|---|---|---|---|---|---|
| Age at diagnosis (years) | n | 237 | 52 | 48 | 89 | 245 | 152 | 586 | 823 |
| | Mean (SD) | 70.32 (8.740) | 69.75 (8.364) | 69.94 (7.575) | 70.09 (7.442) | 70.71 (7.526) | 72.03 (8.855) | 70.81 (7.971) | 70.67 (8.197) |
| | Median | 70.00 | 70.00 | 71.00 | 71.00 | 71.00 | 73.00 | 71.00 | 71.00 |
| | Min, max | 50.0, 90.0 | 49.0, 86.0 | 54.0, 84.0 | 52.0, 87.0 | 52.0, 90.0 | 51.0, 89.0 | 49.0, 90.0 | 49.0, 90.0 |
| Age group at diagnosis (years), n (%) | 30 to <50 | 0 | 1 (1.9) | 0 | 0 | 0 | 0 | 1 (<1.0) | 1 (<1.0) |
| | 50 to <65 | 60 (25.2) | 11 (21.2) | 9 (18.8) | 19 (21.3) | 50 (20.3) | 29 (19.1) | 118 (20.1) | 178 (21.6) |
| | 65 to <80 | 143 (60.1) | 36 (69.2) | 35 (72.9) | 63 (70.8) | 166 (67.5) | 92 (60.5) | 392 (66.8) | 535 (64.8) |
| | ≥80 | 34 (14.3) | 4 (7.7) | 4 (8.3) | 7 (7.9) | 29 (11.8) | 31 (20.4) | 75 (12.8) | 109 (13.2) |
| Number of years since onset of symptoms | n | 238 | 52 | 48 | 89 | 246 | 152 | 587 | 825 |
| | Mean (SD) | 4.70 (2.467) | 4.67 (2.212) | 4.90 (2.195) | 4.31 (1.578) | 4.41 (2.484) | 4.30 (2.219) | 4.43 (2.250) | 4.51 (2.317) |
| | Median | 4.00 | 4.00 | 5.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 |
| | Min, max | 1.0, 17.0 | 1.0, 11.0 | 1.0, 11.0 | 2.0, 10.0 | 1.0, 16.0 | 1.0, 15.0 | 1.0, 16.0 | 1.0, 17.0 |
| Age at onset of symptoms (years) | n | 238 | 52 | 48 | 89 | 246 | 152 | 587 | 825 |
| | Mean (SD) | 68.00 (8.880) | 67.35 (8.220) | 67.13 (7.601) | 67.93 (7.513) | 68.48 (7.815) | 69.95 (9.057) | 68.57 (8.156) | 68.40 (8.370) |
| | Median | 68.00 | 68.00 | 67.50 | 69.00 | 69.00 | 71.00 | 69.00 | 69.00 |
| | Min, max | 46.0, 88.0 | 47.0, 83.0 | 51.0, 82.0 | 50.0, 87.0 | 45.0, 89.0 | 47.0, 87.0 | 45.0, 89.0 | 45.0, 89.0 |
| Age group at onset of symptoms (years) | 30 to <50 | 3 (1.3) | 2 (3.8) | 0 | 0 | 2 (<1.0) | 2 (1.3) | 6 (1.0) | 9 (1.1) |
| | 50 to <65 | 72 (30.3) | 15 (28.8) | 17 (35.4) | 25 (28.1) | 73 (29.7) | 36 (23.7) | 166 (28.3) | 238 (28.8) |
| | 65 to <80 | 142 (59.7) | 33 (63.5) | 29 (60.4) | 61 (68.5) | 155 (63.0) | 87 (57.2) | 365 (62.2) | 507 (61.5) |
| | ≥80 | 21 (8.8) | 2 (3.8) | 2 (4.2) | 3 (3.4) | 16 (6.5) | 27 (17.8) | 50 (8.5) | 71 (8.6) |

AChEIs = acetylcholinesterase inhibitor, AD = Alzheimer's disease, ApoE4 = apolipoprotein ε4 variant, CDR = Clinical Dementia Rating, MCI = mild cognitive impairment.
a: Age was calculated at date of informed consent. Percentages are based on total number of subjects with non-missing values in relevant treatment group.
Source: Table 14.1.4.1.1

Note: Copied from page 152-154 of applicant study report

## Baseline Disease Characteristics

The overall scores (mean [SD]) for Baseline clinical characteristics (ADCOMS, CDR-SB, ADAS–Cog14, MMSE, and FAQ) were similar in the placebo group (0.37 [0.17], 2.89 [1.45], 22.56 [7.66], 26.01 [2.35] and, 6.66 [5.65], respectively) and BAN2401 groups overall (0.38 [0.16], 2.95 [1.37], 22.23 [7.40], 25.62 [2.37], and 7.18 [5.77], respectively) (Table 5). The mean scores in the individual BAN2401 groups and the BAN2401 10 mg/kg monthly and biweekly groups combined were likewise similar.

**Table 5. Baseline Summary Statistics of Clinical Efficacy Endpoints Overall – Full Analysis Set**

| | | | BAN2401 | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Parameter** | | **Placebo (N=238)** | **2.5 mg/kg Biweekly (N=52)** | **5 mg/kg Monthly (N=48)** | **5 mg/kg Biweekly (N=89)** | **10 mg/kg Monthly (N=246)** | **10 mg/kg Biweekly (N=152)** | **Total (N=587)** |
| ADCOMS | n | 238 | 52 | 48 | 89 | 246 | 152 | 587 |
| | Mean (SD) | 0.370 (0.1663) | 0.386 (0.1970) | 0.395 (0.1746) | 0.390 (0.1558) | 0.373 (0.1522) | 0.373 (0.1508) | 0.378 (0.1584) |
| | Median | 0.36 | 0.38 | 0.36 | 0.39 | 0.36 | 0.37 | 0.37 |
| | Min, max | 0.05, 0.94 | 0.07, 0.87 | 0.10, 0.78 | 0.11, 0.78 | 0.06, 0.89 | 0.04, 0.87 | 0.04, 0.89 |
| CDR-SB | n | 238 | 52 | 48 | 89 | 246 | 152 | 587 |
| | Mean (SD) | 2.89 (1.454) | 2.98 (1.584) | 2.94 (1.420) | 3.03 (1.314) | 2.91 (1.320) | 2.97 (1.401) | 2.95 (1.369) |
| | Median | 3.00 | 3.00 | 2.50 | 3.00 | 2.50 | 3.00 | 3.00 |
| | Min, max | 0.50, 9.00 | 0.50, 7.00 | 1.00, 6.00 | 0.50, 6.50 | 0.50, 8.00 | 0.50, 8.50 | 0.50, 8.50 |
| ADAS–Cog14 | n | 237 | 52 | 47 | 89 | 246 | 152 | 586 |
| | Mean (SD) | 22.56 (7.657) | 22.72 (8.050) | 22.94 (7.735) | 22.75 (6.696) | 21.90 (7.302) | 22.06 (7.667) | 22.23 (7.400) |
| | Median | 22.00 | 22.50 | 22.33 | 23.67 | 21.33 | 22.67 | 22.00 |
| | Min, max | 6.00, 46.67 | 10.00, 42.33 | 8.67, 47.33 | 8.67, 39.00 | 3.67, 48.33 | 4.33, 42.00 | 3.67, 48.33 |
| MMSE | n | 238 | 52 | 48 | 89 | 246 | 152 | 587 |
| | Mean (SD) | 26.01 (2.348) | 25.67 (2.487) | 25.25 (2.622) | 25.60 (2.260) | 25.71 (2.364) | 25.61 (2.351) | 25.62 (2.373) |
| | Median | 26.00 | 26.00 | 25.00 | 26.00 | 26.00 | 26.00 | 26.00 |
| | Min, max | 22.00, 30.00 | 22.00, 30.00 | 22.00, 30.00 | 22.00, 30.00 | 21.00, 30.00 | 22.00, 30.00 | 21.00, 30.00 |
| FAQ | n | 238 | 52 | 48 | 89 | 246 | 152 | 587 |
| | Mean (SD) | 6.66 (5.651) | 8.13 (6.471) | 8.12 (6.267) | 6.92 (5.380) | 6.82 (5.430) | 7.28 (6.094) | 7.18 (5.767) |
| | Median | 6.00 | 8.00 | 8.00 | 6.25 | 5.78 | 6.00 | 6.00 |
| | Min, max | 0.00, 24.00 | 0.00, 23.00 | 0.00, 25.00 | 0.00, 25.00 | 0.00, 26.00 | 0.00, 28.89 | 0.00, 28.89 |

ADCOMS = Alzheimer's Disease Composite Score, ADAS-Cog14 = Alzheimer Disease Assessment Scale - Cognitive Subscale with 14 tasks, CDR-SB = Clinical Dementia Rating-Sum of Boxes, FAQ = Functional Assessment Questionnaire, Max = maximum, Min = minimum, MMSE = Mini-Mental State Examination.
Source: Table 14.1.4.1.3

Note: Table was copied from applicant study report page 157

## 3.2.1.4 Results and Conclusions

### *3.2.1.4.1 Applicant's Results*

Table 6 shows summary statistics for the primary efficacy endpoint, ADCOMS, by group at 12 Months in the Full Analysis set.

APPEARS THIS WAY ON ORIGINAL

**Table 6. Summary Statistics for Change From Baseline in ADCOMS at 12 Months - Full Analysis Set**

| | | BAN2401 | | | | | |
|---|---|---|---|---|---|---|---|
| **Parameter<br>Visit<br>Statistic** | **Placebo (N=238)** | **2.5 mg/kg Biweekly (N=52)** | **5 mg/kg Monthly (N=48)** | **5 mg/kg Biweekly (N=89)** | **10 mg/kg Monthly (N=246)** | **10 mg/kg Biweekly (N=152)** | **Combined 10 mg/kg Monthly and Biweekly (N=398)** |
| **ADCOMS – Overall** | | | | | | | |
| Baseline | | | | | | | |
| n | 238 | 52 | 48 | 89 | 246 | 152 | 398 |
| Mean (SD) | 0.370 (0.1663) | 0.386 (0.1970) | 0.395 (0.1746) | 0.390 (0.1558) | 0.373 (0.1522) | 0.373 (0.1508) | 0.373 (0.1515) |
| Median | 0.36 | 0.38 | 0.36 | 0.39 | 0.36 | 0.37 | 0.36 |
| Min, max | 0.05, 0.94 | 0.07, 0.87 | 0.10, 0.78 | 0.11, 0.78 | 0.06, 0.89 | 0.04, 0.87 | 0.04, 0.89 |
| Week 53 (Month 12) | | | | | | | |
| n | 206 | 42 | 45 | 69 | 181 | 98 | 279 |
| Mean (SD) | 0.460 (0.2454) | 0.534 (0.2956) | 0.492 (0.2538) | 0.477 (0.2177) | 0.444 (0.2215) | 0.460 (0.2266) | 0.450 (0.2230) |
| Median | 0.42 | 0.46 | 0.48 | 0.45 | 0.41 | 0.42 | 0.41 |
| Min, max | 0.03, 1.50 | 0.03, 1.23 | 0.10, 1.49 | 0.07, 0.99 | 0.05, 1.26 | 0.07, 1.19 | 0.05, 1.26 |
| Change from Baseline | | | | | | | |
| n | 206 | 42 | 45 | 69 | 181 | 98 | 279 |
| Mean (SD) | 0.102 (0.1554) | 0.149 (0.2007) | 0.106 (0.1687) | 0.098 (0.1411) | 0.079 (0.1627) | 0.076 (0.1442) | 0.078 (0.1562) |
| Median | 0.09 | 0.12 | 0.10 | 0.06 | 0.05 | 0.05 | 0.05 |
| Min, max | -0.41, 0.70 | -0.30, 0.78 | -0.13, 0.87 | -0.16, 0.37 | -0.27, 0.90 | -0.18, 0.57 | -0.27, 0.90 |

Only subjects with non-missing data at both Baseline and the relevant post-Baseline visit are included in the change from Baseline summary statistics.
ADCOMS = Alzheimer's Disease Composite Score, Max = maximum, Min = minimum.
Source: Table 14.2.1.3a and Table 14.2.1.3.1a

Note: copied from page 167 of applicant's study report

The primary efficacy analysis to evaluate the effect of the ED90 dose of BAN2401 on the ADCOMS at 12 months of treatment using Bayesian analysis was conducted after all 856 randomized subjects had received their 12-month assessment or had discontinued before 12 months. The success criterion for the primary analysis was prespecified to require at least 80% probability that ED90 was better than placebo by a CSD of 25% difference from placebo on ADCOMS change from Baseline. The results of this Bayesian analysis are summarized in Table 7 overall for individual treatment groups. Similarly, these analyses were applied to the key secondary endpoints of change from Baseline in ADCOMS, CDR-SB and ADAS-Cog14 at 18 months.

Further details of the Bayesian analyses are provided in the Appendix.

For ADCOMS change from Baseline at 12 months, there was 21% less decline in the BAN2401 10 mg/kg monthly group and 30% less decline in the 10 mg/kg biweekly group compared with placebo. The actual probability for the early success criterion according to the Bayesian analysis at 12 months for the ED90 dose of 10 mg/kg biweekly was 64%; this increased to 76% at 18 months. The probability of BAN2401 being superior to placebo by any difference according to the Bayesian analysis at 12 months was greatest for the 10 mg/kg monthly and 10 mg/kg biweekly doses, 93.2% and 96.7%, respectively.

**Table 7. Bayesian Analysis of ADCOMS at 12 Months-Full Analysis Set**

| | | Change from Baseline | | Posterior Quantities | | | |
|---|---|---|---|---|---|---|---|
| **Treatment Group** | **Total N** | **Mean** | **SD** | **Pr (Max)** | **Pr (ED90)** | **Pr Superiority** | **Pr (CSD)** |
| **ADCOMS – Overall** | | | | | | | |
| Placebo control | 229 | 0.113 | 0.012 | - | - | - | - |
| 2.5 mg/kg biweekly | 51 | 0.141 | 0.024 | 0.005 | 0.005 | 0.142 | 0.013 |
| 5 mg/kg monthly | 48 | 0.114 | 0.021 | 0.047 | 0.062 | 0.490 | 0.097 |
| 5 mg/kg biweekly | 87 | 0.114 | 0.016 | 0.019 | 0.020 | 0.482 | 0.061 |
| 10 mg/kg monthly | 242 | 0.089 | 0.011 | 0.259 | 0.312 | 0.932 | 0.365 |
| 10 mg/kg biweekly | 143 | 0.079 | 0.014 | 0.670 | 0.600 | **0.967** | **0.585** |

Only subjects with non-missing data at both Baseline and the relevant post-Baseline visit are included in the change from Baseline summary statistics.
ADCOMS = Alzheimer's Disease Composite Score, CSD = clinically significant difference, ED90 = dose regimen with at least 90% of the $d_{max}$ treatment effect Max = maximum, Pr = probability.
Note: copied from page 169 of applicant's study report

A cognitive progression model was used to estimate Cognitive Rate Ratio (CRR), i.e, the relative rate of decline in BAN2401 dose compared with placebo using change from Baseline in the ADCOMS at 18 months (Table 8). The dose-response model identified the 10 mg/kg biweekly dose as the most likely dose with maximum effect. dmax, and the most likely ED90 (dose achieving 90% of the maximum effect). The model estimated a mean CRR of 0.658 (95% CI: 0.517, 0.816) for this dose. There was a 99.8% probability that the CRR was at most 0.90. The residual SD for this model was 0.100 with an estimated random effect SD of 0.161.

**Table 8. Model Based Estimates of Cognitive Rate Ratio -Full Analysis Set at 18 Months**

| BAN2401 Treatment Group | Pr (Max) | Pr (ED90) | CRR | | Pr (CRR <0.9) |
|---|---|---|---|---|---|
| | | | Mean | 95% CI | |
| **ADCOMS – Overall** | | | | | |
| 2.5 mg/kg biweekly | 0.017 | 0.027 | 0.911 | (0.716, 1.121) | 0.470 |
| 5 mg/kg monthly | 0.001 | 0.003 | 1.053 | (0.812, 1.304) | 0.107 |
| 5 mg/kg biweekly | 0.002 | 0.003 | 0.974 | (0.794, 1.176) | 0.224 |
| 10 mg/kg monthly | 0.016 | 0.029 | 0.852 | (0.724, 0.992) | 0.763 |
| 10 mg/kg biweekly | 0.964 | 0.938 | 0.658 | (0.517, 0.816) | 0.998 |

CRR = cognitive rate ratio, ED90 = dose regimen with at least 90% of the $d_{max}$ treatment effect, Max = maximum, Pr = probability.
Note: this table copied from page 190 of applicant's study report

**Conventional Analyses on ADCOMS at 18 Months**

The conventional MMRM analyses (censored at the time of initiation or change of AChEIs or memantine treatment regimens) in the Full Analysis Set are as follows. In the individual treatment group analysis the LS mean changes from Baseline in ADCOMS at 18 months were 0.193 for placebo, and 0.173, 0.192, 0.199, 0.166, and 0.136 for BAN2401 2.5 mg/kg biweekly, 5 mg/kg monthly, 5 mg/kg biweekly, 10 mg/kg monthly, and 10 mg/kg biweekly groups, respectively. The respective LS mean differences from placebo for the individual BAN2401 treatment groups were -0.020, -0.001, 0.006, -0.028, and -0.057 for the 2.5 mg/kg biweekly, 5 mg/kg monthly, 5 mg/kg biweekly, 10 mg/kg monthly, and 10 mg/kg biweekly groups, respectively. The LS mean difference of -0.057 (*P*=0.034; 90% CI: -0.102, -0.013) for 10 mg/kg biweekly showed 30% less decline than placebo.

BAN2401 10 mg/kg monthly showed 14% less decline than placebo; the LS mean difference from placebo was -0.028 (*P*=0.228; 90% CI: -0.065, 0.010). In the analysis of the combined BAN2401 10 mg/kg treatment groups versus placebo, the LS mean change from Baseline in ADCOMS at 18 months was 0.152 for the combined 10 mg/kg groups compared to 0.190 for placebo. The LS mean difference of the combined BAN2401 10 mg/kg groups from placebo was -0.039 (P=0.053; 90% CI: -0.071, -0.006), representing 20% less decline than on placebo.

**Conventional Analyses on CDR-SB at 18 Months**

The MMRM analyses (censored at the time of initiation or change of AChEIs or memantine

treatment regimens) in the Full Analysis Set are presented in Table 9. The 18-month results are summarized for the placebo group, individual BAN2401 treatment groups, and combined BAN2401 10 mg/kg (monthly and biweekly) groups.

In the individual treatment group analysis the LS mean changes from Baseline in CDR-SB at 18 months were 1.499 for placebo, and 1.248, and 1.102 for BAN2401 10 mg/kg monthly, and 10 mg/kg biweekly groups, respectively. The respective LS mean differences from placebo were -0.250, and -0.396. These values demonstrated numerically less decline in CDR-SB score compared to placebo for all BAN2401 doses apart from 5 mg/kg monthly, with the biggest effect noted for 10 mg/kg biweekly, which was equivalent to 26% less decline in CDR-SB score compared to placebo. In the analysis of the combined BAN2401 10 mg/kg treatment groups with placebo, the LS mean changes from Baseline in CDR-SB at 18 months were 1.473 for placebo, and 1.171 for the combined 10 mg/kg groups. The LS mean difference of the combined BAN2401 10 mg/kg groups from placebo was -0.302 (P=0.119; 90% CI: -0.620, 0.017), representing 21% less decline than on placebo.

**Table 9 Conventional Analysis of CDRSB at 18 Months-Full Analysis Set**

| Parameter<br>Visit<br>Statistic | Individual Treatment Groups Analysis | | | | | | Combined BAN2401 10 mg/kg Treatment Groups Analysis | |
|---|---|---|---|---|---|---|---|---|
| | | BAN2401 | | | | | | BAN2401 |
| | Placebo (N=238) | 2.5 mg/kg Biweekly (N=52) | 5 mg/kg Monthly (N=48) | 5 mg/kg Biweekly (N=89) | 10 mg/kg Monthly (N=246) | 10 mg/kg Biweekly (N=152) | Placebo (N=238) | Combined 10 mg/kg Monthly and Biweekly (N=398) |
| **CDR-SB – Overall** | | | | | | | | |
| Week 79 (Month 18) | | | | | | | | |
| n | 161 | 34 | 36 | 67 | 149 | 84 | 161 | 233 |
| LS mean | 1.499 | 1.227 | 1.713 | 1.463 | 1.248 | 1.102 | 1.473 | 1.171 |
| SE | 0.16 | 0.338 | 0.334 | 0.250 | 0.169 | 0.213 | 0.158 | 0.136 |
| LS mean difference: active dose – placebo | - | -0.271 | 0.214 | -0.036 | -0.250 | -0.396 | - | -0.302 |
| 90% CI for differences | - | -0.875, 0.332 | -0.384, 0.812 | -0.510, 0.439 | -0.613, 0.112 | -0.821, 0.028 | - | -0.620, 0.017 |
| *P*-value | - | 0.459 | 0.555 | 0.901 | 0.255 | 0.125 | - | 0.119 |

The change from Baseline for each parameter in overall population was analyzed using the MMRM with treatment group/combined treatment groups, visit, disease stage (MCI due to AD, mild AD dementia), ApoE4 status (carrier, non-carrier), presence or absence of concomitant AD treatment (AChEIs and/or memantine) at Baseline, region, treatment group-by-visit interaction as factors, and Baseline value as covariate. The mixed-effects model within each randomization stratum (subgroup) was similar and was reduced by removing corresponding stratification factor from the model in overall population. Subjects were censored at the time of initiation or change of AChEIs or memantine treatment regimens.
AChEIs = acetylcholinesterase inhibitor, AD = Alzheimer's disease, ApoE4 = apolipoprotein ϵ4 variant, CDR-SB = Clinical Dementia Rating-Sum of Boxes, LS = least square, MCI = mild cognitive impairment, MMRM = mixed-effects model with repeated measures.

Note: This table was copied from page 206 of the applicant's study report

In the conventional MMRM analysis of CDR-SB, the LS mean differences from placebo for BAN2401 10 mg/kg monthly and biweekly were -0.399 (P=0.111) and -0.985 (P=0.083), respectively in ApoE4 carriers, representing 24% and 60% less decline, respectively compared to placebo. The respective LS mean differences of 0.037 (P=0.945) and 0.077 (P=0.827) for non-carriers, represented negligible difference compared to placebo.

**Amyloid PET SUVr Normalized to Whole Cerebellum Mask Observed 18-Month Data**

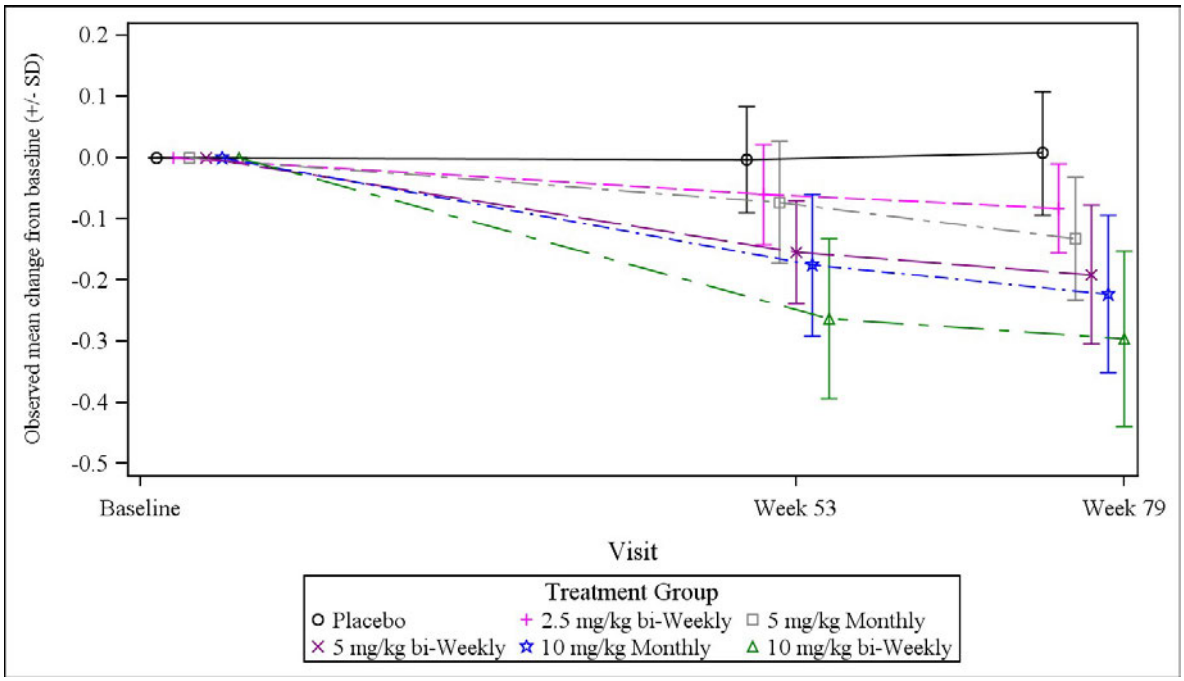
This section presents the results for change from Baseline at the 18-month timepoint for amyloid PET SUVr normalized to whole cerebellum mask; the observed mean change to 18 months is presented in **Figure 1**, and summary statistics for observed values are in Table 10 for the placebo

group, the individual BAN2401 treatment groups, and the combined BAN2401 10 mg/kg groups (monthly and biweekly).

BAN2401 demonstrated a dose-dependent and time-dependent amyloid reduction across all doses versus placebo. The Baseline mean amyloid PET SUVr was 1.4 across all treatment groups. A mean reduction of 0.22 and 0.30 was observed in PET SUVr after 18 months of BAN2401 treatment on 10 mg/kg monthly and biweekly doses, respectively. These amyloid PET reductions resulted in SUVr values approaching the numerical discrimination point of 1.14 for amyloid positivity using the florbetapir imaging agent. The observed mean value of 1.07 at 18 months for the 10 mg/kg biweekly dose was below this threshold. Overall, the results indicate that with BAN2401 treatment there is significant clearance of amyloid from the brain.

The brain amyloid PET SUVr analysis showed a slight increase in amyloid load in the placebo group, reflective of the expected absolute percentage annual SUVr change of approximately 1% to 2% across all reference regions in a population with MCI due to AD and mild AD dementia (Landau et al, 2015).

**Figure 1 Observed Mean (SD) Change From Baseline in Brain Amyloid Levels as Measured by Amyloid PET SUVr Normalized to Whole Cerebellum Mask by Visit - Overall - PD Analysis Set 2**




PD Analysis Set 2 is the group of subjects who had sufficient amyloid PET data to derive at least 1 amyloid PET parameter.
PD = pharmacodynamics, PET = positron emission tomography, SUVr = standard uptake value ratio.
Note: Copied from page 175 of the applicant's study report

**Conventional Analysis of Amyloid PET SUVr Normalized to Whole Cerebellum Mask at 18-Months**

The 18-month results for amyloid PET SUVr normalized to whole cerebellum mask including the LS means and the LS mean differences between active treatment and placebo are summarized for the placebo group, individual BAN2401 treatment groups, and combined BAN2401 10 mg/kg groups in Table 10.

In the individual treatment group analysis the LS mean changes from Baseline in amyloid PET SUVr normalized to whole cerebellum mask at 18 months were 0.004 for placebo, and -0.094, -0.131, -0.197, -0.225, and -0.306 for BAN2401 2.5 mg/kg biweekly, 5 mg/kg monthly, 5 mg/kg biweekly, 10 mg/kg monthly, and 10 mg/kg biweekly groups, respectively. The respective LS mean differences from placebo for the individual BAN2401 treatment groups were -0.099, -0.136, -0.201, -0.229, and -0.310 (P<0.001 for all doses). These results demonstrated a dose-dependent reduction in amyloid PET SUVr normalized to whole cerebellum mask at 18 months treatment with BAN2401 compared to placebo, for which the greatest change was in the BAN2401 10 mg/kg biweekly group.

In the analysis of the combined BAN2401 10 mg/kg treatment groups versus placebo, the LS mean changes from Baseline in amyloid PET SUVr normalized to whole cerebellum mask at 18 months were -0.003 for placebo, and -0.256 for the combined 10 mg/kg groups. The LS mean difference of the combined BAN2401 10 mg/kg groups from placebo was -0.253 (P<0.001).

**Table 10. Summary of MMRM Analyses of Change From Baseline in Brain Amyloid Levels as Measured by PET SUVr Normalized to Whole Cerebellum Mask at 18 Months - PD Analysis Set 2**

| Parameter<br>Visit<br>Statistic | Individual Treatment Groups Analysis | | | | | | Combined BAN2401 10 mg/kg Treatment Groups Analysis | |
|---|---|---|---|---|---|---|---|---|
| | | BAN2401 | | | | | | BAN2401 |
| | Placebo (N=99) | 2.5 mg/kg Biweekly (N=28) | 5 mg/kg Monthly (N=28) | 5 mg/kg Biweekly (N=27) | 10 mg/kg Monthly (N=89) | 10 mg/kg Biweekly (N=44) | Placebo (N=99) | Combined 10 mg/kg Monthly and Biweekly (N=133) |
| **Amyloid PET SUVr – Overall** | | | | | | | | |
| Change from Baseline | | | | | | | | |
| Week 79 (Month 18) | | | | | | | | |
| n | 88 | 23 | 23 | 24 | 82 | 37 | 88 | 119 |
| LS mean | 0.004 | -0.094 | -0.131 | -0.197 | -0.225 | -0.306 | -0.003 | -0.256 |
| SE | 0.011 | 0.020 | 0.020 | 0.021 | 0.012 | 0.016 | 0.012 | 0.010 |
| LS mean difference: active dose – placebo | | -0.099 | -0.136 | -0.201 | -0.229 | -0.310 | | -0.253 |
| 90% CI for differences | - | -0.136, -0.061 | -0.173, -0.098 | -0.238, -0.164 | -0.254, -0.204 | -0.344, -0.277 | - | -0.277, -0.229 |
| *P*-value[a] | - | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | - | <0.001 |
| Dunnett *P*-value[b] | - | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | - | 0.000 |
| Percentage change from Baseline | | | | | | | | |
| Week 79 (Month 18) | | | | | | | | |
| n | 88 | 23 | 23 | 24 | 82 | 37 | 88 | 119 |
| LS mean | 0.645 | -6.589 | -9.124 | -13.893 | -15.357 | -21.516 | 0.082 | -17.708 |
| SE | 0.772 | 1.404 | 1.411 | 1.421 | 0.838 | 1.127 | 0.804 | 0.681 |
| LS mean difference: active dose – placebo | - | -7.234 | -9.769 | -14.538 | -16.002 | -22.161 | - | -17.790 |
| 90% CI for differences | - | -9.794, -4.674 | -12.347, -7.191 | -17.107, -11.969 | -17.726, -14.279 | -24.460, -19.862 | - | -19.417, -16.162 |
| *P*-value[a] | - | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | - | <0.001 |
| Dunnett *P*-value[b] | - | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | - | 0.000 |

PD Analysis Set 2 is the group of subjects who had sufficient amyloid PET data to derive at least 1 amyloid PET parameter. The change from Baseline and percentage
change from Baseline in brain amyloid level parameters were analyzed using the MMRM with treatment group/combined treatment groups, visit, disease stage (MCI due to
AD, Mild AD), the presence or absence of concomitant AD treatment (AChEIs and/or memantine) at Baseline, ApoE4 status (positive, negative), region, treatment group by-
visit interaction as factors, and Baseline brain amyloid levels as covariate. The mixed-effects model within each randomization stratum (subgroup) was similar and was
reduced by removing corresponding stratification factor from the model in overall population. SUVr of global cortical average with respect to reference region. Only PET
data using florbetapir tracer was included.
AChEIs = acetylcholinesterase inhibitor, AD = Alzheimer's disease, ApoE4 = apolipoprotein ϵ4 variant, LS = least square, MCI = mild cognitive impairment, MMRM =
mixed-effects model with repeated measures, PD = pharmacodynamics, PET = positron emission tomography, SUVr = standard uptake value ratio.
a: P-values for comparing each combined BAN2401 group versus the placebo group.
b: adjusted P-values based on Dunnett-Hsu method (see Section 9.7.1.6.2).
Note: This table was copied from page 183 of applicants study report

The LS means (MMRM) of changes from Baseline in ADCOMS in subjects enrolled in the PET substudy in the Full Analysis Set are in

Table **11 .**

**Table 11. Least square means (LSM) of Change from Baseline at 18 months in ADCOMS in subjects enrolled in PET substudy - MMRM**

| | | BAN2401 | | | | |
|---|---|---|---|---|---|---|
| Visit<br>Statistic | Placebo<br>(N = 98) | 2.5 mg/kg<br>bi-Weekly<br>(N = 28) | 5 mg/kg<br>Monthly<br>(N = 27) | 5 mg/kg<br>bi-Weekly<br>(N = 26) | 10 mg/kg<br>Monthly<br>(N = 88) | 10 mg/kg<br>bi-Weekly<br>(N = 44) |
| Week 79 | | | | | | |
| n | 77 | 21 | 21 | 21 | 75 | 32 |
| Least Square Mean | 0.164 | 0.143 | 0.148 | 0.075 | 0.117 | 0.071 |
| SE | 0.021 | 0.038 | 0.038 | 0.039 | 0.022 | 0.031 |
| LS Mean Difference: Active Dose - Placebo | | -0.021 | -0.016 | -0.089 | -0.047 | -0.093 |
| 90% Confidence Interval for Differences | | -0.091, 0.050 | -0.086, 0.054 | -0.161, -0.017 | -0.094, 0.001 | -0.156, -0.031 |
| p-value | | 0.631 | 0.704 | 0.042 | 0.104 | 0.014 |

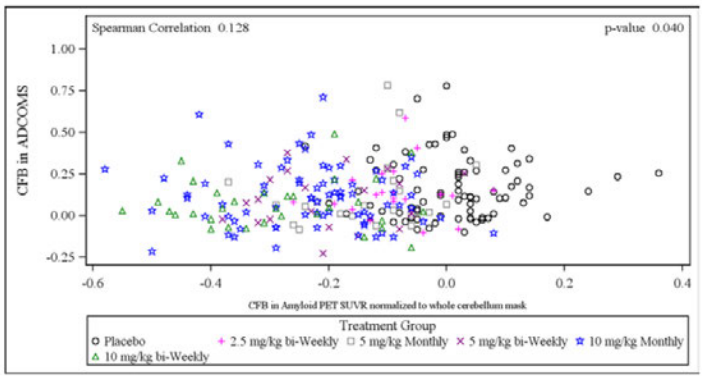
Source: Listing 16.2.6.1.1
The change from baseline for each parameter in overall population is analyzed using the mixed effects model with repeated measures (MMRM) with treatment group, visit, clinical subgroup (MCI due to AD, Mild Alzheimer's Disease Dementia), the presence or absence of ongoing AD treatment (ie, AChEIs and/or memantine) at baseline, APOE4 status (positive, negative), region, treatment group-by-visit interaction as factors, and baseline value as covariate. The mixed effects model within each randomization stratum (subgroup) is similar and is reduced by removing corresponding stratification factor from the model in overall population.
Subjects are censored at the time of initiation or change of AChEIs or memantine treatment regimens.
Only subjects enrolled in PET substudy with non-missing data at 12 and/or 18 months are included.
Program: ../BAN2401/G000-201/CSR/dev/pg/tables/Table 14.2.1.11a.sas 07FEB2019 14:25

Note: This table was copied from page 4893 of the applicant's study report

## Observed Changes From Baseline in Clinical Endpoints vs Changes From Baseline in Amyloid PET SUVr

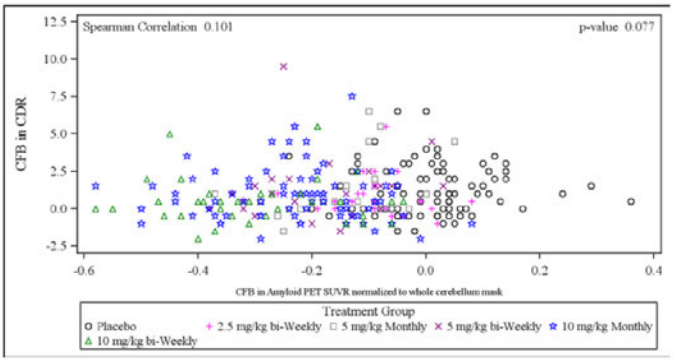
Scatterplots of observed changes from Baseline in ADCOMS, and CDR-SB, and change from Baseline in amyloid PET SUVr normalized to whole cerebellum mask at 18 months are in Figure 2 and Figure 3, respectively. These scatterplots show that points representing subjects in the BAN2401 10 mg/kg biweekly and 10 mg/kg groups in particular are mainly towards the left-hand lower corner of the plot, whereas points representing placebo subjects are mainly on the right-hand side of the graph. These observations indicated there is a correlation between slowing in decline of clinical endpoint scores and decrease in amyloid as measured by PET SUVr after treatment with BAN2401.

**Figure 2. Patient-level Relationship between Week 79 ADCOMS and SUVR Change**




Note: copied from page 288 of study report

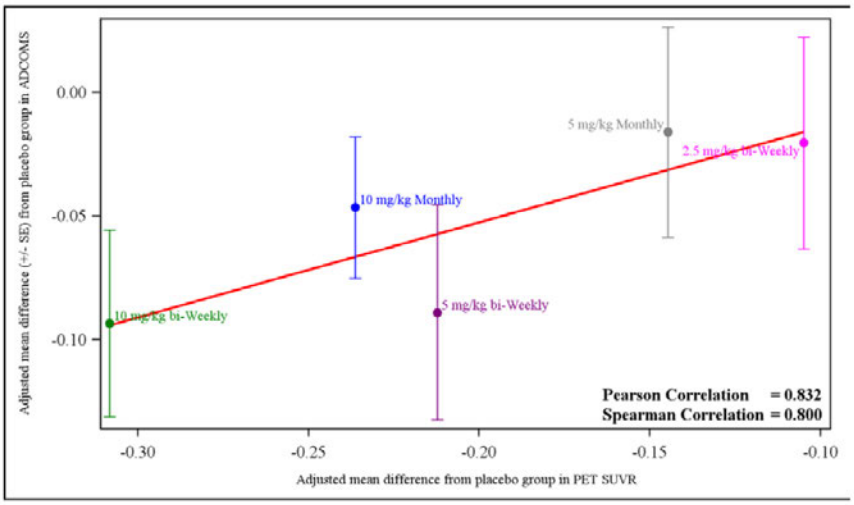
**Figure 3. Patient-level Relationship between Week 79 CDRSB and SUVR Change**




Note: copied from page 289 of study report

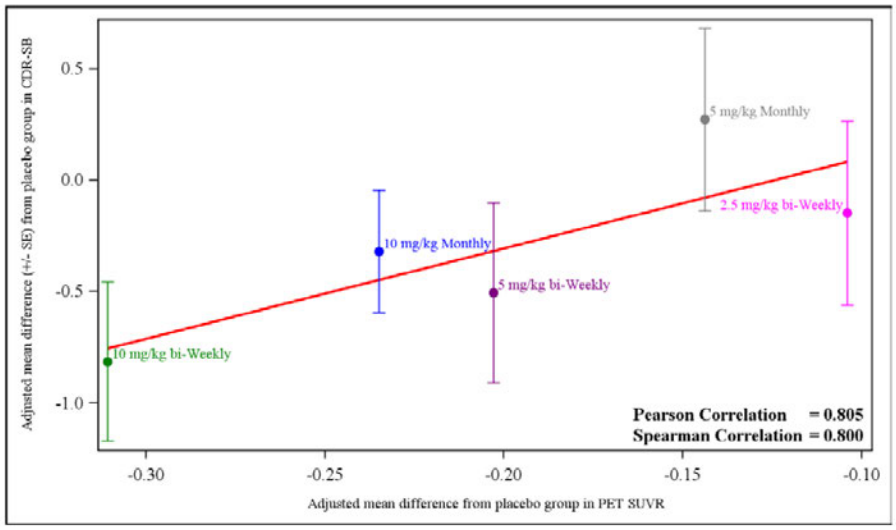
The patient level Spearman correlation between change from baseline in ADCOMS and change from baseline in SUVR at Week 79 was 0.128 (Figure 2). The group level correlation was larger. In particular, an increased treatment effect in amyloid brain clearance as measured by amyloid PET SUVr was associated with increased treatment effect on ADCOMS (Pearson correlation coefficient = 0.832), CDR-SB (Pearson correlation coefficient = 0.805) and ADAS–Cog14 (Pearson correlation coefficient = 0.699) (Figure 4 and Figure 5, respectively).

**Figure 4. Group Level Mean Relationship between Week 79 ADCOMS and SUVR change**




Note: copied from page 290 of study report

**Figure 5. Group Level Mean Relationship between Week 79 CDRSB and SUVR change**




Note: copied from page 291

**Mediation Analysis on Clinical Endpoints and Amyloid PET SUVr**

Mediation analyses were performed to explore the link between the effect of BAN2401 on the clinical endpoints and on brain amyloid load as measured by brain amyloid PET SUVr. Two post-hoc analyses were performed: 1) to assess the degree of the satisfaction of the Prentice criterion (Prentice, 1989) through hypothesis testing, 2) to quantify the Prentice criterion by assessing how much of the treatment effect with respect to the clinical endpoint can be explained by the treatment effect with respect to brain amyloid PET SUVr as a biomarker (Freedman et al, 1992).

An ANCOVA model was used for assessing the treatment effect with respect to the clinical endpoints and amyloid PET SUVr in the PET substudy population at 18 months. For the hypothesis testing approach, if the treatment effect was no longer significant after controlling for the biomarker in the ANCOVA model as a covariate, then the biomarker was considered to have assumed all of the 'significance' of the treatment in predicting the clinical endpoint (Amur et al, 2015). The quantification approach, referred to as the proportion of treatment effect explained (PTE), was defined as the percentage change of treatment effects estimated from 2 ANCOVA models with or without adjusting for the amyloid PET SUVr. The Prentice criterion and the PTE analyses are referred to as the mediation analyses here as they are equivalent to the Baron and Kenny (1986) approach for testing the mediation effect and the Judd and Kenny (1981) difference of coefficients approach to estimate the mediation effect.

The PTE is known for its limitation of high variability when sample size is small and/or when the overall treatment effect is small. The analysis results for all 5 doses were presented for completeness of information. However, the assessment of the mediation effect only focused on the doses where a nominally statistically significant (at the study specified 0.10 significance level) or borderline significant treatment effect on the clinical endpoint were seen in the PET substudy population. This is to ensure the mediation effect was assessed when there was a certain level of confidence that a treatment effect existed.

The mediation analyses are summarized in Table 12. The Prentice criterion are summarized and the PTE on CDR-SB explained by the reduction in amyloid PET SUVr. Treatment effect was significant ($P<0.001$) on the biomarker (PET SUVr) across all doses (Prentice criteria 1) using ANCOVA model as well as MMRM model. The treatment effect was significant on CDR-SB in the MMRM model without adjusting for the biomarker for the 10 mg/kg biweekly. When the MMRM model on clinical endpoint (CDR-SB) was adjusted for the biomarker (change in PET SUVr), treatment effect on 10 mg/kg was no longer significant (Prentice criteria 4; Modeling and Simulation Analysis Report, report number CPMSBAN2401-002R).

For the top dose (10 mg/kg biweekly), almost all the estimated treatment effect on CDR-SB was removed after adjusting for the change from Baseline in amyloid PET SUVr, with the nominal P-values changed from approximately 0.03 to 0.5 (0.12 to close to 1 for ANCOVA model). The estimated PTE is 67% based on MMRM model (96% based on ANCOVA model). This suggests that the treatment effect of the top dose on CDR-SB is related to the

treatment effect on amyloid PET SUVr.

**Table 12. Analyses of Week 79 Change in CDRSB and Change in PET SUVR including Mediation Analysis**

| Statistic at Week 79 (Month 18) | Placebo (N=99) | BAN2401 | | | | |
|---|---|---|---|---|---|---|
| | | 2.5 mg/kg Biweekly (N=28) | 5 mg/kg Monthly (N=28) | 5 mg/kg Biweekly (N=27) | 10 mg/kg Monthly (N=89) | 10 mg/kg Biweekly (N=44) |
| n | 78 | 21 | 21 | 23 | 76 | 35 |
| **Model 1 for Mediation Analysis - Change From Baseline in PET SUVr Normalized to Whole Cerebellum Mask** | | | | | | |
| LS mean difference: active dose – placebo (SE) | - | -0.0899 (0.0247) | -0.1225 (0.0249) | -0.1953 (0.0240) | -0.2277 (0.0164) | -0.3124 (0.0229) |
| 90% CI for differences | - | -0.1307, -0.0490 | -0.1636, -0.0815 | -0.2348, -0.1557 | -0.2548, -0.2006 | -0.3502, -0.2745 |
| *P*-value | - | 0.0003 | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| **Model 2 for Mediation Analysis - CDR-SB at Week 79 (Month 18) Without Adjusting for PET SUVr** | | | | | | |
| LS mean difference: active dose – placebo (SE) | - | -0.4704 (0.4205) | -0.0915 (0.4238) | -0.5417 (0.4072) | -0.4553 (0.2790) | -0.6275 (0.3894) |
| 90% CI for differences | - | -1.1647, 0.2239 | -0.7914, 0.6083 | -1.2140, 0.1305 | -0.9159, 0.0053 | -1.2705, 0.0155 |
| *P*-value | - | 0.2644 | 0.8292 | 0.1846 | 0.1039 | 0.1084 |
| **Model 3 for Mediation Analysis - CDR-SB at Week 79 (Month 18) After Adjusting for PET SUVr** | | | | | | |
| LS mean difference: active dose – placebo (SE) | - | -0.2987 (0.4299) | 0.1397 (0.4420) | -0.1698 (0.4571) | -0.0218 (0.3710) | -0.0329 (0.5140) |
| 90% CI for differences | - | -1.0085, 0.4112 | -0.5900, 0.8694 | -0.9245, 0.5849 | -0.6344, 0.5907 | -0.8815, 0.8158 |
| *P*-value | - | 0.4879 | 0.7522 | 0.7106 | 0.9531 | 0.9490 |

The ANCOVA mediation analyses included the following factors and covariates: Model 1 included treatment group as a factor and the Baseline PET SUVr value and ApoE4 status as covariates; Model 2 included treatment group as a factor and Baseline CDR-SB value, Baseline PET SUVr and ApoE4 status as covariates; Model 3 included treatment group as a factor and Baseline CDR-SB value, Baseline PET SUVr, change from Baseline in PET SUVr at 18 months, and ApoE4 status as covariates. Subjects were censored at the time of initiation or change of AChEIs or memantine treatment regimens. Only subjects with both change from Baseline in PET SUVr normalized to whole cerebellum mask at 18 months and change from Baseline in CDR-SB at 18 months were included. Only subjects with both change from Baseline in PET SUVr normalized to whole cerebellum mask at 18 months and change from Baseline in CDR-SB at 18 months were included.
AChEIs = acetylcholinesterase inhibitor, ANCOVA = analysis of covariance, ApoE4 = apolipoprotein ε4 variant, CDR-SB = Clinical Dementia Rating-Sum of Boxes, LS = least square, PET = positron emission tomography, SUVr = standard uptake value ratio.
Source: Table 14.2.5.5a, Table 14.2.5.5b and Table 14.2.5.5c

Note: copied from page 298 of study report

*Reviewer's Comment:*

*The mediation analysis of the relationship between CDRSB change and SUVR change, including Model 3 in* Table 12*, was not specified in the statistical analysis plan.*

### *3.2.1.4.2 Reviewer's Results*

In the PET substudy population with Week 79 assessments of CDRSB and SUVR there were some differences in baseline demographics and disease characteristics among the groups. The 10 mg/kg bi-Weekly group had 37.8 % Mild AD for baseline disease stage as compared to 27.3% for placebo (Table 13). Also, 62.2% were Male in the 10 mg/kg bi-Weekly group as compared to 40.9 % Male in the Placebo group.

**Table 13. PET Substudy Analysis Population Baseline Demographics and Disease Characteristics**

| | | Planned Treatment for Period 01 | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 10 mg/kg Monthly | 10 mg/kg bi-Weekly | 2.5 mg/kg bi-Weekly | 5 mg/kg Monthly | 5 mg/kg bi-Weekly | Placebo | All |
| **Sex** | | | | | | | | |
| **F** | **N** | 35 | 14 | 10 | 10 | 11 | 52 | 132 |
| | **%** | 42.7 | 37.8 | 43.5 | 43.5 | 45.8 | 59.1 | 47.7 |
| **M** | **N** | 47 | 23 | 13 | 13 | 13 | 36 | 145 |
| | **%** | 57.3 | 62.2 | 56.5 | 56.5 | 54.2 | 40.9 | 52.3 |
| **Race** | | | | | | | | |
| **BLACK OR AFRICAN AMERICAN** | **N** | 1 | 2 | 1 | | | 2 | 6 |
| | **%** | 1.2 | 5.4 | 4.3 | | | 2.3 | 2.2 |
| **WHITE** | **N** | 81 | 35 | 22 | 23 | 24 | 86 | 271 |
| | **%** | 98.8 | 94.6 | 95.7 | 100.0 | 100.0 | 97.7 | 97.8 |
| **Age** | **Mean** | 71.7 | 72.9 | 70.8 | 69.0 | 72.0 | 71.2 | 71.4 |
| **STRC2** | | | | | | | | |
| **MCI due to AD** | **N** | 62 | 23 | 15 | 19 | 16 | 64 | 199 |
| | **%** | 75.6 | 62.2 | 65.2 | 82.6 | 66.7 | 72.7 | 71.8 |
| **Mild AD** | **N** | 20 | 14 | 8 | 4 | 8 | 24 | 78 |
| | **%** | 24.4 | 37.8 | 34.8 | 17.4 | 33.3 | 27.3 | 28.2 |
| **STRC3** | | | | | | | | |
| **ApoE4 Negative** | **N** | 12 | 33 | 5 | 6 | 3 | 23 | 82 |
| | **%** | 14.6 | 89.2 | 21.7 | 26.1 | 12.5 | 26.1 | 29.6 |
| **ApoE4 Positive** | **N** | 70 | 4 | 18 | 17 | 21 | 65 | 195 |
| | **%** | 85.4 | 10.8 | 78.3 | 73.9 | 87.5 | 73.9 | 70.4 |

| | | Planned Treatment for Period 01 | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 10 mg/kg Monthly | 10 mg/kg bi-Weekly | 2.5 mg/kg bi-Weekly | 5 mg/kg Monthly | 5 mg/kg bi-Weekly | Placebo | All |
| **STRC4** | | | | | | | | |
| **With AChEIs and/or Memant** | **N** | 40 | 16 | 14 | 12 | 15 | 38 | 135 |
| | **%** | 48.8 | 43.2 | 60.9 | 52.2 | 62.5 | 43.2 | 48.7 |
| **Without AChEIs and/or Mem** | **N** | 42 | 21 | 9 | 11 | 9 | 50 | 142 |
| | **%** | 51.2 | 56.8 | 39.1 | 47.8 | 37.5 | 56.8 | 51.3 |
| **BCDRSB** | **Mean** | 2.8 | 2.7 | 2.8 | 2.3 | 2.8 | 2.6 | 2.7 |
| **BSUVR** | **Mean** | 1.42 | 1.36 | 1.42 | 1.46 | 1.41 | 1.40 | 1.41 |
| **All** | **N** | 82 | 37 | 23 | 23 | 24 | 88 | 277 |

The applicant presented the exploratory mediation analysis estimating the proportion of treatment effect on CDRSB explained by the treatment effect on SUVR as shown in Table 12. For the corresponding analysis for the ADCOMS the reviewer found that the 10 mg/kg biweekly proportion of treatment effect on Week 79 Change in ADCOMS explained by Week 79 PET SUVR change is by ANCOVA 50% for 10 mg/kg Monthly and 40% for 10 mg/kg bi-weekly using applicant's code (50% ANCOVA and 58% by MMRM both with censoring of data after change of AD medications as the applicant did in Table 12). These proportions for the mediation analysis corresponding to the primary endpoint are considerably smaller than those reported for the CDRSB. There also seems to be an impact of APOE on the proportions with less explained in non-carriers or no significant clinical effects in non-carriers making these proportions uninterpretable. In APOE4 carriers only the proportion of Week 79 CDRSB treatment effect explained for 10 mg/kg monthly is 72% and only 31% for 10 mg/kg biweekly. In non-carriers the treatment effects unadjusted for biomarker change are not close to nominal significance: 10 mg/kg Monthly has an estimated difference from placebo at Month 18 of -.057 p=0.92 and for 10 mg/kg biweekly this is -0.23 (SE=.43) p=0.59. Therefore, the proportions explained by the biomarker effect are not meaningful in APOE4 non-carriers.

**Table 14. Exploratory Mediation Analysis for CDRSB at Week 79 by ApoE4 Status**

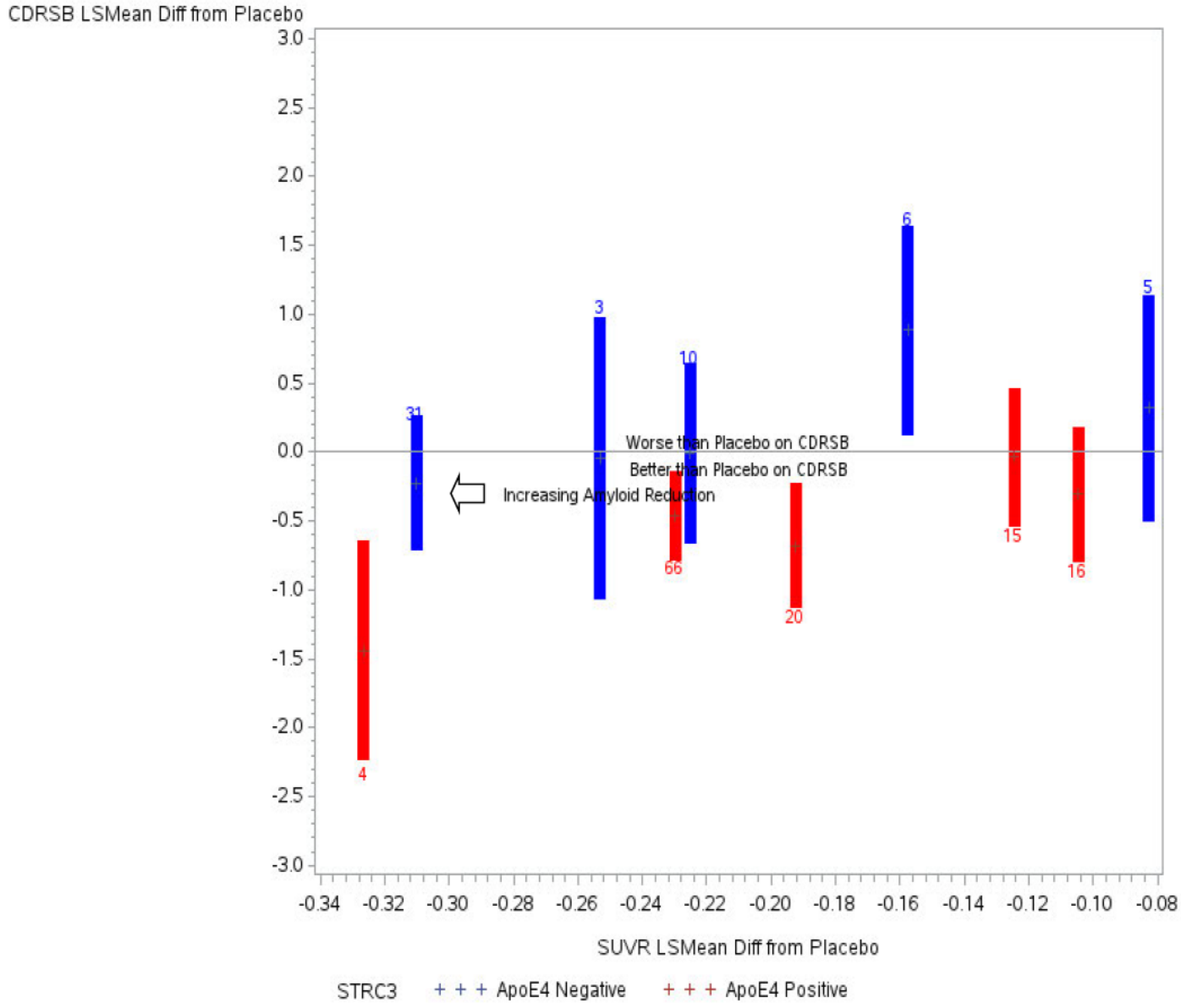
| APOE4 Subgroup | Dose Group | Biomarker Adjusted Clinical Treatment Difference Estimate | Biomarker Unadjusted Clinical Treatment Difference Estimate | Proportion of Clinical Treatment Effect Explained |
|---|---|---|---|---|
| APOE4 Negative | 10 mg/kg Monthly(N=17) | 0.53 (SE=.71) | -0.057 (SE=0.58) p=0.9221 | |
| | 10 mg/kg bi-Weekly(N=74) | 0.58 (S.E.=0.71) | -0.23 (S.E.=0.43) p=0.5907 | |
| | 5 mg/kg bi-Weekly(N=6) | 1.02 (S.E.=1.03) | +0.35 (S.E.=0.93) p=0.7107 | |
| APOE4 Positive | 10 mg/kg Monthly(N=132) | -0.16 (S.E.=0.45) | -0.57 (S.E.=0.32) (p=0.0810) | 0.7183 |
| | 10 mg/kg bi-Weekly(N=10) | -1.14 (S.E.=1.00) | -1.66 (S.E.=0.93) (p=0.0760) | 0.3137 |
| | 5 mg/kg bi-Weekly(N=61) | -0.38 (S.E.=0.53) | -0.71(S.E.=0.47) p=0.1320 | 0.47 |

Note: There were 114 placebo carriers and 47 placebo non-carriers.

Source: Reviewer's Analysis

Figure 6 shows the relationship between Week 79 LSMean Changes in CDRSB vs. LSMean Changes in SUVR by Dose separated by APOE4 status. The relationship may be weaker for the blue bars, i.e., the ApoE4 non-carriers, than for the red bars, the APOE4 carriers.

**Figure 6 MMRM CDRSB LSMeans by SUVR LSmeans subgrouped by ApoE4 status**




## 3.3 Evaluation of Safety

Safety in general is not addressed in this review. Please see the Clinical safety review.

# 4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

## 4.1 Gender, Race, Age, and Geographic Region

### 4.1.1 Gender, Race, and Age

Here we examine Amyloid PET SUVR treatment effects at Week 79 in subgroups.

In Females, the estimated difference for 10 mg/kg biweekly vs. placebo is 0.350[S.E.=0.029].

In Males, the estimated difference for 10 mg/kg biweekly vs. placebo is 0.277[S.E.=0.028].

In Females, the estimated difference for 10 mg/kg monthly vs. placebo is 0.214[S.E.=0.022].

In Males, the estimated difference for 10 mg/kg monthly vs. placebo is 0.241[S.E.=0.022].

The PET population was 90% White, so there are too few subjects in other Race categories to permit meaningful subgroup analyses by Race.

Treatment differences in PET SUVR were biggest in the middle age group, ages 65-80, which was by far the largest subgroup accounting for 67% of the PET substudy population(Table 15). Sample sizes for the other age subgroups, <65 (19%) and >80 (14%), are relatively small.

**Table 15. Week 79 LSMean Change in SUVR by Age Group (MMRM)**

| Age Group | Dose Group vs. Placebo | LSMean Difference from Placebo | Std Error of Difference |
|---|---|---|---|
| <65 | 10 mg/kg M | -0.1583 | 0.03384 |
| 65-80 | 10 mg/kg M | -0.2529 | 0.01831 |
| ≥80 | 10 mg/kg M | -0.1968 | 0.03774 |
| <65 | 10 mg/kg bw | -0.1694 | 0.04931 |
| 65-80 | 10 mg/kg bw | -0.3536 | 0.02295 |
| ≥80 | 10 mg/kg bw | -0.2612 | 0.05081 |

### 4.1.2 Geographic Region

All of the PET SUVR data was collected in North America, so there are no subgroup analyses by region.

#### 4.1.2.1 Individual Sites

There do not appear to be any highly influential sites for efficacy in terms of the biomarker Amyloid PET SUVR.

### 4.2 Other Special/Subgroup Populations

Subgroup effects on SUVR for APOE +/- appeared consistent although the sample size in APOE- was rather small (APOE- 10 mg/kg bi-weekly vs. placebo: 0.310 [S.E.=0.027], APOE+ 10 mg/kg bi-weekly vs. placebo: 0.327 [S.E.=0.044]).

For 10 mg/kg monthly vs. placebo at Week 79 the estimated SUVR difference in APOE carriers was 0.230 [S.E.=0.017], while in non-carriers it was 0.225 [S.E.=0.035].

There were no apparent significant differences in SUVR effects by Baseline disease stage: MCI due to AD or Mild Alzheimer's.

# 5 SUMMARY AND CONCLUSIONS

## 5.1 Statistical Issues

There are many statistical issues with this Bayesian response adaptive randomization trial (Study 201). Multiple dose groups were compared to placebo to identify the best performing dose (ED90). Multiple interim analyses were performed to allow for possibly stopping the trial earlier and having multiple chances to declare the trial is positive. Sample size can depend on the observed treatment difference on the primary endpoint.

Randomization of APOE4 carriers to the 10 mg/kg bi-weekly dose group, the best performing dose, was stopped in the middle of the study at the request of European regulators due to increased risk of Amyloid Related Imaging Abnormalities for carriers at this dose. This led to a large imbalance in APOE4 carrier status between 10 mg/kg bi-weekly and placebo (70% and 29% APOE4 non-carriers for 10 mg/kg bi-weekly and placebo, respectively).

The applicant acknowledged during the IND stage that the analysis plan did not control the type I error at the level of 0.05 two-sided or equivalently 0.025 one-sided, due to the Bayesian dose selection and interim analyses without any multiplicity correction. The simulation study suggested a one-sided type I error of approximately 0.10 or greater. Because of this and other uncertainties with the trial design, the study was to be considered exploratory and a learning trial.

The trial did not meet the success criteria at the interim analyses or at the final analysis. The final posterior probability for the primary endpoint ADCOM is 58.5%, which is smaller than the prespecified probability threshold of 80%, for having at least a difference from placebo greater than or equal to the clinical threshold. The applicant reported that the final posterior probability of superiority to placebo is 96.7%. This probability is smaller than the appropriate probability threshold which should be much larger than the nominal level of 97.5% because of needed adjustments for multiple interim analyses and multiple doses involved in the selection of the best performing dose.

Due to the failure of the primary endpoint and the large number of secondary endpoints, all secondary endpoints should be considered exploratory.

The PET substudy was a voluntary substudy, so the balance of baseline demographics and disease characteristics within the substudy may not be guaranteed and the substudy sample may not be representative of the randomized population.

There do not appear to be clinical effects in APOE4 non-carriers despite there being effects on SUVR in non-carriers comparable to those in APOE4 carriers. This seems to not align with the biomarker treatment effect being thought of as reasonably likely to predict a corresponding clinical treatment effect for all patients.

## 5.2 Collective Evidence

Collective evidence is not considered in this review since there was only one double-blind, controlled trial.

## 5.3 Conclusions and Recommendations

The applicant submitted Study 201 to evaluate lecanemab for treatment of Alzheimer's disease. In this phase 2 study, the applicant's primary endpoint was an Alzheimer's disease composite, ADCOMS. Study 201 had a Bayesian adaptive randomization design with multiple interim analyses. Randomization ratios were adapted to focus on the best performing dose based on the interim results.

Randomization of APOE4 carriers to the 10 mg/kg bi-weekly dose group, the best performing dose, was stopped in the middle of the study at the request of European regulators due to increased risk of Amyloid Related Imaging Abnormalities for carriers at this dose. This led to a large imbalance in APOE4 carrier status between 10 mg/kg bi-weekly and placebo (70% and 29% APOE4 non-carriers for 10 mg/kg bi-weekly and placebo, respectively).

As prespecified by the applicant, the trial was to be considered a success if either 1) early success was declared at an interim analysis or 2) the trial continued to completion and was declared a success at the 12 months Bayesian analysis. If the trial continued to completion, the trial was to be considered a success if there was at least an 80% probability that the best performing dose achieved a clinical threshold of a 25% difference from placebo on ADCOMS change from Baseline at Month 12.

The trial did not meet the success criteria at the interim analyses or at the final analysis. The final posterior probability indicates that there is a 58.5% probability of achieving a difference between placebo and lecanemab greater than or equal to the clinical threshold. This posterior probability does not meet the prespecified success probability of 80%.

The applicant specified several key secondary endpoints and several other secondary endpoints. Due to the failure of the primary endpoint and the large number of secondary endpoints, all secondary endpoints should be considered exploratory. Among the key secondary endpoints the applicant focused on the effect of lecanemab compared to placebo on brain amyloid pathophysiology at 18 months of treatment in subjects with Early Alzheimer's Disease as measured by amyloid positron emission tomography (PET). Eight hundred and fifty-four patients were randomized in the study, 315 (36.8%) of whom participated voluntarily in the PET substudy. The mixed model for repeated measures analyses of change from baseline in brain amyloid levels as measured by PET SUVr normalized to whole cerebellum mask at 18 months

estimated a reduction in percent change of 21.5 for 10 mg/kg biweekly compared to placebo (nominal p-value of <0.001).

The impact of amyloid reduction on the clinical outcome is uncertain. The Spearman correlation at the patient level is 0.128 between change in ADCOMS and change in amyloid PET SUVR at Month 18. This correlation is small. Furthermore, there is no apparent treatment effect on the clinical endpoint in APOE4 non-carriers despite the comparable amyloid reductions to APOE4 carriers (see Figure 6 and Table 14).

# 6 Appendix

**NDLM Dose-Response Model (Primary Analysis of Study 201)**

The 79-week outcomes were modelled as normally distributed,

$Y_{i,79} \sim N(\theta_d, \sigma^2)$

where d = 0 (placebo control), d=1 (2.5 mg/kg bi-weekly), d=2 (5 mg/kg bi-weekly), d=3 (10 mg/kg bi-weekly), d=4 (5 mg/kg monthly), and d=5 (10 mg/kg monthly).
The mean change from baseline to 79-weeks for dose group d was labeled as $\theta_d$.
A dose-response model was constructed for the mean change from baseline for each treatment arm.
The arms are modeled with a two-dimensional first-order normal dynamic linear model (NDLM). This model is a Gaussian random walk model.
For $\theta_0$ and $\theta_1$, the priors were pre-specified only for the ADCOMS endpoint. In order to account for differences in scale for the different cognitive endpoints, an empirical Bayes approach was used to adjust the scale for the other cognitive endpoints, scaling based on the standard deviation across all patients for the baseline visit.

The priors for $\theta_0$ and $\theta_1$

ADCOMS: $\theta_0 \sim N(0, 0.5^2)$; $\theta_1 \sim N(0, 0.5^2)$

and the NDLM structure for the bi-weekly doses:

$$\theta_2 \sim N(\theta_1, \tau^2)$$
$$\theta_3 \sim N(\theta_2, \tau^2)$$

The monthly doses are connected to the dose-response model through the respective bi-monthly doses. The connection through the NDLM first order structure results in the following priors

$$\theta_4 \sim N(\theta_2, \tau^2)$$

$$\theta_5 \sim N\left(.5(\theta_3 + \theta_4)\ , \tau^2\right)$$

The drift parameter (variance component) $\tau^2$ was pre-specified for the ADCOMS data. Since they were not prespecified for the other cognitive measures, an empirical Bayes approach was used for setting the priors. This approach scales the central value of $\tau^2$ to match the relative scaling of the ADCOMS scale parameter to the standard deviation across all patients for the baseline visit. These priors are as follows:

ADCOMS: $\tau^2 \sim IG(0.25, 0.0025)$

where IG(a,b) is the inverse gamma distribution with shape parameter a and scale parameter b. The distribution of the variance of the primary endpoint, $\sigma^2$, was pre- specified for the ADCOMS endpoint. Again, an empirical Bayes approach is used to specify the prior parameters for the other endpoints, adjusting the scale parameter standard deviation across all patients for the baseline visit:

1. ADCOMS: $\sigma^2 \sim IG(2.5, 0.056)$
2. CDR-SB: $\sigma^2 \sim IG(2.5, 4.90)$

*Longitudinal Modeling with Bayesian Linear Regression*

A linear regression model is created for the correlation between the 13-, 27-, 39-, 53-, 65- and 79-week value (we suppress the subject index *i* and refer to the change from baseline at weeks 13, 27, 39, 53, 65, and 79 as $Y_{13}$, $Y_{27}$, $Y_{39}$, $Y_{53}$, $Y_{65}$ and $Z_{79}$). A common model is used for each active dose and a separate model for the control.

Each of the model instances is identical, with the following structure for early time period j:

$$Z_{79}|Y_j \sim N(\alpha_j + \beta_j Y_j, \lambda^2)$$

$$\alpha_j \sim N(0, 0.05^2)$$

$$\beta_j \sim N(0.8, 0.1^2)$$

$$\lambda^2 \sim IG(2.5, 0.025)$$

The joint posterior distribution of $\alpha_j$, $\beta_j$, and $\lambda_j$, for j=13, 27, 39, 53, and 65 are based on the observed values of $Z_{79}$.

*Cognitive Progression Model*

For a patient i at a visit j, the ADCOMS score $Yij$ is modeled as

$$Y_{ij} = \mu_{g_i} + \gamma_i + exp(\theta_{ti}) \sum_{v=0}^{j} \alpha_{vg_i} + \epsilon_{ij}.$$
$$\epsilon_{ij} \sim N(0, \sigma^2)$$
$$\gamma_i \sim N(0, \sigma_\gamma)$$

*jj*

$Y_{ij} = \mu_{g_i} + \gamma_i + e(\theta_{ti}) \sum_{v=0 \text{ to } j} \alpha_{vg_i} + \epsilon_{ij}.)$

$\epsilon_{ij} \sim N(0, \sigma^2)$

$\gamma_i \sim N(0, \sigma_\gamma)$

In this model $\gamma_i$ is a random patient effect and $e(\theta_{ti})$ is the cognitive rate ratio (CRR) for treatment t where t = 1 is placebo control. The term $\mu_{gi}$ is the mean baseline score for group $gi$, and $\alpha_{vgi}$ is mean change in the ADCOMS score from visit v – 1 to visit v and is restricted to be monotonic. The $\epsilon_{ij}$ represents residual error.

Groups are defined by the four categories: g=1: MCI/APOE-, g=2: MCI/APOE+, g=3: Mild AD/APOE-, g=4: Mild AD/APOE+.

Priors are specified as follows.

For ADCOMS:

$\mu_g \sim N(0.3, 0.15^2)$, for $g$ = 1,2

$\mu_g \sim N(0.5, 0.15^2)$, for $g$ = 3,4

$\alpha_{v,g} \sim N(0.025, 0.52)$, for all $v, g$

$\sigma_2 \sim IG(10, 0.081)$

$\sigma_{\gamma 2} \sim IG(100, 1.44)$

---

**This is a representation of an electronic record that was signed electronically. Following this are manifestations of any and all electronic signatures for this electronic record.**

---

/s/

---

TRISTAN S MASSIE
12/19/2022 12:39:16 PM

HSIEN MING J HUNG
12/19/2022 01:05:00 PM