

Peripheral and Central Nervous System (PCNS) Drugs Advisory Committee Meeting
November 6, 2020

NDA/BLA# 761178

Drug name: Aducanumab

Applicant: Biogen

Combined FDA and Applicant PCNS Drugs Advisory Committee Briefing Document

DISCLAIMER STATEMENT

The attached package contains background information prepared by the Applicant and the Food and Drug Administration (FDA) for the panel members of the Advisory Committee. The package contains assessments and/or conclusions and recommendations written by individual FDA reviewers. Such conclusions and recommendations do not necessarily represent the final position of the individual reviewers, nor do they necessarily represent the final position of the Review Division or Office. We have brought the drug aducanumab to this Advisory Committee to gain the Committee's insights and opinions, and the background may not include all issues relevant to the final regulatory recommendation and instead is intended to focus on issues identified by the Agency for discussion by the Advisory Committee. The FDA will not issue a final determination on the issues at hand until input from the Advisory Committee process has been considered and all reviews have been finalized. The final determination may be affected by issues not discussed at the Advisory Committee meeting.

This document is available for public release without redaction.

TABLE OF CONTENTS

TABLE OF CONTENTS.....	2
LIST OF TABLES	4
LIST OF FIGURES	5
LIST OF ABBREVIATIONS	7
1. INTRODUCTION	9
1.1. Applicant Proposed Indication	9
1.2. Purpose of the Meeting	10
2. BACKGROUND	11
2.1. Overview of Alzheimer's Disease	11
2.2. Unmet Medical Need	12
2.3. Rationale for Aducanumab for the Treatment of Alzheimer's Disease	13
2.3.1. β -Amyloid Pathway	13
2.3.2. Mechanism of action of aducanumab and nonclinical data	13
2.4. Historical Context for the Development of Anti-A β Treatments for Alzheimer's Disease	14
2.4.1. Early anti-A β clinical programs.....	14
2.4.2. Molecular differences among anti-A β antibodies.....	15
2.4.3. Amyloid-related amyloid imaging abnormalities (ARIA)	15
2.4.4. Current/active anti-A β clinical programs	16
2.4.5. Advances in Alzheimer's disease research.....	16
2.4.6. Methodologic considerations in clinical trial design	17
2.4.7. Summary	18
2.5. Regulatory History.....	18
3. ADUCANUMAB CLINICAL TRIALS.....	21
3.1. Overview of the Clinical Development Plan for Aducanumab in Alzheimer's disease.....	21
3.1.1. Overview of Aducanumab Clinical Development.....	21
3.1.2. Study 103 Design.....	23
3.1.3. Study 301 and 302 Design.....	28
3.2. Early Termination of the Phase 3 Studies.....	34
3.2.1. Prespecified interim analysis and outcome.....	34

3.2.2.	Futility results (based on pooled December 2018 data) did not accurately reflect results of the individual studies	35
3.3.	Post-Futility	36
3.3.1.	Analysis of a larger dataset.....	36
3.3.2.	Biogen-FDA Collaborative Investigation.....	37
3.3.3.	Interpretability of Phase 3 data	37
3.3.4.	Conclusions.....	38
3.4.	Summary of Study 302	39
3.4.1.	Disposition, Demographics, and Baseline Characteristics	39
3.4.2.	Efficacy.....	41
3.4.3.	Clinical Pharmacology.....	54
3.4.4.	Conclusion	56
3.5.	Summary of Study 301	57
3.5.1.	Disposition, Demographics, and Baseline Disease Characteristics.....	57
3.5.2.	Efficacy	59
3.6.	Effect of Study 301 on the Persuasiveness of Study 302	62
3.6.1.	Introduction.....	62
3.6.2.	Hypotheses Tested	62
3.6.3.	Results.....	64
3.6.4.	Conclusions.....	84
3.7.	Summary of Study 103	85
3.7.1.	Disposition, Demographics, and Baseline Characteristics	85
3.7.2.	Efficacy.....	87
3.7.3.	Conclusion	89
3.8.	Efficacy Summary and Conclusions.....	91
4.	SAFETY	95
4.1.	Overall Extent of Exposure	96
4.2.	Summary of Adverse Events	110
4.3.	Summary of Adverse Events of Special Interest	114
4.3.1.	ARIA Incidence	119
4.3.2.	Characteristics of ARIA	122
4.4.	Summary of Other Safety Areas of Interest	129
4.5.	Safety Summary and Conclusions	130

5.	POINTS FOR THE ADVISORY COMMITTEE TO CONSIDER	132
6.	REFERENCES	133
7.	APPENDIX 1.....	140
8.	APPENDIX 2.....	246

LIST OF TABLES

Table 1:	Aducanumab Clinical Studies.....	22
Table 2:	Demographics and Baseline Characteristics – ITT Population	40
Table 3:	Primary Efficacy Endpoint Analysis: Change from Baseline on the CDR-SB at Week 78 in Study 302.....	42
Table 4:	Change from Baseline on CDR Items at Week 78 in Study 302.....	42
Table 5:	Change from Baseline in CDR-SB at Week 78 in Study 302: Uncensored Analysis and OTC Analysis	44
Table 6:	Change from Baseline to Week 78 on the CDR-SB, Primary Analysis and Prespecified and Post Hoc Sensitivity Analyses in Study 302, ITT Population	46
Table 7:	Primary and Secondary Endpoints at Week 78 in Study 302: ITT Population	47
Table 8:	Primary, Secondary, and Tertiary Endpoints at Week 78 in Study 302: ITT Population	48
Table 9:	Demographics and Baseline Characteristics – ITT Population	58
Table 10:	Primary and Secondary Endpoints at Week 78 in Study 301: ITT Population	60
Table 11:	Overview of Hypotheses and Outcomes for Why Results for the High-Dose Groups Were Discordant Between Studies 301 and 302.....	65
Table 12:	Change from Baseline in CDR-SB at Week 78: All Participants With and Without Post-ARIA Observations Excluded, Studies 301 and 302, ITT Dataset	66
Table 13:	Analyses to Investigate Efficacy Outcomes in the High-Dose Group by Dosing Subgroups.....	77
Table 14:	Demographics and Baseline Characteristics—Placebo-Controlled Period—Study 103	86
Table 15:	CDR-SB and MMSE Change from baseline at 12 Months—Placebo-Controlled Period--Study 103.....	88
Table 16:	Study 302 (18 Months) and Study 103 (12 Months): Change from baseline on the CDR-SB, MMSE, ADAS-Cog13, ADCS-ADL-MCI, and NPI-10.....	92

Table 17:	Studies 301 and 302 (18 Months) and Study 103 (12 Months): Change from Baseline in Biomarkers.....	93
Table 18:	Overall Extent of Exposure – Pool A1	98
Table 19:	Overall Extent of Exposure – Pool A2	103
Table 20:	Overall Summary of Exposure – Safety Population – Aducanumab Summary of Clinical Safety	109
Table 21:	Overall Summary of Adverse Events – Pool A1	112
Table 22:	Adverse Events With at Least 5% Incidence in Aducanumab 10 mg/kg and 2% Higher Incidence Than Placebo – Pool A1	113
Table 23:	Severity of ARIA-E and ARIA-H on MRI.....	117
Table 24:	ARIA Management by Protocol Version - Studies 301 and 302.....	118
Table 25:	ARIA Events Based on AE eCRF - Pool A1 Safety MRI Population.....	120
Table 26:	Summary of ARIA-H Events Based on Adverse Event eCRF by ARIA-E Status – Pool A1 Safety MRI Population	121

LIST OF FIGURES

Figure 1:	Description of the β -amyloid pathway	13
Figure 2:	Evolution in Nomenclature for Categorization of Alzheimer's Disease Across the Alzheimer's Disease Continuum (Since 2010).....	17
Figure 3:	Study 103 Placebo-Controlled Period Schematic.....	26
Figure 4:	Schematic of Study Design including Aducanumab Doses (Pre- and Post-Protocol Version 4).....	33
Figure 5:	Change From Baseline on the CDR-SB Over Time in Study 302	43
Figure 6:	Changes in Brain Amyloid Plaque, Measured by PET and Quantified as SUVR in Study 302	49
Figure 7:	CSF p-Tau and t-Tau Change at Week 78 in Study 302, CSF Modified Analysis Population	50
Figure 8:	Change from baseline in Tau PET in Studies 301 and 302 (Pooled Data): Frontal, Temporal, and Medial Temporal Composite Regions; Association between Tau PET Change in Medial Temporal Composite Region and Dose.....	51
Figure 9:	Subgroup Analyses (Disease Characteristics) for Clinical Endpoints, Study 302	52
Figure 10:	Subgroup Analyses (Demographics) for Clinical Endpoints, Study 302	53
Figure 11:	Population Mean Simulations of Nominal Brain Amyloid PET-Time Profiles by Dose (3, 6, and 10 mg/kg) – Study 302	55

Figure 12: Effect of Dose Interruption due to ARIA on Nominal Brain Amyloid-Time Profile – Study 302	55
Figure 13: Change From Baseline on the CDR-SB Over Time, Study 301.....	61
Figure 14: Treatment Difference for Change on CDR-SB at Week 78 by ApoE ϵ 4 Status with and Without Post-ARIA Observations	68
Figure 15: CDR-SB Results With and Without Rapid Progressors, Studies 301 and 302	70
Figure 16: Individual Dosing Profiles for High-Dose Participants in Studies 301 and 302	74
Figure 17: Median Number of 10 mg/kg Doses by Enrollment Time	75
Figure 18: Dosing Profiles (Heat Maps) for Zero, Intermediate, and High/Consistent Dosing Subgroups Within the High-Dose Groups of Studies 301 and 302	76
Figure 19: Heterogeneity in Dosing Subsets of the High-Dose Group by Mean Cumulative Dose, Studies 301 and 302	78
Figure 20: Comparison of CDR-SB in Subsets with Opportunity to Receive 14 Doses of 10 mg/kg – OTC Population.....	79
Figure 21: Consistent CDR-SB Outcomes in Participants Receiving Sufficient Doses of 10 mg/kg	81
Figure 22: Consistent CDR-SB Outcomes in Participants Receiving Sufficient Doses of 10 mg/kg – Excluding Rapid Progressors	81
Figure 23: Brain amyloid PET Change from Baseline at 6 and 12 Months—Placebo-Controlled Period—Study 103	89
Figure 24: Pooling Strategy for the Integrated Safety Analysis.....	96
Figure 25: MRI Monitoring Frequency in the Placebo-Controlled Treatment Period of the Phase 3 Studies 301 and 302.....	115
Figure 26: Kaplan-Meier Plot of Time to First ARIA-E Event – Pool A1 Safety MRI Population Aducanumab 10 mg/kg	123
Figure 27: Kaplan-Meier Plot of Time to First ARIA-E Event in Pool A2 For Those in the Pool A1 Aducanumab 10 mg/kg Group – Pool A2 Safety MRI Population	124
Figure 28: Kaplan-Meier Plot of Time to First ARIA-E Event by ApoE Status – Pool A1 Safety MRI Population Aducanumab 10 mg/kg ApoE ϵ 4-	125
Figure 29: Kaplan-Meier Plot of Time to First ARIA-E Event by ApoE Status – Pool A1 Safety MRI Population Aducanumab 10 mg/kg ApoE ϵ 4+	126

LIST OF ABBREVIATIONS

A β	amyloid beta
ADAS-Cog 13	Alzheimer's Disease Assessment Scale – Cognitive 13-Item Scale
ADCS-ADL-MCI	Alzheimer's Disease Cooperative Study – Activities of Daily Living – Mild Cognitive Impairment
AE	adverse event
ANCOVA	analysis of covariance
anti-A β	anti-amyloid beta (antibody)
APP	amyloid precursor protein
ApoE	apolipoprotein E
ApoE ϵ 4	apolipoprotein E ϵ 4
ARIA	amyloid-related imaging abnormalities
ARIA-E	amyloid-related imaging abnormalities – edema
ARIA-H	amyloid-related imaging abnormalities – hemorrhage or superficial siderosis
β -amyloid	peptide derived from membrane bound amyloid precursor protein
β -amyloid ₁₋₄₂	42-amino acid form of β -amyloid
BLA	Biologics Licensing Application
CDR	Clinical Dementia Rating scale
CDR-SB	Clinical Dementia Rating-Sum of Boxes
CNS	central nervous system
CSF	cerebrospinal fluid
C-SSRS	Columbia Suicide Severity Rating Scale
ECG	electrocardiogram
eCRF	electronic case report form
FDA	Food and Drug Administration
IgG1	immunoglobulin gamma 1
ITT	intent-to-treat
IV	intravenous(ly)
IWG	International Working Group
LTE	long-term extension
MCI	mild cognitive impairment

MedDRA	Medical Dictionary for Regulatory Activities
MMRM	mixed model repeated measures
MMSE	Mini-Mental State Examination
MRI	magnetic resonance imaging
NPI-10	Neuropsychiatric Inventory-10
OTC	opportunity to complete (Week 78)
PET	positron emission tomography
PK	pharmacokinetic(s)
PT	preferred term
p-Tau	phosphorylated tau
SAE	serious adverse event
SD	standard deviation
SE	standard error
SOC	system organ class
SPA	Special Protocol Assessment
SUVr	standard uptake value ratio
t-Tau	total tau
US	United States

1. INTRODUCTION

This briefing document presents results from three studies in the aducanumab Alzheimer's disease clinical development program, Studies 302, 301, and 103, along with discussion and analyses of the findings from these studies. In general, this document includes the Applicant's position followed by the FDA's position, to reduce redundancy and improve readability.

1.1. Applicant Proposed Indication

Proposed indication: Aducanumab is an amyloid beta targeting antibody indicated to delay clinical decline in patients with Alzheimer's disease.

1.2. Purpose of the Meeting

The FDA's Position:

The purpose of this meeting is to discuss the data submitted by the Applicant, with endorsement by the FDA of such submission, intended to establish the effectiveness of aducanumab, and to provide an overview of the safety profile of aducanumab.

Despite intense basic and clinical research and the existence of several approved therapies, there is an enormous unmet medical need for effective treatments for Alzheimer's disease, especially treatments that are intended to address the biological basis of the disease with a goal of favorably altering its long-term course. Currently approved treatments do not target the underlying pathology of Alzheimer's disease and their beneficial effects are modest and transitory. Furthermore, there are no treatments explicitly approved for the relatively earlier stage of Alzheimer's disease included in the aducanumab clinical development program. There has not been an approval of a novel medication for the treatment of Alzheimer's disease since 2004. Aducanumab targets amyloid beta, a fundamental pathological hallmark of the disease. Although there have been previous failures of other drugs that have been intended to target amyloid beta in some fashion, there are features of aducanumab's pharmacologic profile and of the design of its clinical development program that are novel and distinguish it from prior efforts with these other agents.

After promising early clinical and biomarker data emerged from the early phase Study 103, the Applicant embarked upon two trials of essentially identical design, Study 301 and Study 302, that were intended to establish the effectiveness of aducanumab. These studies were initiated in 2015. After conducting a prespecified interim analysis for futility in early 2019, the Applicant terminated Studies 301 and 302 and made a public announcement to this effect on March 21, 2019. Subsequent examination of individual study results that included additional data that had accrued during the time the futility analysis was being conducted revealed findings that differed from the results of the prespecified futility analysis, most notably including apparently positive results in Study 302. The Applicant promptly brought these results to the FDA for discussion and advice on their appropriate interpretation. After an initial consideration of these findings, the FDA recognized that additional work was necessary to achieve a maximum understanding of the results and established a collaborative plan for further rigorous analyses of the data. These analyses ultimately led to the FDA advising the Applicant that submission of a marketing application seeking approval of aducanumab was reasonable.

The notably long duration and thorough nature of pre-submission review affords a particularly complete consideration by the FDA of the evidence of effectiveness to be discussed at this meeting. The evidence presented by the Applicant in the application in support of aducanumab's effectiveness is essentially unchanged from that which has been considered throughout the pre-submission phase.

The FDA seeks input from the committee on whether aducanumab is an effective treatment for Alzheimer's disease.

2. BACKGROUND

2.1. Overview of Alzheimer's Disease

The Applicant's Position

Alzheimer's disease is an irreversible, progressive neurodegenerative disorder that slowly destroys memory and thinking skills. Initial impairment in memory may be followed by behavioral and neuropsychiatric symptoms, and finally a person's ability to perform usual daily life activities. Alzheimer's disease is the most common cause of dementia among older adults and is ultimately fatal.

In the US, more than 5.8 million people are living with Alzheimer's disease. By 2050, this number is projected to more than double [Alzheimer's Association 2020]. Alzheimer's disease was the sixth leading cause of death in the US and the fifth leading cause for people aged 65 years and older in 2018 [Alzheimer's Association 2020].

Life expectancy after an Alzheimer's disease diagnosis depends on various factors, including age at onset and severity at diagnosis [Brodaty 2012; Guehne 2007; Guehne 2005; Xie 2008]. An important distinction for Alzheimer's disease patients is that a significant period of their remaining lives may be spent in the severe disabling disease state, adding to the burden on the health care system [Rizzuto 2012].

In addition to the effect on patients, Alzheimer's disease places a burden on families and caregivers. Informal caregiving for patients with Alzheimer's disease or dementia has been estimated at 18 billion hours per year in the US, valued at \$230 billion annually [Alzheimer's Association 2020; Dunbar 2018]. Increased care demands result in increased financial, psychological, and physical stress for the caregiver [Alzheimer's Association 2020; Suehs 2014].

Alzheimer's disease is a pathophysiological and clinical continuum where specific pathologic changes in the brain that occur as a continuous process throughout the course of the disease precede by many years the eventual emergence of clinical symptoms of cognitive decline that also progress continuously over a long period [Jack 2018].

Alzheimer's disease is defined biologically by the presence of two abnormal protein deposits: extracellular deposits of brain amyloid plaques (comprising β -amyloid peptides) and neurofibrillary tangles (comprising abnormal tau protein) [Hyman 2012].

Deposition of β -amyloid peptides into amyloid plaques begins decades prior to observable clinical symptoms [Vermunt 2019]. Deposition of β -amyloid is followed sequentially by markers of neurodegeneration [Hardy and Selkoe 2002], accumulation of tau pathology [Hanseeuw 2019], and brain volume loss [Jonsson 2012], all of which initiate prior to the onset of clinical symptoms. This pre-symptomatic phase of Alzheimer's disease precedes the emergence of clinical symptoms by 10 to 20 years [Villemagne 2013]. At this time of the emergence of the earliest clinical findings, Alzheimer's disease has already been present for decades, and there is already an advanced pathologic disease state in the brain.

The most salient known risk factors for Alzheimer's disease are the unmodifiable contributors of older age, genetics, and family history. Of these, increasing age has the largest known impact on risk of developing Alzheimer's disease. While several genes have been found to increase the risk of Alzheimer's disease, the $\epsilon 4$ allele of the apolipoprotein E (ApoE) gene is the strongest known

genetic risk factor [Elias-Sonnenschein 2011; Mattsson 2018]. Compared with the most common ApoE genotype of $\epsilon 3/\epsilon 3$, $\epsilon 4$ heterozygosity increases risk of Alzheimer's disease by 3 to 4 times, and $\epsilon 4$ homozygosity increases risk by 8 to 12 times [Alzheimer's Association 2020]. Approximately two-thirds of pathology-confirmed Alzheimer's disease cases are $\epsilon 4$ positive (homozygous or heterozygous), compared with about 15% to 20% of the general population [Mattsson 2018]. Autosomal dominant genetic mutations are estimated to account for less than 1% of Alzheimer's disease cases [Bekris 2010].

Developing treatments to halt, slow, or reverse the course of disease is a goal for the scientific community, the clinical health care system, long-term care services, community/support environments serving patients and caregivers, and the federal government [U.S. Department of Health and Human Services 2019].

With pathophysiological changes in the brain starting a decade or more before clinical symptom onset, treatments directed at this goal must begin prior to the development of irreversible neurodegeneration and well before there are advanced clinical symptoms [FDA 2018; U.S. Department of Health and Human Services 2019; Vermunt 2019].

The FDA's Position:

The Applicant's overview accurately discusses the contemporary understanding of the pathophysiology and evolution of Alzheimer's disease.

FDA considers it critical to intervene in Alzheimer's disease, and neurodegenerative diseases in general, as early as possible given the complexity and possible irreversibility of pathophysiological deterioration responsible for clinical findings.

2.2. Unmet Medical Need

The Applicant's Position

Currently approved Alzheimer's disease treatments include the cholinesterase inhibitors donepezil, rivastigmine, and galantamine, and the N-methyl-D-aspartate antagonist memantine.

None of these agents address the underlying pathology of the disease. Their effects are reversible and lessen over time due to the continued progression of the disease process [Birks 2006; McShane 2006].

There are no treatments available to halt, slow, or cure Alzheimer's disease. Clearly, there is an unmet need for therapies that halt or slow the disease.

The FDA's Position:

The FDA agrees there is an urgent and unmet medical need for effective treatments for Alzheimer's disease. In addition to the general need for more effective treatments, there is a particular unmet need for effective treatments to delay, halt, or reverse the pathophysiological processes that ultimately lead to the clinical deficits of Alzheimer's disease.

2.3. Rationale for Aducanumab for the Treatment of Alzheimer's Disease

The Applicant's Position

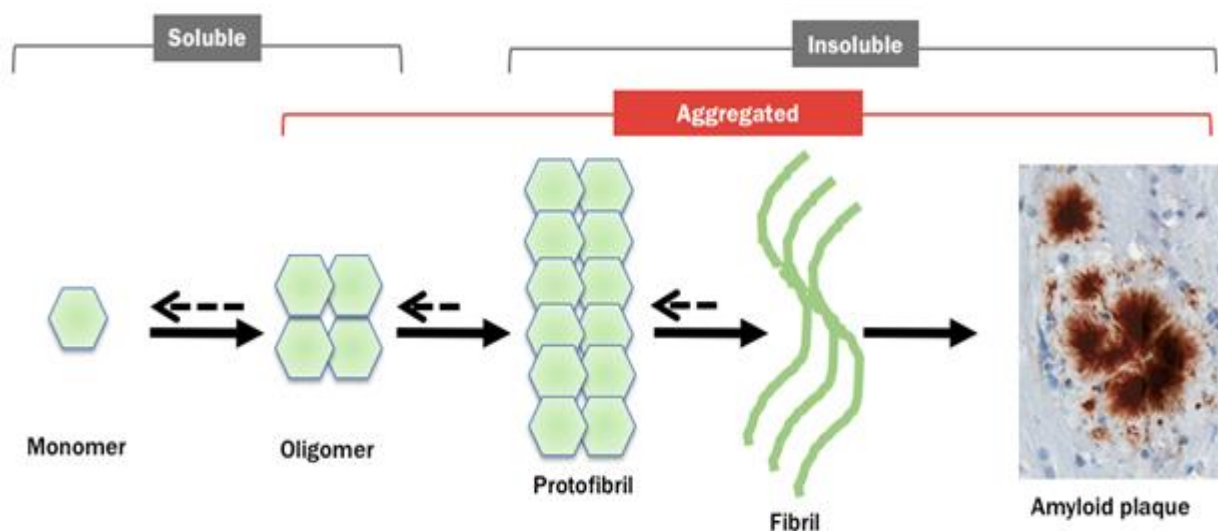
2.3.1. β -Amyloid Pathway

β -amyloid is a peptide generated by sequential cleavage of its precursor protein (APP) [Kang 1987; Selkoe 1991]. Monomers of β -amyloid can be of differing lengths and can self-aggregate, forming a range of aggregates from small soluble oligomers to larger aggregated forms which ultimately deposit in the brain (Figure 1).

Aggregated forms of β -amyloid, including soluble oligomers and amyloid plaques, are directly toxic to synapses and neurons [Benilova 2012; Koffie 2009], while monomers are not considered to exhibit these detrimental properties.

Genetic, neuropathological, and cell biological evidence suggest that reducing or removing $A\beta$ could be beneficial for patients with Alzheimer's disease [Ising 2015; Selkoe 2013].

Figure 1: Description of the β -amyloid pathway



Source: Image adapted from [Huang 2013].

2.3.2. Mechanism of action of aducanumab and nonclinical data

Unlike any other anti- $A\beta$ monoclonal antibody in development, aducanumab was derived from human B-cells collected from healthy elderly subjects with no signs of cognitive impairment and from cognitively impaired elderly subjects with unusually slow clinical decline. The screening of libraries of human memory B cells for reactivity against aggregated $A\beta$ led to molecular cloning, sequencing, and recombinant expression of aducanumab [Sevigny 2016].

Aducanumab is a human IgG1 anti- $A\beta$ monoclonal antibody selective for $A\beta$ aggregates, including soluble oligomers and insoluble fibrils, but not monomers, as demonstrated in a variety of biochemical and structural analyses.

Binding of aducanumab to aggregated β -amyloid promotes the removal of amyloid from the brain, through a microglia-mediated phagocytosis mechanism. Experimental evidence has shown that removal of amyloid stops or reduces neurotoxicity [Rozkalne 2009; Spires-Jones 2009] that might otherwise lead to neurodegeneration and, ultimately, cognitive impairment.

Nonclinical studies have confirmed in vivo dose-dependent target engagement and dose-dependent reduction in brain amyloid burden for aducanumab [Sevigny 2016].

The FDA's Position:

FDA acknowledges the Applicant's rationale for the development of aducanumab in Alzheimer's disease and notes the amyloid pathway has been widely considered a target for drug development.

FDA notes the unique provenance of aducanumab. Its origin, based upon reactivity to aggregated amyloid beta in humans who appear phenotypically resistant to cognitive deterioration, is unique amongst agents targeting amyloid beta.

2.4. Historical Context for the Development of Anti-A β Treatments for Alzheimer's Disease

The Applicant's Position

2.4.1. Early anti-A β clinical programs

Multiple investigational products have been developed to either reduce production of β -amyloid or lower levels of aggregated β -amyloid present in the brain. Enhancing elimination of β -amyloid by immunotherapy has been one of the most pursued approaches for the development of a treatment for Alzheimer's disease. However, several development programs have been terminated due to either lack of efficacy and/or safety considerations.

The failures of these previous anti-A β programs are recognized to be due in part to the following clinical development considerations:

- Inclusion of individuals in those clinical trials without evidence of brain β -amyloid pathology (i.e., individuals without Alzheimer's disease) [Selkoe and Hardy 2016].
- Unknown or no target engagement prior to initiation of Phase 3 (i.e., inadequate effect on β -amyloid in the brain or CSF) [Karran and Hardy 2014].
- Use of dosages too low to elicit sufficient target engagement, effect on pathophysiology, or clinical progression [Selkoe and Hardy 2016].
- Inclusion of participants at later stages of Alzheimer's disease dementia, when significant irreversible neurodegeneration has already occurred [Musiek and Holtzman 2015; Selkoe and Hardy 2016].

None of the failed anti-A β antibodies achieved a clinical proof of concept, i.e., effect on clinical progression prior to initiating Phase 3 studies. Aducanumab was the first program to achieve proof of concept prior to Phase 3, as will be discussed in Section 3.1.2.

2.4.2. Molecular differences among anti-A β antibodies

In addition to the clinical development considerations discussed above, anti-A β antibodies differ considerably with respect to their molecular characteristics including antibody isotype, binding epitopes and biophysical properties. Anti-A β antibodies therefore differ in specificity for and mechanisms of action toward the different forms of β -amyloid. This directly affects their ability to target and remove the toxic forms of β -amyloid and, consequently, the potential for efficacy.

Of the anti-A β antibodies previously in development for symptomatic Alzheimer's disease, antibodies that bound to monomeric forms of β -amyloid appeared to influence directly the clearance of β -amyloid from the brain into the plasma [DeMattos 2001]. However, binding to β -amyloid monomers can also lead to an unproductive mobilization of antibody bound to a large pool of monomers found outside of the brain, and might not represent a desirable feature if one considers the optimal profile for an anti-A β antibody [Seubert 2008].

Several mechanisms of action have been described for the anti-A β antibodies targeting aggregated forms of β -amyloid that are currently in development for symptomatic Alzheimer's disease. Those mechanisms can affect either the β -amyloid aggregation process and ultimately the ability of these anti-A β antibodies to modulate the levels of neurotoxic oligomers, and/or the clearance of β -amyloid aggregates

Nonclinical studies have shown that aducanumab, similar to other anti-A β antibodies targeting aggregated forms of β -amyloid such as gantenerumab and bapineuzumab, induces clearance of brain amyloid plaques through a microglia-mediated phagocytosis process [Sevigny 2016]. Additionally, aducanumab due to its particular affinity and binding stoichiometry toward β -amyloid aggregates uniquely among anti-A β antibodies directly inhibits the molecular process through which oligomers form (secondary nucleation), thereby reducing the formation of neurotoxic A β oligomers [Linse 2020].

2.4.3. Amyloid-related amyloid imaging abnormalities (ARIA)

Another feature that relates to the molecular characteristics of anti-A β antibodies is amyloid-related imaging abnormalities (ARIA). ARIA, including ARIA-E (previously referred to as vasogenic edema), was observed in the very first clinical trials as a treatment-related side effect in patients receiving an anti-A β monoclonal antibody [Ostrowitzki 2012; Sperling 2011b]. ARIA has since been observed with some but not all other anti-A β antibodies, particularly those targeting aggregated and/or deposited β -amyloid [Greenberg 2020; Penninkilampi 2017]. Consistent findings are that ARIA occurs early during treatment, is typically asymptomatic, and is more frequent in ApoE ϵ 4 carriers than noncarriers especially at higher doses [Greenberg 2020; Ketter 2017; Salloway 2014; Sperling 2012; Sperling 2011b]. The biological mechanism of ARIA-E is not well understood. Published hypotheses suggest that increased cerebrovascular permeability could be caused either by increased A β clearance from the brain parenchyma, leading to saturation of the perivascular drainage system, and/or by direct antibody interaction with deposited vascular amyloid, leading to its clearance and weakening of the vessel walls [Greenberg 2020; Zago 2013].

ARIA is discussed in more detail in Section 4.

2.4.4. Current/active anti-A β clinical programs

The anti-A β antibodies currently in development for Alzheimer's disease specifically target and remove the neurotoxic aggregated forms of β -amyloid (e.g., aducanumab, BAN2401, and gantenerumab). Each antibody also has effector-function enabling microglial-mediated clearance of aggregated β -amyloid.

With these properties, and in contrast to the first generation of anti-A β antibodies, these agents have each demonstrated robust and sustained reductions in brain amyloid in early symptomatic Alzheimer's disease participants as measured by PET imaging [Klein 2018; Sevigny 2016; Swanson 2018].

Aducanumab and BAN2401 have also shown a clinical proof of concept prior to initiating Phase 3 studies [Swanson 2018]. Aducanumab was, with Study 103, the first of this generation of antibodies to demonstrate robust reductions in brain amyloid in clinical trials, and to have shown a reduction in clinical decline, i.e., the first anti-A β program to establish proof of concept before initiating Phase 3 (results detailed in Section 3.1.2). Accordingly, aducanumab is the first program of this generation of anti-A β antibodies to have data from Phase 3 trials (results detailed in Sections 3.4 and 3.5).

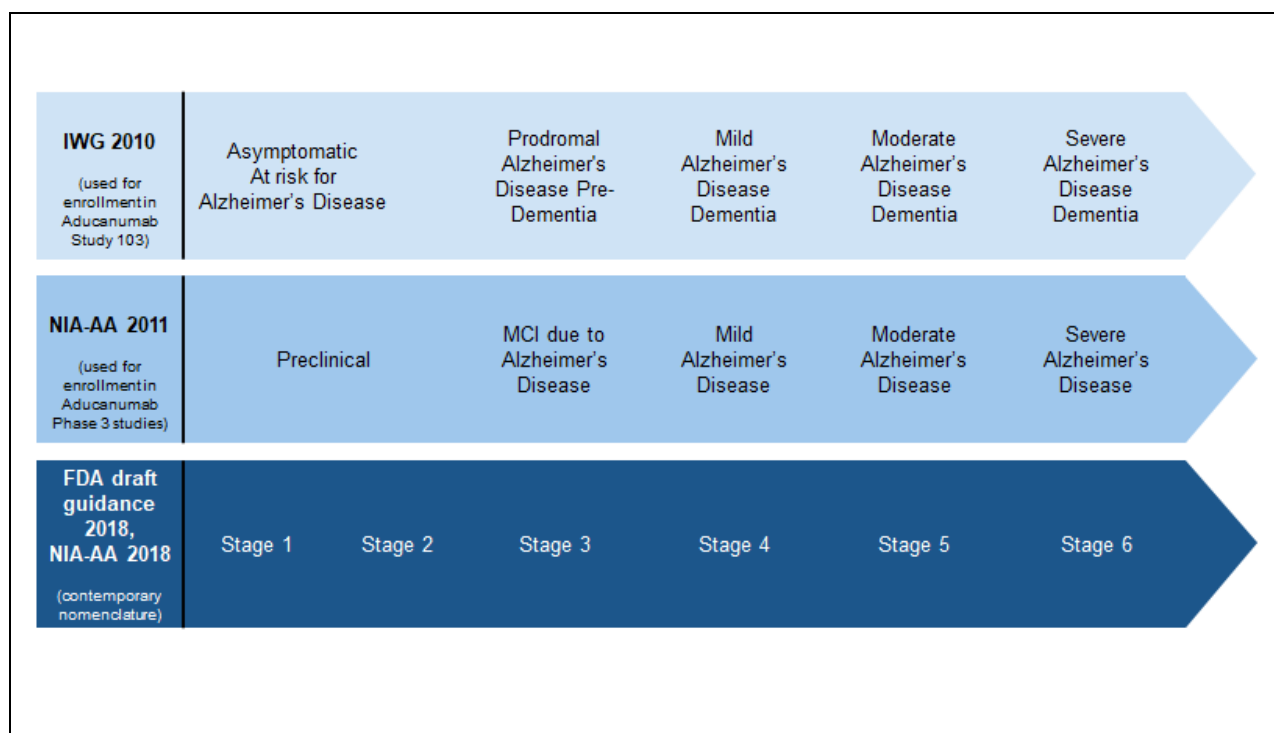
2.4.5. Advances in Alzheimer's disease research

Along with the key observations from the previous anti-A β clinical trials, advances have been made by the research community in understanding Alzheimer's disease [Weiner 2017]. It is established that the disease is a continuum, and that the deposition of β -amyloid into amyloid plaques begins decades prior to the onset of symptoms [Jack 2018].

For treatments targeting amyloid pathology these findings led to a shift to conducting clinical trials in participants at earlier stages of disease.

Nomenclature for diagnostic guidelines has kept pace with the scientific progress in Alzheimer's disease. Those germane to the 10-year time frame of the aducanumab clinical development program are listed in Figure 2.

Figure 2: Evolution in Nomenclature for Categorization of Alzheimer's Disease Across the Alzheimer's Disease Continuum (Since 2010)



Abbreviations: IWG = International Working Group; MCI = mild cognitive impairment; NIA-AA = National Institute on Aging and the Alzheimer's Association.

Data sources: [\[Albert 2011; Dubois 2010; FDA 2018; Jack 2018; McKhann 2011; Sperling 2011a\]](#).

2.4.6. Methodologic considerations in clinical trial design

The conduct of clinical trials in participants at early stages of symptomatic disease comes with additional methodological considerations.

Among them are developing reliable methods for identifying participants with early symptomatic disease. This includes developing methods to detect the underlying core pathology of Alzheimer's disease present early in the disease process (β -amyloid), because symptoms at this stage are subtle and may be due to other dementias (e.g., vascular dementia) or other non-neurodegenerative diseases (e.g., depression). Methods approved in the US at the time the Phase 3 protocols were being developed include the use of the amyloid PET tracers [^{18}F]-florbetapir, [^{18}F]-flutemetamol and [^{18}F]-florbetaben [\[Amyvid USPI 2012; Neuraceq USPI 2014; Vizamyl USPI 2013\]](#). In addition, there are several CSF tests for β -amyloid confirmation that are currently in development in the US. These CSF tests appear to show similar sensitivity and specificity ranges to those of the amyloid PET tracers and are routinely used in the enrollment of participants in contemporary Alzheimer's disease studies.

In addition, appropriate clinical outcome measures must be identified that are sensitive to change at this disease stage.

Finally, designing and powering clinical trials requires accurate estimates of the rate of decline on these measures in untreated patients. These design considerations and how they were each addressed in the aducanumab clinical trials are discussed in Sections [3.1.2](#) and [3.1.3](#).

2.4.7. Summary

The following are key points from the preceding discussion:

- The failure of previous anti-A β programs to demonstrate efficacy is likely attributable to both the molecular characteristics of the drug and clinical trial design features
- Anti-A β antibodies cannot be considered as a single class. They are distinct at the molecular level, and the differences have an impact on their mechanism of action, including:
 - Differences in binding characteristics (i.e., which form or forms of β -amyloid that they bind to)
 - Whether they reduce amyloid plaque
 - Whether they are associated with ARIA
- With the increased understanding of Alzheimer's disease, there has been a shift in the focus of the scientific community to conducting clinical trials of anti-A β agents in patients at earlier stages of the disease.
- Clinical research in patients at the earlier stages of disease poses additional considerations for design and conduct of clinical trials.
- Among several anti-A β antibodies currently in development for Alzheimer's disease, aducanumab was the first of this generation to have Phase 3 clinical data.

The FDA's Position:

The FDA agrees that anti-A β therapies do not represent a single class of drugs. As noted, aducanumab has molecular properties that in part distinguish it from the anti-A β therapies that have been previously tested in clinical trials. Aducanumab may be more appropriately considered in the context of later generation therapies which have demonstrated target engagement and brain amyloid reduction at sufficiently high drug levels. The design of the pivotal trials for aducanumab benefited from the lessons learned from previous trials regarding patient selection, outcome assessment, and dosing. Therefore, previous late-stage failures of anti-A β therapies do not constitute a demonstrated "class failure" and are not particularly informative for the assessment of the effectiveness of aducanumab.

2.5. Regulatory History

The Applicant's Position

Biogen submitted to FDA a Biologics License Application (BLA) in July 2020 for approval to market aducanumab.

Applications are being prepared for review in other regions.

Biogen sought guidance from FDA prior to and during the development of aducanumab as a treatment for Alzheimer's disease. In 2015, Special Protocol Assessment (SPA) agreements were sought and obtained from the FDA for the Phase 3 studies. FDA agreed that the design and planned analysis of each Phase 3 study addressed the objectives necessary to support a regulatory submission. This included use of the CDR-SB as the primary efficacy endpoint.

On March 21, 2019, Biogen announced the termination of the Phase 3 program (Studies 301 and 302), based on results of a prespecified interim futility analysis with a data cutoff date of December 26, 2018. As discussed further in Section 3.3, subsequent individual study results of analyses based on data through March 20, 2019 differed from the results of the prespecified futility analysis, with one study showing apparently positive results (Study 302). At a Type C meeting in June 2019, Biogen sought guidance from the FDA on the appropriateness and interpretation of the latter set of results. Subsequently, over a period of approximately 1 year, Biogen engaged with the FDA in an ongoing bilateral scientific effort to understand the similarities and differences between the two Phase 3 studies (Studies 301 and 302). Biogen and FDA met at three further Type C meetings to discuss the ongoing scientific effort (October 2019, February 2020, and June 2020); the last of these also addressed Type B pre-BLA meeting topics to support submission of a BLA.

June 2019:

At this meeting, the FDA concluded in the minutes that “it would have been more appropriate if futility had not been declared for those studies.” After noting that “the effect of early termination of the studies on the interpretability of the observed efficacy data and associated analyses is a matter for further detailed consideration,” the FDA further noted, “on face, that the effects of aducanumab in that [302] study might not only be interpreted as being supportive of the efficacy of that compound in Alzheimer's disease, but might also be considered exceptionally persuasive on several of the instruments used to evaluate efficacy.”

The FDA noted, “Further complicating the interpretation of the available data for Studies 301 and 302 are the partially conflicting results ... for Study 301 as compared with those for Study 302, with particular attention to the discordant high dose results of each study (while noting an apparent degree of consistency of the low-dose results between the studies). A detailed understanding, informed by plans for further analyses ..., of the overall results, and especially these discordant results, is critical to any consideration of whether Study 302 (with or without possible support from Study 301, as might be determined from further explorations of the data) might provide evidence adequate to establish the effectiveness of aducanumab for the treatment of Alzheimer's disease.”

The FDA further stated that “the submission of a marketing application for aducanumab based primarily on the results of Study 302 as a single positive efficacy study may also be considered. It is possible that the results of Study 301 may have a role in supporting the results of Study 302 or may be understood well enough to be dismissible (i.e., to not represent evidence that the drug is ineffective), assuming that further analyses do not lead to a conclusion that Study 301 is clearly negative.”

After citing the possibility that aducanumab could be an effective drug for the treatment of Alzheimer's disease based on the Study 302 data presented, the FDA stated, “It is imperative that extensive resources be brought to bear on achieving a maximum understanding of the

existing data. Given the wholly unique situation that is the current state of the aducanumab development program ..., those further analyses would best be conducted as part of a bilateral effort involving the Agency and sponsor, i.e., through a 'workstream' or a 'working group' collaboration."

October 2019:

The FDA indicated that the analyses conducted since the June 2019 Type C meeting had established that the Phase 3 results were interpretable. The minutes state, "The analyses conducted since the June 14, 2019, Type C meeting, have established not only that the results of Studies 301 and 302 are interpretable, but on face, suggest an understanding of the discordant results of Studies 301 and 302 sufficient to allow for independent consideration of whether Study 302 might provide evidence adequate to establish the effectiveness of aducanumab for the treatment of Alzheimer's disease."

February 2020:

Additional analyses to further characterize the dose-exposure-response relationship were presented and discussed. With additional analyses complete, the minutes stated, "The additional analyses presented in this meeting package address the questions that remained following the October 21, 2019, Type C meeting and provide additional insight regarding the relationship between aducanumab dose/exposure and response in Study 301 and Study 302."

June 2020:

Biogen summarized for the FDA its proposal to submit a BLA, with Study 302 providing the primary contribution to substantial evidence of the effectiveness of aducanumab, with reasons for differences between the Phase 3 studies understood and Study 301 subgroups providing support for Studies 302, and with Study 103 providing an additional contribution to substantial evidence of aducanumab's effectiveness. While a matter of review, FDA indicated that Biogen's plan to submit a BLA was reasonable.

The FDA's Position:

The Applicant provides an accurate summary of the regulatory history. The first formal request from Biogen to the FDA regarding a discussion of the decision to terminate the aducanumab Phase 3 studies and the resultant analyses of the data from those studies came on May 15, 2019, in the form of a Type C Meeting Request. Given the unmet medical need and unique nature of the data, FDA and the Applicant agreed upon a plan for further analyses and FDA formally engaged with the Applicant through Type C Meetings on four occasions between June 2019 and June 2020, the last of which included a discussion of pre-BLA questions submitted by the Applicant.

3. ADUCANUMAB CLINICAL TRIALS

3.1. Overview of the Clinical Development Plan for Aducanumab in Alzheimer's disease

3.1.1. Overview of Aducanumab Clinical Development

The Applicant's Position

The aducanumab clinical program comprises 7 clinical studies in Alzheimer's disease and 1 study in healthy volunteers ([Table 1](#)).

Aducanumab was administered by IV infusion in all Alzheimer's disease studies and was administered every 4 weeks in all multiple-dose studies.

Table 1: Aducanumab Clinical Studies

Study	Design / Dosing Regimen	Study Population per categorization ¹ at the time of enrollment (no. subjects enrolled)
First-in-human study		
101	Phase 1 randomized, blinded, PC, single-ascending-dose / Single doses of ADU (0.3 to 60 mg/kg) or placebo	Mild to moderate Alzheimer's disease (N=53)
Clinical studies pertinent to the efficacy of aducanumab		
103	Phase 1b randomized, DB, PC, multiple dose, dose escalation, staggered cohort / <u>12-month PC period:</u> ADU (1, 3, 6, 10 mg/kg in a fixed-dose regimen or 10 mg/kg after a 44-week titration) or PBO in a 3:1 or 3:2 ratio <u>LTE period:</u> dose-blinded ADU	Prodromal Alzheimer's disease and mild Alzheimer's disease dementia (N=197)
302	Phase 3 multicenter, randomized, DB, PC, parallel-group / <u>18-month PC period:</u> Low-dose ADU (3 or 6 mg/kg after titration), high-dose ADU ² (10 mg/kg after a 24-week titration), or PBO in a 1:1:1 ratio. Randomization was stratified by ApoE ε4 carrier status. <u>LTE period:</u> dose-blinded ADU	MCI due to Alzheimer's disease and mild Alzheimer's disease dementia (N=1643)
301	Identical in design to Study 302	Same as Study 302 (N=1653)
Other aducanumab clinical studies		
102	Randomized, open-label, bioavailability study / Single dose of ADU (6 mg/kg IV OR 420 mg SC)	Healthy volunteers (N=28)
104	Randomized, DB, PC, single- and multiple-ascending-dose study in Japanese participants / ADU (single and multiple doses of 1 or 3 mg/kg; 6 mg/kg after titration; 10 mg/kg after titration) or PBO in a 4:1 ratio.	Mild to moderate Alzheimer's disease (N=21)
205	Multicenter, randomized, parallel-group, DB, controlled 54-week study with an LTE period / ADU 10 mg/kg after titration	MCI due to Alzheimer's disease and mild Alzheimer's disease dementia (N=52)
304 (ongoing)	Single-arm, longitudinal, multicenter, open-label 24-month study / ADU 10 mg/kg after titration	Participants who had previously participated in Studies 103, 205, 301, or 302 (N=120) ³

Abbreviations: ADU = aducanumab; DB = double-blind; IV = intravenous; LTE = long-term extension; MCI = mild cognitive impairment; PBO = placebo; PC = placebo-controlled; SC = subcutaneous

1: See [Figure 2](#) for changes over time in the categorization of Alzheimer's disease across the disease continuum.

2: High-dose aducanumab in ApoE ε4 carriers was 6 mg/kg (after titration) under Protocol Versions 1-3 and was increased to 10 mg/kg (after titration) under Protocol Versions 4-6. High dose aducanumab in noncarriers was 10 mg/kg (after titration) throughout the study.

3: As of June 15, 2020.

3.1.2. Study 103 Design

The Applicant's Position

When results from the placebo-controlled period became available in 2015, Study 103 was the first study of an anti-A β agent that demonstrated:

- A substantial dose and time dependent reduction in brain amyloid
- A reduction in clinical decline

Key elements were incorporated into the design of Study 103 based on the scientific advances from the research community and previous clinical trials with anti-A β antibodies (Section 2.4) including:

- A study population at earlier stages of disease: Previous trials of other agents were conducted in mild and moderate stages of Alzheimer's disease dementia. Study 103 recruited participants at an earlier stage, including participants in the prodementia stage.
- Methods to reliably identify this earlier population: Previous anti-A β studies enrolled more than 20% participants who did not have amyloid pathology [Salloway 2013]. This raised concern that the failures may have been due at least in part to a significant proportion of the study populations not having Alzheimer's disease, and the pathology not being targeted by the agent. Study 103 addressed this by requiring all participants to have brain amyloid pathology confirmed by newly available technology ([¹⁸F]-florbetapir PET imaging). Study 103 was the first study of an anti-A β agent to enroll participants with amyloid pathology-confirmed early symptomatic Alzheimer's disease.
- Appropriate clinical outcome measures sensitive to change at this earlier stage of disease: The symptoms at the early symptomatic stage are less pronounced than in later stages of Alzheimer's disease and require measures different from those used in later stages. In Study 103 the CDR-SB, was included as an outcome measure. The CDR-SB is an integrated scale that assesses both daily function and cognitive effects and was shown to be sufficiently sensitive and specific to detect change over time in early symptomatic Alzheimer's disease participants [Cedarbaum 2013], an outcome that was confirmed within Study 103. CDR-SB is accepted by the FDA in early Alzheimer's disease clinical studies, and it was agreed to by the FDA and EMA as the primary endpoint for aducanumab's Phase 3 clinical trials (see Section 3.1.3).

Below is a summary of key Study 103 design features, followed by a summary of key Study 103 findings that informed the Phase 3 trial design.

Design

The study was a randomized, multicenter study that included a 12-month randomized, double-blind, placebo-controlled period followed by a dose-blinded LTE period. A total of 196 participants at 27 clinical sites in the United States were randomized and dosed.

Doses studied in the first 3 cohorts (Arms 1-7) were 1, 3, 6, or 10 mg/kg administered every 4 weeks (specifically, 14 doses over the 12-month placebo-controlled period) [Figure 3]. The randomization was stratified by ApoE ϵ 4 status (carrier or noncarrier).

Participants in the fourth cohort (Arms 8 and 9), comprising ApoE ϵ 4 carriers only, were randomized to either aducanumab (2 doses of 1 mg/kg, then 4 doses of 3 mg/kg, 5 doses of 6 mg/kg, and 3 doses of 10 mg/kg) or placebo during the placebo-controlled period. Arms 8 and 9 were added to Study 103 to assess whether the incidence of ARIA, a side effect of brain amyloid removal, could be mitigated in ApoE ϵ 4 carriers by titration.

Objectives

The safety and tolerability of aducanumab was the primary aim of the study.

Secondary outcomes were: (1) the effect of aducanumab on brain amyloid plaque content as measured by [^{18}F]-florbetapir PET, (2) the PK of aducanumab, and (3) the immunogenicity of aducanumab.

Clinical efficacy endpoints were prespecified in the study protocol as exploratory.

During the placebo-controlled period, pharmacodynamic and clinical assessments were performed at 6 months and 1 year.

Population

The study was conducted in participants 50 to 90 years of age (inclusive) with early symptomatic Alzheimer's disease, defined as prodromal Alzheimer's disease (IWG nomenclature [Dubois 2010]) or mild Alzheimer's disease dementia, who were positive for brain amyloid pathology as assessed by [^{18}F]-florbetapir PET. Participants must have had a baseline MMSE score of 20 to 30 (inclusive) and a CDR global score of 0.5 or 1. Both ApoE ϵ 4 carriers and ApoE ϵ 4 noncarriers were enrolled.

Clinical endpoints

Clinical efficacy endpoints included change from baseline on the CDR-SB and MMSE.

CDR-SB and MMSE were later used in the Phase 3 studies; for descriptions of these endpoints, see Section 3.1.3.

Similar to the Phase 3 studies, the CDR-SB and MMSE were administered to participants by different raters, and neither rater was permitted to perform other study assessments (e.g., safety assessments). This was done to protect against the potential for functional unblinding.

Safety monitoring

In Study 103, ARIA was an adverse event of interest, given the experience with previous anti-A β antibodies (Section 2.4), and ARIA management rules were based on published guidelines [Sperling 2011b]. The rules in place at the start of the study were conservative, as ARIA was at that time less well understood. Over time, accumulating data have supported a reduced need for dose terminations and interruptions, and dose reduction after interruptions. All changes in ARIA management in the aducanumab program have been made in consultation with the independent Data Monitoring Committee.

In Study 103, cases of ARIA were managed using centrally read MRI findings and clinical symptoms, if present. These entailed requiring dose suspension or dose termination (depending on the radiographic severity of the ARIA findings observed on MRI, the presence/absence of

symptoms, and, if symptoms were present, their clinical severity), followed by resumption of dosing after the findings resolved/stabilized.

The potential for functional unblinding due to ARIA was minimized by having different raters for efficacy endpoints and for adverse event management.

ARIA monitoring and management are described in more detail in Section 4.

Statistical methods

While safety and tolerability were the main aims of Study 103, the sample size was based on the primary pharmacodynamic endpoint: change from baseline to Week 26 in [¹⁸F]-florbetapir PET signal.

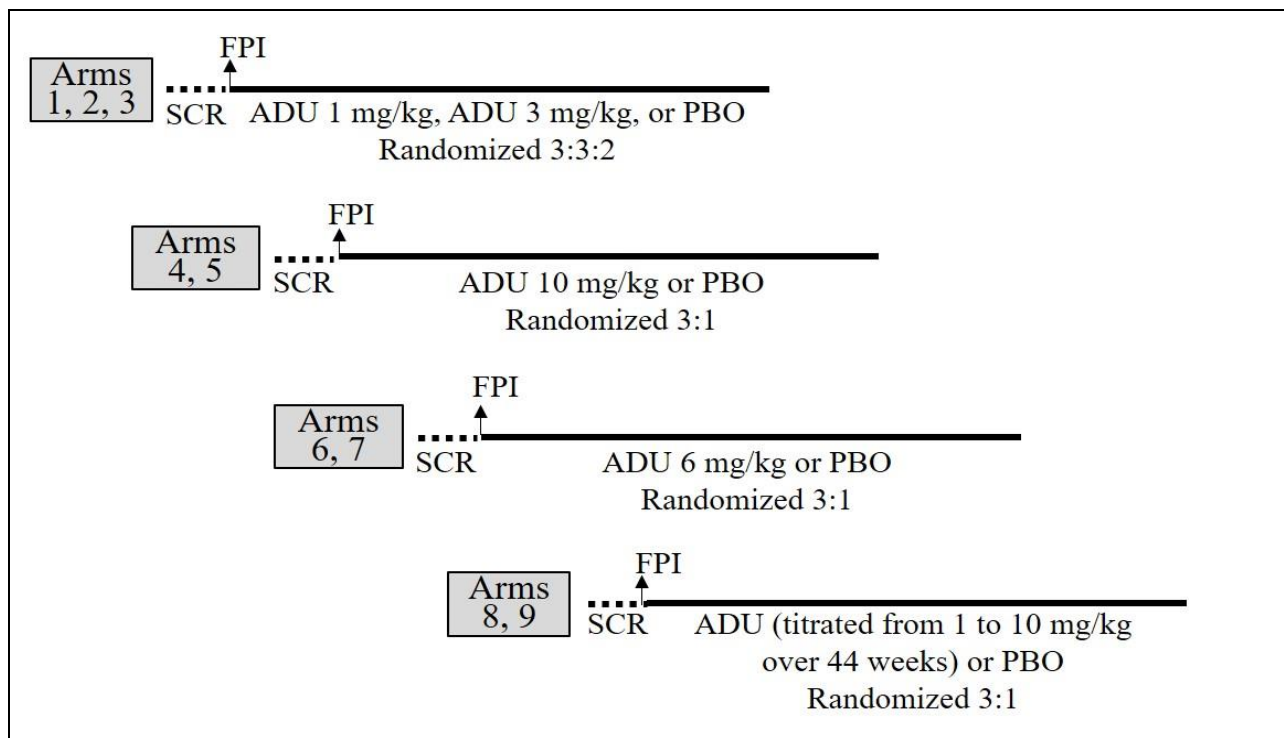
Results presented for Study 103 are based on prospectively determined analyses conducted per the prespecified statistical analysis plan.

To support clinical development needs, including safety monitoring, data were available from several prespecified interim analyses. No changes to the planned statistical analyses were made based on these data.

The original statistical analysis plan was finalized on February 7, 2014, before unblinding of data from the fixed-dose cohorts. The titration cohort was added to the study per Protocol Version 6, dated July 8, 2014, and Addendum 1 to the statistical analysis plan was finalized on August 5, 2016, before unblinding of the titration cohort. No changes to the original statistical methods were made.

No formal adjustment for multiplicity was performed. All p-values in Study 103 were nominal.

Figure 3: Study 103 Placebo-Controlled Period Schematic



Abbreviations: ADU = aducanumab; FPI = first patient in; SCR = screening; PBO = placebo

Note: Arms 1-7 included both ApoE ϵ 4 carriers and noncarriers. Arms 8 and 9 included ApoE ϵ 4 carriers only. Arms 8 (2 doses of 1 mg/kg, then 4 doses of 3 mg/kg, 5 doses of 6 mg/kg, and 10 mg/kg thereafter) had only 3 doses of 10 mg/kg administered during the 12-month placebo-controlled period.

Study 103 findings informed Phase 3 program

The following key findings from Study 103 informed the design of the Phase 3 program. Study 103 results are presented in greater detail in Section 3.7.

- 10 mg/kg was identified as the most efficacious dose, as determined by target engagement and clinical endpoints:
 - Aducanumab's effect on amyloid plaque was shown to be dose and time dependent, with the greatest reduction observed with the 10 mg/kg dose. The magnitude of the reduction in brain amyloid plaque observed with aducanumab 10 mg/kg in Study 103 was subsequently observed to be quantitatively almost identical to that seen with the high dose of aducanumab tested in the Phase 3 Study 302 (Section 3.8).
 - Aducanumab's effect on CDR-SB and MMSE scores was shown to be dose dependent, with the greatest effect observed with the 10 mg/kg dose, despite the study not being powered to detect an effect on clinical measures.
 - These data supported further investigation of the 10 mg/kg dose in the Phase 3 program.

- CDR-SB was confirmed to be sensitive enough to detect change in the early symptomatic Alzheimer's disease participants and therefore appropriate for use in Phase 3.
- While efficacy was observed to be similar regardless of baseline clinical stage, greater variability was observed on the CDR-SB in participants in Study 103 with more advanced disease at baseline (MMSE score of 20-26). Therefore, the Phase 3 enrollment criteria were adjusted such that all participants were required to have a baseline MMSE score of 24-30 (inclusive) and CDR global score of 0.5.
- Data suggested that titration lowered the incidence of ARIA in ApoE ϵ 4 carriers compared with fixed dosing. This finding, combined with efficacy and pharmacodynamic data showing 10 mg/kg (administered as 14 doses over 12 months) to be the most efficacious dose, confirmed selection of both (1) the use of a titration regimen in Phase 3 and (2) the 18-month duration of the placebo-controlled period. Specifically, in the Phase 3 studies, a titration regimen was employed which included a 6-month titration period followed by 14 doses of 10 mg/kg over 12 months; hence, an 18-month treatment duration was implemented in the Phase 3 studies.

The FDA's Position:

The FDA agrees with the general description of Study 103. Although designed primarily as a safety and tolerability study, Study 103 was an adequate and well-controlled study that explicitly included assessment of prespecified clinical and biomarker endpoints and a prospectively identified statistical analysis plan. Although not prospectively controlled for multiplicity, the FDA notes that the choice of endpoints and analytical approach is consistent with that which would have been anticipated should an analytical hierarchy have been in place, and that the prespecified elements were respected. It is important to note that there are design and analysis limitations in Study 103 as compared to a typical confirmatory Phase 3 study design. In particular, because of the staggered dose design and randomization scheme there is no direct concurrent randomization to the various treatment arms to inform a dose-response analysis or a dose vs. placebo comparison based upon concurrent randomization. However, the placebo arms were pooled across cohorts in compliance with the prespecified statistical analysis plan because of the uneven randomization to placebo in each cohort within the staggered design. The study informed the design of Studies 301 and 302 but also included similar elements as these studies, including the requirement of a positive amyloid PET scan and blinded assessment of clinical endpoints. Notable differences include the smaller sample size, inclusion of patients further along the disease continuum, and enrollment of patients only in the United States. The FDA's comments on the results of Study 103 are included in Section 3.7.3.

3.1.3. Study 301 and 302 Design

The Applicant's Position

Design

The identically designed Studies 301 and 302 were large, global, randomized, double-blind, placebo-controlled, parallel-group studies designed to assess the efficacy, safety, PK, and pharmacodynamics of aducanumab.

The studies included an 18-month double-blind placebo-controlled period followed by a dose-blinded long-term extension (LTE).

Combined, the studies enrolled 3285 participants at 348 sites in 20 countries (Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Japan, the Netherlands, Poland, Portugal, South Korea, Spain, Sweden, Switzerland, Taiwan, the United Kingdom, and the United States).

Objectives

The primary objective of the studies was to evaluate the efficacy of aducanumab in reducing clinical decline, as measured on the CDR-SB.

Secondary objectives were to evaluate the effect of aducanumab on clinical decline as measured on the MMSE, ADAS-Cog13, and ADCS-ADL-MCI.

There was one tertiary efficacy objective: to assess the effect of aducanumab on behavior, as measured on the NPI-10.

Other tertiary objectives included the safety and tolerability of aducanumab and the PK of aducanumab.

The following biomarkers were assessed in substudies of both Studies 301 and 302 to evaluate the effects of:

- aducanumab on biomarkers proximal to its mechanism of action (brain amyloid pathology, as assessed by [¹⁸F]-florbetapir PET in the brain amyloid PET substudy, with a subset of sites in Japan using [¹⁸F]-flutemetamol, and by CSF β -amyloid₁₋₄₂ levels in the CSF substudy);
- aducanumab-mediated β -amyloid lowering on downstream disease-related pathophysiological processes (intracellular tau accumulation as measured by CSF p-tau levels, and neurodegeneration as measured by CSF t-Tau levels) in the CSF substudy; and
- aducanumab on downstream tau pathophysiology as assessed by [¹⁸F]-MK-6240 Tau PET in a substudy.

Additionally, the effects of aducanumab on brain volume change as measured by MRI was assessed in all participants; this is considered a nonspecific measure of neurodegeneration.

In addition to the baseline assessment, clinical measures were evaluated at 6 months, 1 year, and 18 months, amyloid PET was done at 6 months and 18 months (Figure 4), CSF and Tau PET

assessments were done at 18 months, and MRI (as a biomarker assessment) was done at 6 months and 18 months.

Population

The Phase 3 studies were conducted in participants 50 to 85 years of age (inclusive) with early symptomatic Alzheimer's disease who were positive for brain amyloid pathology as assessed by PET. Participants must have had a baseline MMSE score of 24 to 30 (inclusive) and a CDR global score of 0.5. Both ApoE ϵ 4 carriers and ApoE ϵ 4 noncarriers were enrolled.

Per the study protocols, ~80% participants in both studies had a baseline clinical diagnosis of MCI due to Alzheimer's disease [Albert 2011], with ~20% having a diagnosis of mild Alzheimer's disease dementia [McKhann 2011].

Clinical endpoints

Once symptoms are noticeable, Alzheimer's disease is characterized by progressive decline in cognition and function, including memory, language, executive and visuospatial function, personality, and behavior, which causes loss of abilities to perform instrumental and/or basic activities of daily living.

The clinical efficacy scales used in the Phase 3 studies were chosen because they are regarded as appropriate measures for the patient population included in the studies, i.e., with apparent detectable abnormalities on neuropsychological measures, and mild but detectable functional impairment. Disease progression in this early symptomatic patient population is expected to be slow over this timeframe.

Five clinical efficacy scales were selected to measure the broad array of symptoms experienced by individuals with Alzheimer's disease: CDR-SB, MMSE, ADAS-Cog13, ADCS-ADL-MCI, and NPI-10. All of these scales are well-validated and are commonly used in Alzheimer's disease clinical research. Furthermore, assessments on these 5 scales collectively reflect several independent sources of information: the patient, caregiver, and independent clinical assessors. The total scores of these 5 scales indicate the severity of the disease, with changes over time reflecting clinical progression.

The primary outcome, CDR-SB, is a measure of both cognition and function and has been established as an appropriate measure in early symptomatic Alzheimer's disease [Cedarbaum 2013]; this was shown in our own program in Study 103 (Sections 3.1.2 and 3.7). The secondary (MMSE, ADAS-Cog13, ADCS-MCI-ADL) and tertiary (NPI-10) efficacy outcome measures all complement the CDR-SB, and each provides important and largely independent information, with minimal overlap across measures. This wide array of cognitive, functional, and neuropsychiatric measures provides a comprehensive and thoroughly validated assessment of the Alzheimer's disease state, with demonstrated sensitivity to change as the disease progresses.

Alzheimer's disease severity subsumes a range of different manifestations that are heterogeneous in nature (i.e., they each arise independently and at possibly different rates, out of a common disease pathophysiology that progresses with some anatomic and physiologic variation). Within this universe of disease manifestations, test results may be associated with one another. Some may be correlated because they are duplicative (e.g., orientation is a construct that is similarly tested in the CDR-SB, the MMSE, and the ADAS-Cog13). Alternatively, associations may be

causative (i.e., memory impairment and dyspraxias rated on the ADAS-Cog13 can causally contribute to difficulties in cooking a meal [rated on the ADCS-ADL-MCI]). Therefore, the constructs will change over time in tandem as disease progresses, and in that sense, they can correlate with one another.

To assess the independent contribution of each scale, Principal Component Analyses of individual items from the primary and 3 secondary outcome measures were conducted on natural history data (Alzheimer's Disease Neuroimaging Initiative [ADNI]) and the placebo groups of Studies 301 and 302.

This showed that a large number of factors is needed to explain the majority (approximately 80%) of the information in the 4 scales, thus demonstrating that the CDR, MMSE, and ADAS-Cog13 each measure different aspects of the cognitive state of the patient, and that there is very little overlap among them. Similarly, while CDR and ADCS-ADL-MCI both measure the functional state of the patient, each provides information that is minimally covered by the other. The NPI-10 measures several dimensions of neuropsychiatric and behavioral symptoms that are distinct from one another and from the cognitive and functional symptom items on the other scales.

It is important to distinguish principal components from correlated changes in symptoms over time. Principal components address questions about what underlying constructs the scales measure, based on the individual items of the scales. It is possible, often likely, that independent principal components will have highly correlated changes over time as the symptoms of Alzheimer's disease progress together. In other words, changes in total scores of efficacy scales can be highly correlated yet the scales measure largely independent factors.

The CDR-SB and the secondary efficacy assessments were conducted by 2 different trained health care professionals (raters), and neither rater was permitted to perform other study assessments (e.g., safety assessments) or have access to any other clinical data. This was done to protect against the potential for functional unblinding.

The primary clinical efficacy outcome measure, CDR-SB, was discussed with and accepted by FDA as part of its Special Protocol Assessment (SPA) of the Phase 3 protocols in 2015 as a single primary efficacy outcome measure. Disease progression, measured as change from baseline on CDR-SB, reflects evolving disability on both cognition and functional abilities, and was deemed adequate to assess the effects of the drug. The FDA endorsed a statistically significant effect on the CDR-SB as a clinically meaningful outcome. The secondary clinical efficacy measures (MMSE, ADAS-Cog13, and ADCS-ADL-MCI) were accepted as providing supportive information.

Clinical Dementia Rating Scale – Sum of Boxes (CDR-SB)

The primary endpoint was the change from baseline in CDR-SB at Week 78. The scale integrates assessments from 3 domains of cognition (memory, orientation, judgment/problem-solving) and 3 domains of function (community affairs, home/hobbies, personal care). Following caregiver interview and systematic patient examination, the rater assigns a score that best describes the patient's current performance level in each of these essential domains of life functioning. Detailed anchors guide scoring to reflect the severity of performance disability in each domain. Prespecified severity anchors range from none = 0, questionable = 0.5, mild = 1, moderate = 2 to severe = 3 (the personal care domain omits the 0.5 score). The "sum of boxes" scoring

methodology sums the score for each of the 6 domains and provides a value ranging from 0 to 18 that can change in increments of 0.5 or greater. Higher scores indicate greater disease severity.

Mini-Mental State Examination (MMSE)

The first-ranked secondary efficacy endpoint was the change from baseline in MMSE at Week 78. The MMSE is a widely used performance-based test of global cognitive status. It has several known limitations impacting sensitivity to change, particularly in earlier disease stages: substantial ceiling effects, sensitivity to practice effects, scores are impacted by patients' educational achievement, and learning effects are observed [Chapman 2016; Franco-Marina 2010; Galasko 1993; Spencer 2013]. Also, the MMSE lacks items reflecting executive dysfunctions often seen in early clinical stages. It consists of 11 tasks that assess orientation, word recall, attention and calculation, language abilities, and visuospatial functions [Folstein 1975]. The scores from the 11 tests are combined to obtain the total score, which ranges from 0 to 30, with lower scores over time indicating increasing cognitive impairment.

Alzheimer's Disease Assessment Scale – Cognitive Subscale (13-Item version) (ADAS-Cog13)

The second-ranked secondary efficacy endpoint was the change from baseline in ADAS-Cog13 at Week 78. ADAS-Cog13 comprises both cognitive tasks and clinical ratings of cognitive performance [Mohs 1997; Rosen 1984]. The scale items capture word recall, ability to follow commands, the ability to correctly copy or draw an image, naming, the ability to interact with everyday objects, orientation, word recognition, memory, comprehension of spoken language, word-finding, and language ability, with a measure for delayed word recall and concentration/distractibility. The total score ranges from 0 to 85. An increase in score over time indicates increasing cognitive impairment.

Alzheimer's Disease Cooperative Study - Activities of Daily Living Inventory (Mild Cognitive Impairment version) [ADCS-ADL-MCI]

The third-ranked secondary efficacy endpoint was the change from baseline in ADCS-ADL-MCI at Week 78. The ADCS-ADL-MCI consists of 17 instrumental items (e.g., shopping, preparing meals, using household appliances) and 1 basic item (getting dressed). Ratings reflect caregiver observations about the patient's actual functioning over the previous month and provide an assessment of change in the functional state of the participant over time. The total score ranges from 0 to 53, with lower values over time reflecting functional deterioration. The ADCS-ADL has been used as an endpoint in Alzheimer's disease clinical trials of symptomatic therapies [Bullock 2005; Tariot 2000] and more recently as a co-primary endpoint with ADAS-Cog in an anti-A β trial [Doody 2014].

Neuropsychiatric Inventory-10 (NPI-10)

The study included a single tertiary clinical efficacy endpoint: change from baseline in NPI-10 at Week 78. The NPI-10 is completed by an interviewer with the study partner informant and systematically indexes the presence, frequency, and severity of 10 neuropsychiatric symptoms: delusions, hallucinations, depression/dysphoria, anxiety, apathy, euphoria, irritability/lability, disinhibition, agitation/aggression, and aberrant motor behavior [Cummings 2020]. The total score ranges from 0 to 120, with higher scores indicating worse symptoms.

Doses studied

The Phase 3 studies compared the effects of 2 dosing regimens of aducanumab versus placebo over the 18-month placebo-controlled period. Participants were randomized 1:1:1 to aducanumab low dose, aducanumab high dose, or placebo. The randomization was stratified by ApoE ϵ 4 status.

Aducanumab 10 mg/kg was hypothesized to be the most efficacious dose, based on results from nonclinical studies (Section 2.3.2) and Study 103 (Section 3.1.2). However, at initiation of the Phase 3 studies, due to a more limited understanding of ARIA than today and with the aim to reduce the incidence of ARIA, both the aducanumab low and high doses differed based on the participants' ApoE ϵ 4 status, as shown below. Consequently, at the beginning of the studies, only ApoE ϵ 4 noncarriers (approximately one-third of the participants in the high-dose arm) were randomized to 10 mg/kg. Additionally, to mitigate ARIA, each target dose was reached after a titration period. For the hypothesized efficacious dose of 10 mg/kg, the titration period was 6 months (i.e., 6 months of titration to the target dose was followed by 14 doses of 10 mg/kg over the next 12 months).

Dosing Protocol Versions 1-3

- Low dose:
 - ApoE ϵ 4 carriers: 3 mg/kg after titration over 8 weeks
 - ApoE ϵ 4 noncarriers: 6 mg/kg after titration over 24 weeks
- High dose:
 - ApoE ϵ 4 carriers: 6 mg/kg after titration over 24 weeks
 - ApoE ϵ 4 noncarriers: 10 mg/kg after titration over 24 weeks

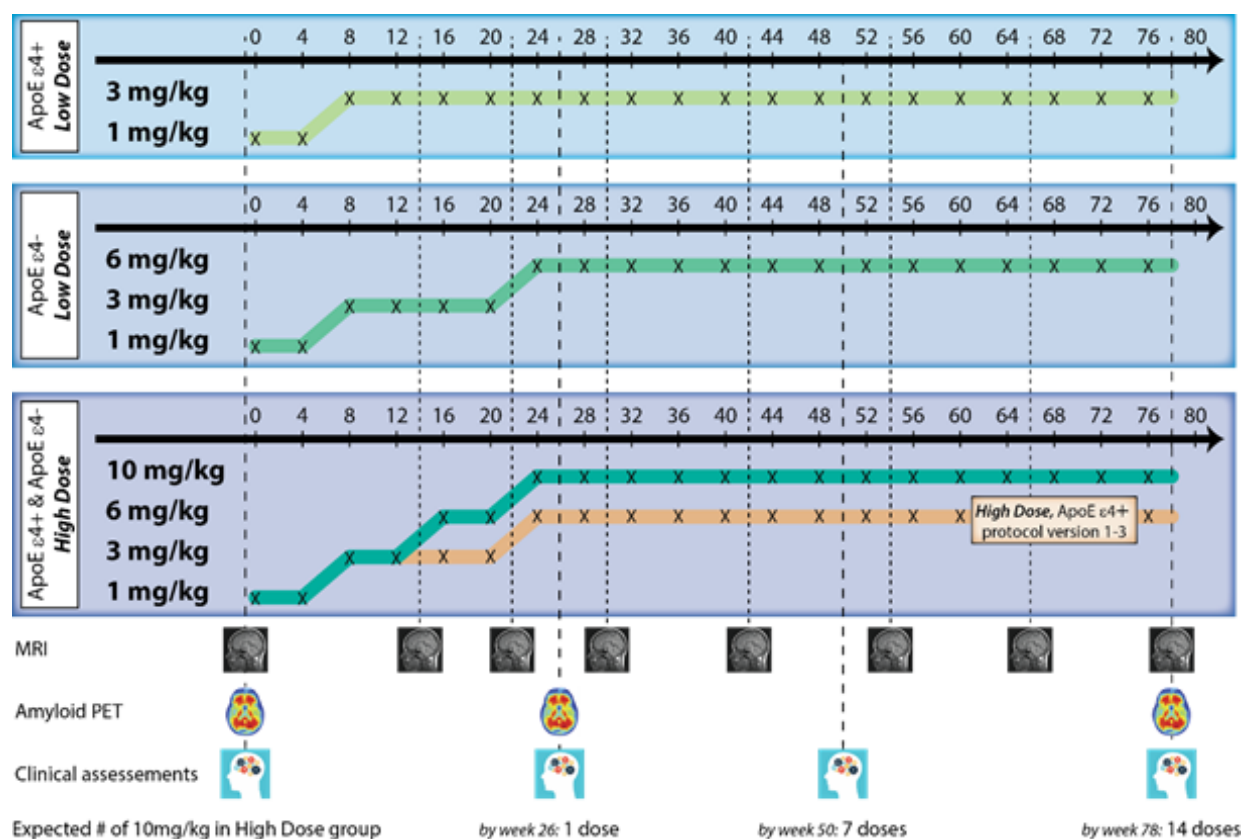
In 2017, based on ARIA data from the final enrolled cohort of Study 103, the Phase 3 protocols were amended mid-study (in Protocol Version 4) to increase the high dose in ApoE ϵ 4 carriers to 10 mg/kg. These changes were made in consultation with the independent Data Monitoring Committee used for the aducanumab program as well as global regulators.

Dosing Protocol Versions 4-6

- Low dose:
 - Unchanged
- High dose:
 - 10 mg/kg (after titration over 24 weeks) in all participants regardless of the participant's ApoE ϵ 4 status

This change in dosing is illustrated in [Figure 4](#).

Figure 4: Schematic of Study Design including Aducanumab Doses (Pre- and Post-Protocol Version 4)



ARIA management

At the start of the Phase 3 program, ARIA was managed using rules largely similar to those then in place in Study 103.

More specifically, Protocol Version 1 mandated dosing suspension under certain circumstances (e.g., if ARIA was accompanied by mild or moderate symptoms, or if ARIA-E was radiographically moderate or severe regardless of whether symptoms were present). Dosing could be resumed at the next lower dose after the event resolved; these participants were then required to remain at that dose for the duration of the trial. Under other circumstances (e.g., if symptoms of greater severity were present), dosing was required to be discontinued permanently.

With Protocol Version 3 several changes in ARIA dose management were implemented:

- For participants who suspended dosing due to ARIA, after resolution of the findings (1) participants could resume dosing at the same dose, and (2) participants could continue titration to the target dose.
- Participants with severe or serious symptoms (other medically important events only) could suspend dosing rather than discontinue permanently.

These changes had the effect of enabling more subjects to continue on aducanumab, and also enabled more of those participants to reach their assigned target dose.

Statistical methods

The primary, secondary, and tertiary efficacy endpoints as well as brain amyloid endpoints were analyzed according to the prespecified analysis plan. In some instances (noted in the data tables), post hoc analyses were conducted; these included analyses of some exploratory endpoints and some sensitivity analyses.

For the primary and secondary endpoints, a prespecified sequential testing procedure that was agreed to by FDA as part of the SPA agreement was used to control the type 1 error rate (rate of false positive results). For other endpoints, statistical significance is considered nominal.

The Phase 3 Statistical Analysis Plans were dated September 12, 2018, and an Addendum to the plan dated November 4, 2019 was finalized prior to database lock.

The Statistical Analysis Plan Addendum did not alter the prespecified primary analysis methods as documented in the September 2018 Statistical Analysis Plan, except for specifying that the primary analysis would exclude efficacy data collected after March 20, 2019, per agreement with FDA at the October 21, 2019 Type C meeting. Additionally, no censoring of data was applied to biomarker, PK, and safety data.

The impact of early termination is discussed in Section 3.2.

The FDA's Position:

The SPA agreements for Studies 301 and 302 indicate FDA's concurrence with the critical elements of the overall protocol design.

FDA agrees with the description of the design of Studies 301 and 302.

CDR-SB is a scale that adequately and meaningfully assesses both daily function and cognitive effects in an integrated manner and is consistent with FDA guidance on clinical endpoints appropriate for Stage 3 patients. FDA accepts a statistically significant change on an inherently meaningful instrument such as CDR-SB as evidence of a clinically meaningful effect.

3.2. Early Termination of the Phase 3 Studies

The Applicant's Position

3.2.1. Prespecified interim analysis and outcome

An interim analysis for futility was prespecified in the study protocols and statistical analysis plan, to allow for the early termination of the studies in the event that the analysis showed the drug to be obviously ineffective, thereby limiting exposure to placebo and the drug.

The Phase 3 protocols and statistical analysis plans specified that an interim analysis would be performed after approximately the first 50% of participants in the studies had the opportunity to

complete (OTC) the Week 78 primary efficacy assessment. The data cutoff date for the prespecified futility analysis was December 26, 2018.

The interim analysis was pre-specified as a futility analysis only, with no possibility of stopping for positive efficacy. Therefore, no need exists for alpha adjustment.

To maintain the integrity of the study, an independent group external to Biogen, which was not involved in the conduct of the study, performed the analysis. The independent Data Monitoring Committee reviewed the unblinded results of the interim analysis and, based on prespecified criteria, made the recommendation to Biogen to terminate the studies. Biogen leadership reviewed the results and accepted the independent Data Monitoring Committee's recommendation.

The prespecified criteria for futility were primarily based on conditional power for CDR-SB, which is the probability calculated on the data at the interim analysis that the final analysis would show statistical significance in favor of aducanumab. The studies were to be considered futile if the conditional power was less than 20%. The conditional power for each study was calculated on a future estimate based on pooled data from Studies 301 and 302.

The use of pooled data was based on the fact that the pooling approach has better operating characteristics than the approach based on single-trial data, if there is only small to moderate heterogeneity on the treatment effects between the 2 studies [Deng 2020]. As the two Phase 3 studies were identically designed, large heterogeneity was not expected.

The results of the prespecified futility analysis showed that estimated conditional power values, based on the pooled data, for the CDR-SB in the high-dose group were 12% for Study 302 and 0% for Study 301. As such, the probability of a statistically significant difference was below the prespecified cutoff of 20%. With the futility criteria met, on March 21, 2019 Biogen announced the termination of the aducanumab Phase 3 program.

3.2.2. Futility results (based on pooled December 2018 data) did not accurately reflect results of the individual studies

While the prespecified futility criteria were met (based on the pooled December 2018 data, as prespecified), the conditional power values differed between studies (0% for 301 and 12% for 302), and this was reflected in the individual study results based on that same dataset using the prespecified primary efficacy analysis methods. Specifically, the individual study results using the prespecified primary efficacy analysis methods on the December 2018 data showed:

- A percent treatment difference favoring high-dose aducanumab of -18% on the CDR-SB was seen in Study 302, whereas a percent treatment difference of 15% on CDR-SB favoring placebo was seen in Study 301.

This outcome violated a key assumption in the estimation of conditional power, namely that the treatment effect in the two studies would be similar. Because of this, after the futility announcement, and in response to a subsequent FDA request, the conditional power was re-estimated for the individual studies. Results of this non-pooled analysis yielded conditional power estimates of 0% and 59% on the primary endpoint for the high dose in Studies 301 and 302, respectively.

In brief, using the non-pooled analysis, the studies would not have met futility criteria.

A second assumption in estimating conditional power was also violated: the assumption that the treatment effect would not change substantially over time in these studies. As will be discussed in Section 3.3, analysis of data up to the time of the futility decision (March 20, 2019) showed an improved treatment effect relative to the December (futility) dataset for both Studies 302 and 301. As expected, this improved treatment effect (relative to the December [futility] dataset) was also confirmed by the final data analysis, as described in Sections 3.4.2 and Section 3.5.2.

Therefore, although the criteria for the futility analysis were prespecified, two key assumptions on which the futility analysis was based were invalid, and results of the futility analysis yielded inaccurate predictions for the final outcomes.

Reasons for these inaccurate predictions, as well as for the differences between Studies 302 and 301, are explored in Section 3.6.

The FDA's Position:

The FDA acknowledges that the Applicant followed the prespecified plan by announcing the termination of the aducanumab Phase 3 studies in response to the futility analysis. At the June 14, 2019, Type C meeting, however, FDA stated, "it would have been more appropriate if futility had not been declared for those studies." Upon initial receipt of the summary of the futility analysis results, the FDA also noted the apparently divergent effects of the high dose of aducanumab between Study 301 and 302 on primary and secondary efficacy parameters. In preliminary comments to the June 14, 2019, Type C Meeting, the FDA therefore asked the Applicant to provide conditional power estimates if non-pooled futility analyses had been performed for each study independently. The assumption that the treatment effect would be similar in the two studies was clearly not realized. The assumption that the treatment effect would not substantially change during the study was also not met.

3.3. Post-Futility

3.3.1. Analysis of a larger dataset

After the futility announcement, in an effort to understand why the studies had divergent outcomes based on the December 2018 dataset (see Section 3.2.2), Biogen ran the primary analyses using all data collected up to March 20, 2019 (futility decision).

All data included in this larger dataset (March 2019 ITT dataset) had been collected under rigorous, unchanged, protocol-specified, double-blind conditions, as the studies had continued to be conducted per the clinical study protocols until the futility announcement.

In the March 2019 ITT dataset, 66% of participants from Study 301 and 60% of participants from Study 302 had the opportunity to complete the Week 78 assessments. The March 2019 ITT data also included 100% of participants who had the opportunity to complete the Week 26 assessments and over 80% with the opportunity to complete the Week 50 assessments (both studies).

With the addition of these substantial additional data in the March 2019 ITT dataset, analysis of the data using the prespecified primary analysis methods yielded results that differed from the

results in the December 2018 dataset. The March ITT data were preliminary, in that following the March 21, 2019, futility announcement, collection of Phase 3 data and data cleaning continued (although dosing had been stopped, with no doses administered after March 20, 2019). Results based on the final data, following database lock in November 2019, showed that the final results were consistent with these preliminary March ITT results. Specifically:

- Study 302 showed an improved treatment effect, with the percent treatment difference favoring high-dose aducanumab changing from -18% on the CDR-SB in the interim (December 2018) results to -22%.
- In Study 301, the results for the high dose no longer showed marked worsening as compared with placebo, with the treatment difference favoring placebo changing from 15% to 2% on the CDR-SB.

3.3.2. Biogen-FDA Collaborative Investigation

Biogen shared the March 2019 ITT results with the FDA, seeking the Agency's counsel and expert opinion on the appropriateness and interpretation of the analyses. The FDA indicated in the June 14, 2019 Type C meeting minutes that, "it is possible, on face, that the effects of aducanumab in Study 302 might not only be interpreted as being supportive of the efficacy of that compound in Alzheimer's disease, but might also be considered exceptionally persuasive on several of the instruments used to evaluate efficacy." Furthermore, based on the data presented by Biogen, FDA referred to the prespecified plan for the futility analysis in the meeting minutes as "flawed," and indicated that "it would have been more appropriate if futility had not been declared for those studies."

As described in Section 2.5, the FDA and Biogen agreed that further analyses of the Phase 3 data were needed to determine the next steps for the aducanumab development program. Considering the significant unmet need for Alzheimer's disease, and the possibility that aducanumab could be an effective drug for the treatment of Alzheimer's disease, the FDA considered it "imperative that extensive resources be brought to bear on achieving a maximum understanding of the existing data," as stated in the June 2019 Type C meeting minutes. The FDA further stated in the minutes, "Given the wholly unique situation that is the current state of the aducanumab development program ..., those further analyses would best be conducted as part of a bilateral effort involving the Agency and sponsor, i.e., through a 'workstream' or a 'working group' collaboration."

3.3.3. Interpretability of Phase 3 data

A first objective agreed upon between Biogen and the FDA was to determine whether the Phase 3 data were valid and interpretable, given the early termination of the studies. It was agreed that this assessment was a prerequisite to embarking on any other analyses.

To address the question of interpretability, in the context of the studies having been inappropriately terminated early, Biogen and the FDA agreed to take the novel approach of "virtually completing" the Phase 3 trials. Virtual completion of the trials was done using statistical modeling and simulation based on information from the observed data from the March 2019 dataset to understand the range of plausible final trial outcomes if the trials had run to

completion. Two prospectively defined approaches were agreed on prior to being conducted by Biogen:

- Approach 1 addressed the question: Given the data seen at the time of the futility declaration, what was the range of plausible results had the studies run to completion? Approach 1 simulated completion of the existing trials (i.e., supplemented the large amount of observed data with 1000 simulated outcomes for the data that were administratively censored due to the early termination of the trials).
- Approach 2 addressed the question: Given the data seen at the time of the futility declaration, what was the range of plausible results if a large number of trials like these were run start to finish? Approach 2 virtually completed 5000 trials with fully simulated data, informed by the observed data.

Within each approach, several parameterizations were used to predict the unobserved data in order to assess outcomes over the plausible range. Approach 1 was considered the primary assessment because observed outcomes were included, and what would have happened at completion of the existing trials was considered the key question. Approach 2 provided supportive evidence.

For each study, consistent results were obtained across all parameterizations, and the results for each study were also consistent with the observed results of the terminated studies.

- For Study 302, of the 1000 simulated studies in Approach 1, 94% yielded a statistically significant difference for high-dose aducanumab versus placebo on the CDR-SB, and for the 5000 simulated studies in Approach 2, the corresponding percentage was 90%. Higher percentages of a statistically significant result were seen for the ADAS-Cog13 and ADCS-ADL-MCI (up to 99%), and lower percentages for MMSE (as low as 76%).
- In contrast, Study 301 yielded low success rates for high-dose aducanumab.

In other words, the results of the virtually completed trials were highly consistent with those of the terminated trials.

3.3.4. Conclusions

Based on these results, the FDA and Biogen jointly concluded that early termination of the Phase 3 program did not compromise the ability to interpret the results. Results of the two trials were jointly concluded to be reliable and interpretable on face and reflected an accurate representation of the effects of aducanumab. As such, the data were suitable for further analysis. The October 2019 Type C meeting minutes state, “we agree that the results of Studies 301 and 302 are interpretable and suitable for additional consideration.”

At this juncture Biogen and the FDA also agreed that the dataset to be used in the final analysis of Phase 3 data would be based on all data through database lock (November 2019), with efficacy data after March 20, 2019, censored.

The FDA's Position:

The FDA agrees with the discussion and conclusions presented by the Applicant in Section 3.3.

3.4. Summary of Study 302

3.4.1. Disposition, Demographics, and Baseline Characteristics

The Applicant's Position

Final Study 302 data were analyzed using the prespecified analysis plan.

A total of 1638 participants were randomized and dosed in the study. Of these, 90.4% of participants were either still on study or had completed the study as of March 20, 2019. Among those who had the opportunity to complete the Week 78 visit by March 20, 2019, 87.6% completed the study.

Overall, 9.6% of all randomized and dosed participants had withdrawn from the study as of March 20, 2019. Among those with the opportunity to complete the Week 78 visit, 12.4% withdrew from study prematurely. The degree of participants prematurely withdrawing from the study was low. In general, Alzheimer's disease studies of this duration have reported a premature withdrawal rate of 20% to 30% [Doody 2014; Salloway 2014; Sevigny 2016]. To address this, the study power calculation included an assumption of a 20% dropout rate.

Demographics and baseline characteristics were similar across treatment groups. Per protocol design, most participants had a diagnosis of MCI due to Alzheimer's disease (81.6%), while 18.4% of participants had mild Alzheimer's disease dementia (Table 2). Mean baseline MMSE scores were 25.9 for participants with mild Alzheimer's disease dementia and 26.4 for participants with MCI due to Alzheimer's disease. Demographics and baseline characteristics were well balanced across treatment groups within the PET substudy.

Table 2: Demographics and Baseline Characteristics – ITT Population

	PBO (N=548)	Low dose (N=543)	High dose (N=547)	Total (N=1678)
Age in years, mean \pm SD	70.8 \pm 7.40	70.6 \pm 7.45	70.6 \pm 7.47	70.7 \pm 7.43
Sex, Female n (%)	290 (52.9)	269 (49.5)	284 (51.9)	843 (51.5)
Race				
Asian n (%)	47 (8.6)	39 (7.2)	42 (7.7)	128 (7.8)
White n (%)	431 (78.6)	432 (79.6)	422 (77.1)	1285 (78.4)
Education years, mean \pm SD	14.5 \pm 3.68	14.5 \pm 3.63	14.5 \pm 3.60	14.5 \pm 3.63
Alzheimer's disease medications used, n (%)	282 (51.5)	281 (51.7)	285 (52.1)	848 (51.8)
ApoE ϵ 4, n (%)				
Carriers	368 (67.2)	362 (66.7)	365 (66.7)	1095 (66.8)
Noncarriers	178 (32.5)	178 (32.8)	181 (33.1)	537 (32.8)
Clinical stage, n (%)				
MCI due to Alzheimer's disease	446 (81.4)	452 (83.2)	438 (80.1)	1336 (81.6)
Mild Alzheimer's disease	102 (18.6)	91 (16.8)	109 (19.9)	302 (18.4)
PET SUVR, mean composite \pm SD	1.375 \pm 0.1748	1.394 \pm 0.1837	1.383 \pm 0.1833	1.384 \pm 0.1805
(n) - PET substudy only	(159)	(159)	(170)	(488)

Abbreviations: PBO = placebo; SUVR = standard uptake value ratio.

Data source: 221AD302/CSR/T-DM-PC; 221AD302/CSR/T-BL-CHAR-PC; 221AD302/CSR/T-BL-CHAR-PET-PC

The FDA's Position:

The FDA agrees with the Applicant's presentation of disposition, demographics, and baseline disease characteristics for Study 302.

3.4.2. Efficacy

The Applicant's Position

Study 302 is a robustly positive study and is the primary contributor to substantial evidence of the effectiveness of aducanumab.

Aducanumab slowed clinical decline in early symptomatic Alzheimer's disease participants, as evidenced by statistically significant effects across the primary endpoint, all 3 secondary endpoints, and the tertiary efficacy endpoint. These 5 scales each measure different aspects of the disease (Section 3.1.3). Therefore, the consistent benefit of aducanumab versus placebo in these diverse endpoints reflects broad efficacy across different and important aspects of Alzheimer's disease. The clinical results confirmed the earlier results seen in the independently conducted Study 103 (Sections 3.1.2 and 3.7).

Primary efficacy endpoint – primary analysis

The primary efficacy endpoint analyses based on the ITT population demonstrated the following:

- Aducanumab high dose resulted in a statistically significant reduction in clinical decline compared with placebo at Week 78, as measured on the primary efficacy outcome measure, CDR-SB (22% less decline, $p = 0.0120$) [Table 3, Figure 5].
- Aducanumab low dose showed a numerical advantage over placebo on the CDR-SB at Week 78. This result is indicative of a dose-response relationship (Table 3, Figure 5).
- In an additional analysis using the same methods as the primary analysis, aducanumab high dose had a consistent effect across all 6 domains of the CDR (i.e., orientation, memory, judgment and problem solving, community affairs, home and hobbies, and personal care) as compared with placebo, with treatment differences ranging from -15% to -32% (Table 4).

Table 3: Primary Efficacy Endpoint Analysis: Change from Baseline on the CDR-SB at Week 78 in Study 302

	Difference vs PBO ^a (%)		
	p-value		
	PBO decline (N=548)	Low dose (N=543)	High dose (N=547)
CDR-SB	n=288 1.74	n=290 -0.26 (-15%) 0.0901	n=299 -0.39 (-22%) 0.0120

Abbreviations: N = numbers of all randomized and dosed participants that were included in the analysis; n: numbers of randomized and dosed subjects with primary endpoint assessment at Week 78; PBO = placebo

^aDifference vs placebo at Week 78. Negative percentage means less progression in the treated arm.

Note: The primary analysis was conducted on the ITT population excluding data collected after March 20, 2019.

Note: A mixed model for repeated measures was used as the primary analysis to analyze change from baseline in the CDR-SB using fixed effects of treatment group, time (categorical), treatment group-by-time interaction, baseline CDR-SB, baseline CDR-SB by time interaction, baseline MMSE, Alzheimer's disease symptomatic medication use at baseline, region, and laboratory ApoE ε4 status.

Data source: 221AD302/CSR/T-CDR-MMRM-PC

Table 4: Change from Baseline on CDR Items at Week 78 in Study 302

	Difference vs PBO ^a (%)	
	p-value	
	PBO decline (N=548)	High dose (N=547)
CDR-SB	n=288 1.74	n=299 -0.39 (-22%) 0.0120
Memory	0.25	-0.07 (-28%)
Orientation	0.34	-0.08 (-24%)
Judgment and problem solving	0.28	-0.07 (-25%)
Community Affairs	0.31	-0.07 (-23%)
Home and Hobbies	0.29	-0.07 (-24%)
Personal Care	0.20	-0.03 (-15%)

Abbreviations: N = numbers of all randomized and dosed participants that were included in the analysis; n: numbers of randomized and dosed subjects with primary endpoint assessment at Week 78; PBO = placebo

^aDifference vs placebo at Week 78. Negative percentage means less progression in the treated arm.

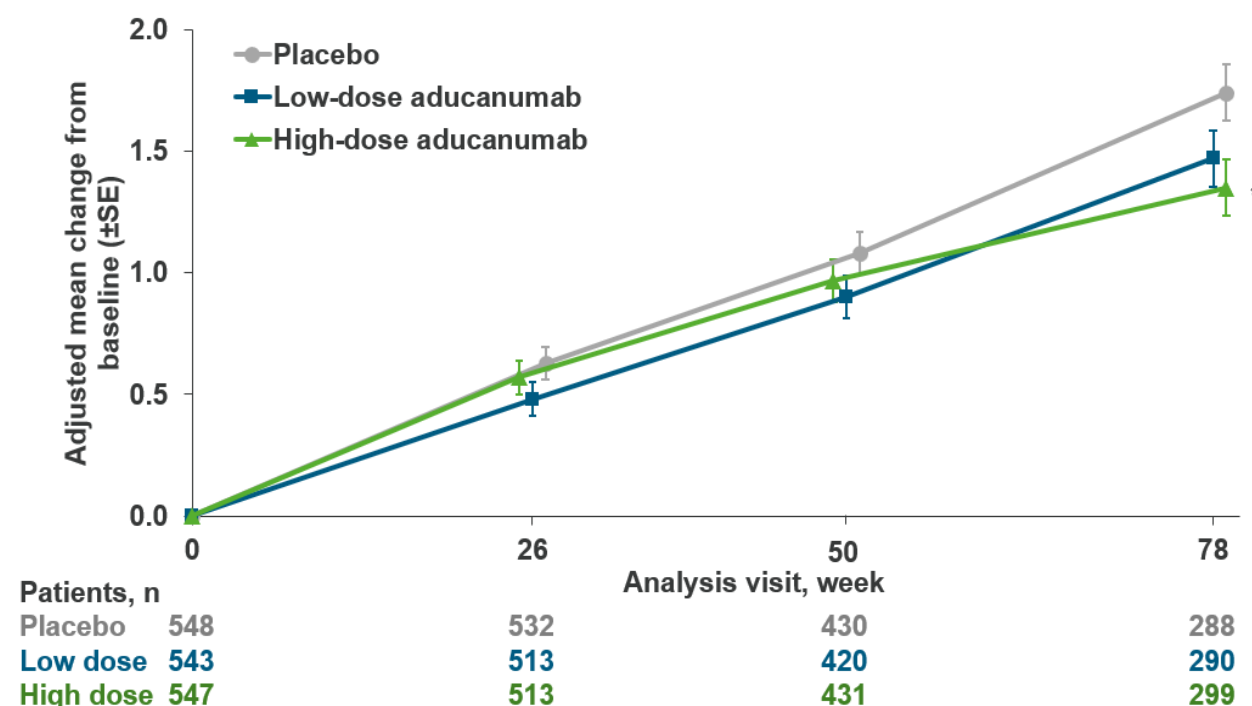
Note: The CDR item analysis was post hoc.

Note: The analysis was conducted on the ITT population excluding data collected after March 20, 2019.

Note: A mixed model for repeated measures was used as the primary analysis to analyze change from baseline in the CDR score using fixed effects of treatment group, time (categorical), treatment group-by-time interaction, baseline CDR score, baseline CDR score by time interaction, baseline MMSE, Alzheimer's disease symptomatic medication use at baseline, region, and laboratory ApoE ε4 status.

Data source: FDA/20200903/F-CDR-ITEM-BAR-MMRM-PC-302

Figure 5: Change From Baseline on the CDR-SB Over Time in Study 302



* $p < 0.05$

Note: Results were based on a mixed model for repeated measures, with change from baseline in CDR-SB as dependent variable and with fixed effects of treatment group, categorical visit, treatment-by-visit interaction, baseline CDR-SB, baseline CDR-SB by visit interaction, baseline MMSE, Alzheimer's disease symptomatic medication use at baseline, region, and laboratory ApoE ε4 status.

Data source: 221AD302/CSR/T-CDR-MMRM- PC

Primary efficacy endpoint – supplementary analyses

The primary analysis excluded all data that became available after the decision to discontinue the aducanumab program was made public, on March 21, 2019, to limit the potential introduction of bias. In response to FDA feedback, an analysis based on all data from all participants (i.e., with no censoring of data after March 20, 2019) from the ITT population was included in the Statistical Analysis Plan Addendum (finalized prior to database lock). Results from this uncensored ITT analysis (Table 5) show that the magnitude of treatment effect (25% reduction) and statistical significance ($p = 0.0029$) at the primary endpoint were maintained in the uncensored ITT analysis despite the participants' having been off treatment for up to 18 weeks prior to the last assessment.

An additional analysis specified in the Addendum was performed using data from participants who had the opportunity to complete (OTC) the Week 78 visit by March 20, 2019. The advantage of including only those participants with the opportunity to complete the Week 78 visit by that date was that the amount of missing Week 78 data (necessitating imputation) was low, which is more similar to what would have been the case had the studies run to completion. Results of this additional OTC analysis were consistent with those for the primary analysis and the uncensored ITT analysis, showing a 22% reduction in decline for high dose aducanumab as compared with placebo ($p = 0.0368$).

Table 5: Change from Baseline in CDR-SB at Week 78 in Study 302: Uncensored Analysis and OTC Analysis

	Uncensored ITT Analysis ^a			OTC Analysis ^b		
	Diff vs PBO ^c (%)			Diff vs PBO ^c (%)		
	p-value			p-value		
	PBO decline (N=548)	Low dose (N=543)	High dose (N=547)	PBO decline (N=313)	Low dose (N=329)	High dose (N=340)
CDR-SB	n = 408	n = 399	n = 403	n = 288	n = 290	n = 298
	1.79	-0.22 (-12%) 0.1273	-0.44 (-25%) 0.0029	1.61	-0.27 (-17%) 0.1188	-0.36 (-22%) 0.0368

Abbreviations: Diff = difference; N = numbers of randomized and dosed subjects that were included in the analysis; n: numbers of randomized and dosed subjects with primary endpoint assessment at Week 78; OTC = opportunity to complete (Week 78); PBO = placebo.

^a Uncensored analysis was conducted on the ITT population including all data collected in the placebo-controlled period; p values are nominal.

^b OTC analysis was conducted on the ITT population who had the opportunity to complete the Week 78 visit by March 20, 2019; p-values are nominal.

^c Difference vs placebo at Week 78. Negative percentage means less progression in the treated arm.

Data source: 221AD302/CSR/T-CDR-MMRM-ALL-PC; 221AD302/CSR/T-CDR-MMRM-OTC-PC

Primary efficacy endpoint – Sensitivity analyses

Sensitivity analyses assessed the robustness of the treatment effect to early treatment discontinuation, the use of or changes in Alzheimer's disease symptomatic medications, and different assumptions for missing data. Several sensitivity analyses were prespecified, as listed below. The jump to reference analysis was not prespecified but was conducted given that this method is a most conservative assessment of missing data.

As seen in [Table 6](#), in these sensitivity analyses, results for the high dose are consistent with the primary analysis.

- **Censoring after intercurrent events:** Excluding data after intercurrent events (i.e., after early withdrawal or change in concomitant medications) estimates the result expected if all participants adhered to their randomized treatment.
- **Copy increment from reference:** This approach was prespecified to estimate what would have happened if participants who prematurely withdrew from study and a similar proportion of those ongoing participants who could have withdrawn prematurely if the study had not been terminated (i.e., a total of 12% of the ITT population) retained the benefit gained from their prior treatment but progressed as if they were on placebo after withdrawal.
- **Jump to reference:** This approach was not prespecified but is the most conservative method among this family of sensitivity analysis approaches. It assumes that participants who withdrew early and a similar proportion of those ongoing participants who could have withdrawn prematurely if the study had not been terminated (i.e., a total of 12% of the ITT population) immediately lost benefit from previous study treatment and progressed as placebo afterward.

Therefore, across several different sensitivity analyses, statistical significance of the high-dose arm was consistently maintained.

Finally, in the **tipping point analysis**, this prespecified delta adjustment approach showed that in order to overturn statistical significance achieved in the primary efficacy endpoint analysis, subjects in the high-dose group who withdrew prematurely by March 20, 2019, and a similar proportion of those ongoing subjects who could have withdrawn prematurely if the study had not been terminated (i.e., a total of 12% of the ITT population) would have had to progress faster than the subjects in the placebo group. This is implausible, given that the large amount of post-treatment data collected after study termination showed no sign of high-dose participants performing worse than the placebo participants after discontinuing aducanumab (see the supplementary Uncensored ITT Analysis, above). The implausibility of the worsening needed to overturn the statistical significance reinforces the robustness of the primary analysis.

Two additional post hoc sensitivity analyses were used to assess the consequences of non-normality in the data ([Table 6](#)). The non-normality (i.e., right-skewness in the data) increased heterogeneity and could possibly bias estimates of treatment effects. The consequences of the non-normality were assessed by repeating the primary analysis on data that were “normalized” using a log-transformation, and a non-parametric procedure that used the ranked order of outcomes rather than the actual values. In the non-parametric and log-transformed analyses, the magnitude of the treatment effect was increased compared to the primary analysis (-0.44 and -0.49, respectively). The data transformations also decreased heterogeneity. Consequently, the p-value from the non-parametric analysis was 3-fold smaller than the primary analysis (0.0120 vs. 0.0041) and the p-value from the log-transform analysis was 20-fold smaller (0.0120 vs. 0.0006).

In summary, the treatment effect observed in the Study 302 high-dose arm was robust to a range of missing data assumptions and departures from normality assumptions.

Table 6: Change from Baseline to Week 78 on the CDR-SB, Primary Analysis and Prespecified and Post Hoc Sensitivity Analyses in Study 302, ITT Population

		Difference vs PBO (%) ^a	
	PBO decline (N=548)	p-value	High dose (N=547)
CDR-SB (primary analysis)	1.74	-0.39 (-22%)	0.0120
Censoring after intercurrent events ^b	1.64	-0.41 (-25%)	0.0119
Copy Increment from reference ^c	1.68	-0.34 (-20%)	0.0289
Jump to reference ^d	1.68	-0.33 (-20%)	0.0350
Log transformation to correct skewness ^e	1.40	-0.49 (-35%)	0.0006
Nonparametric test ^f	--	Hodges-Lehmann estimate of median difference ^f	-0.44 0.0041

Abbreviations: N = numbers of all randomized and dosed participants that were included in the analysis; PBO = placebo.

Note: The primary analysis was conducted on the ITT population excluding data collected after March 20, 2019.

Note: A mixed model for repeated measures was used as the primary analysis to analyze change from baseline in the CDR-SB using fixed effects of treatment group, time (categorical), treatment group-by-time interaction, baseline CDR-SB, baseline CDR-SB by time interaction, baseline MMSE, Alzheimer's disease symptomatic medication use at baseline, region, and laboratory ApoE ε4 status.

^a Difference vs placebo at Week 78. Negative percentage means less progression in the treated arm.

^b Data collected after premature discontinuation of treatment and any change to concomitant Alzheimer's disease symptomatic medication were excluded.

^c Missing data for participants in active drug groups were imputed using the copy increment from reference method. The missing data for participants still active in the placebo-controlled period on March 20, 2019 were imputed using the missing at random method, except for a portion that were assumed to be potential dropout if the study had not been terminated and were imputed using the copy increment from reference method

^d Missing data for participants in active drug groups was imputed using the jump to reference method. The missing data for participants still active in the placebo-controlled period on March 20, 2019 were imputed using the missing at random method, except for a portion that were assumed to be potential dropout if the study had not been terminated and were imputed using the jump to reference method.

^e A mixed model for repeated measures was fitted based on the log-transformed data and the results were then transformed back to the original scale. The p-values are for the difference at the log scale.

^f Non-parametric test p-values were from a rank ANCOVA model of the change from baseline CDR-SB at Week 78 based on a dataset imputed with multiple imputation. Hodges-Lehmann estimator of median difference at Week 78 was calculated based on a dataset imputed with multiple imputation.

Note: Except for the primary analysis, all other p-values are nominal.

Data source: 221AD302/CSR/T-CDR-CIR-PC; 221AD302/CSR/T-CDR-CNSR-PC; ISE/CSR/T-CDR-MMRM-J2R-PC-302; ISE/CSR/T-CDR-MMRM-PC-LOG-302; 221AD302/CSR/T-CDR-NONPARAM-PC

Secondary efficacy endpoints

The primary efficacy endpoint results are supported by statistically significant differences favoring aducanumab high dose on all 3 secondary endpoints at Week 78. Aducanumab high dose resulted in a reduction in clinical decline as measured on the MMSE (18% less decline, p = 0.0493), ADAS-Cog13 (27% less decline, p = 0.0097), and ADCS-ADL-MCI (40% less

decline, $p = 0.0006$), as compared with placebo (Table 7). Statistical significance was achieved using the prespecified sequential testing procedure to control the overall type 1 error rate.

As was seen with the primary efficacy results, aducanumab low dose had a numerical advantage over placebo on 2 of the 3 secondary endpoints at Week 78, which is suggestive of a dose-response relationship when considered with the difference observed between aducanumab high dose and placebo (Table 7).

Table 7: Primary and Secondary Endpoints at Week 78 in Study 302: ITT Population

	PBO decline (N=548)	Difference vs PBO ^a (%)	
		Low dose (N=543)	High dose (N=547)
CDR-SB	n=288 1.74	n=290 -0.26 (-15%) 0.0901	n=299 -0.39 (-22%) 0.0120
MMSE	n=288 -3.3	n=293 -0.1 (3%) 0.7578	n=299 0.6 (-18%) 0.0493
ADAS-Cog13	n=287 5.162	n=289 -0.701 (-14%) 0.1962	n=293 -1.400 (-27%) 0.0097
ADCS-ADL-MCI	n=283 -4.3	n=286 0.7 (-16%) 0.1515	n=295 1.7 (-40%) 0.0006

Abbreviations: N = numbers of all randomized and dosed participants that were included in the analysis; n: numbers of randomized and dosed subjects with endpoint assessment at Week 78.

PBO = placebo.

^a Difference vs placebo at Week 78. Negative percentage means less progression in the treated arm.

Note: The analysis was conducted on the ITT population excluding data collected after March 20, 2019.

Note: A mixed model for repeated measures was used as the primary analysis to analyze change from baseline in the CDR-SB using fixed effects of treatment group, time (categorical), treatment group-by-time interaction, baseline CDR-SB, baseline CDR-SB by time interaction, baseline MMSE, Alzheimer's disease symptomatic medication use at baseline, region, and laboratory ApoE ε4 status. The change from baseline in MMSE, ADAS-Cog13 and ADCS-ADL-MCI scores were also analyzed using a mixed model for repeated measures.

Data source: Table 3; 221AD302/CSR/T-MMSE-MMRM-PC; 221AD302/CSR/T-ADAS-MMRM-PC; 221AD302/CSR/T-ADL-MMRM-PC.

A consistent treatment effect ranging from 26% to 51% was observed across all ADAS-Cog13 items that measure episodic memory, orientation, and attention. Furthermore, item-level analysis for the ADCS-ADL-MCI indicated a reduction in decline across a variety of daily activities, such as tasks related to personal hygiene and use of household appliances.

Results were robust across sensitivity and supplementary analyses.

Tertiary efficacy endpoint – NPI-10

Aducanumab high dose also resulted in less decline in neuropsychiatric symptomatology than placebo, as measured on the NPI-10, a tertiary endpoint, at Week 78 (87% less decline, nominal $p = 0.0215$) [Table 8]. An analysis assessing the consequences of non-normality in the data, similar to what was done for the other clinical endpoints, yielded directionally consistent results (nonparametric analysis, point estimate for treatment difference -0.8, nominal $p=0.0992$).

Table 8: Primary, Secondary, and Tertiary Endpoints at Week 78 in Study 302: ITT Population

		Difference vs PBO ^a (%)	
		p-value	
	PBO decline (N=548)	Low dose (N=543)	High dose (N=547)
CDR-SB	n=288 1.74	n=290 -0.26 (-15%) 0.0901	n=299 -0.39 (-22%) 0.0120
MMSE	n=288 -3.3	n=293 -0.1 (3%) 0.7578	n=299 0.6 (-18%) 0.0493
ADAS-Cog13	n=287 5.162	n=289 -0.701 (-14%) 0.1962	n=293 -1.400 (-27%) 0.0097
ADCS-ADL-MCI	n=283 -4.3	n=286 0.7 (-16%) 0.1515	n=295 1.7 (-40%) 0.0006
NPI-10	n=282 1.5	n=283 -0.5 (-33%) 0.3921	n=291 -1.3 (-87%) 0.0215

Abbreviations: N = numbers of all randomized and dosed participants that were included in the analysis; n = numbers of randomized and dosed subjects with endpoint assessment at Week 78.

NPI-10 = Neuropsychiatric Index-10; PBO = placebo

^a Difference vs placebo at Week 78. Negative percentage means less progression in the treated arm.

Note: In the analysis of change from baseline on the NPI-10, the p-value is nominal.

Note: The analysis was conducted on the ITT population excluding data collected after March 20, 2019.

Note: A mixed model for repeated measures was used as the primary analysis to analyze change from baseline in the CDR-SB using fixed effects of treatment group, time (categorical), treatment group-by-time interaction, baseline CDR-SB, baseline CDR-SB by time interaction, baseline MMSE, Alzheimer's disease symptomatic medication use at baseline, region, and laboratory ApoE ε4 status. The change from baseline in MMSE, ADAS-Cog13, ADCS-ADL-MCI, and NPI-10 scores were also analyzed using a mixed model for repeated measures.

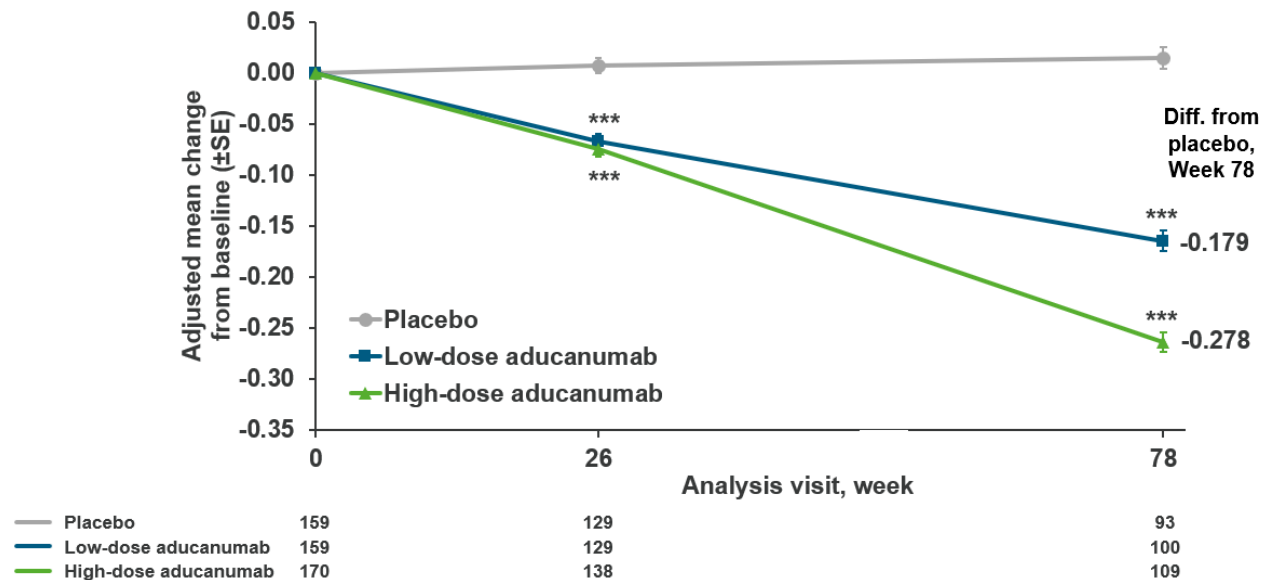
Data source: Table 3; Table 7; 221AD302/CSR/T-NPI-MMRM-PC.

Effects on biomarkers proximal to aducanumab's mechanism of action (brain amyloid plaque and CSF β -amyloid₁₋₄₂ levels)

Aducanumab treatment produced statistically significant, dose-dependent reductions in brain amyloid plaque versus placebo, as measured by PET, at both Weeks 26 and 78 ($p < 0.0001$); the magnitude of reduction was larger at Week 78 than Week 26 (Figure 6). The magnitude of the high-dose effect was similar to the result for the 10 mg/kg fixed-dose group in Study 103. Furthermore, a statistically significant, dose-dependent increase in CSF β -amyloid₁₋₄₂ levels (a second biomarker of target engagement) was observed at Week 78 for both the low-dose and high-dose groups as compared with placebo ($p < 0.0001$).

Together, results for these biomarkers proximal to aducanumab's mechanism of action demonstrated target engagement (brain amyloid plaque as measured by PET, CSF β -amyloid₁₋₄₂ levels) and lowering of brain amyloid plaque levels in the brain (a hallmark of Alzheimer's disease pathophysiology) and indicated that a higher dose and longer duration of aducanumab treatment were associated with greater effects.

Figure 6: Changes in Brain Amyloid Plaque, Measured by PET and Quantified as SUVR in Study 302



Abbreviations: Diff = difference; SUVR = standard uptake value ratio; *** $p < 0.001$ (nominal).

Note: Results were based on a mixed model for repeated measures, with change from baseline as dependent variable and with fixed effects of treatment group, categorical visit, treatment-by-visit interaction, baseline SUVR value, baseline SUVR value by visit interaction, baseline MMSE, baseline age, and laboratory ApoE $\epsilon 4$ status.

Data source: 221AD302/CSR/T-PET-MMRM-COMP-CERE-PC

Effects on downstream biomarkers of intracellular tau pathology and neurodegeneration

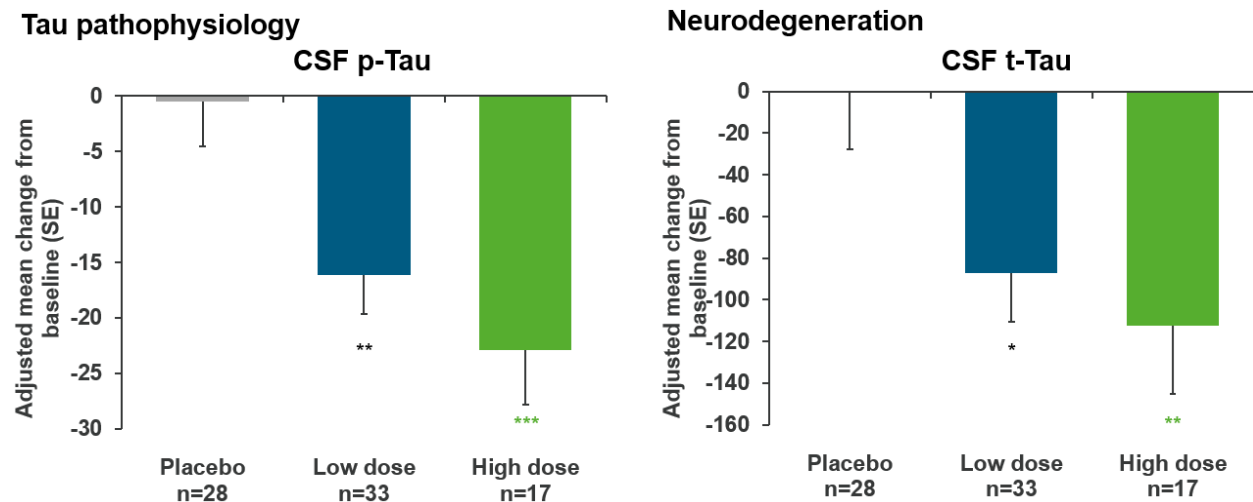
Aducanumab treatment produced statistically significant, dose-dependent reductions in CSF p-Tau levels (high dose, $p = 0.0005$; low dose, $p = 0.0035$) and CSF t-Tau levels (high dose, $p = 0.0088$; low dose, $p = 0.0148$) at Week 78 (Figure 7).

These findings on a downstream Alzheimer's disease-related biomarker of tau pathology (p-Tau) and a downstream marker of neurodegeneration (t-Tau) demonstrated that aducanumab lowered tau accumulation and neurodegeneration in a dose-dependent manner.

Furthermore, a higher cumulative dose of aducanumab was associated with a greater change in both CSF p-Tau and t-Tau levels.

Collectively, these results, with effects on brain amyloid plaque described earlier, show a dose-response relationship consistent with a direct effect of aducanumab on lowering brain amyloid pathology and a downstream action of aducanumab on slowing the progression of tau pathology and neurodegeneration.

Figure 7: CSF p-Tau and t-Tau Change at Week 78 in Study 302, CSF Modified Analysis Population



* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (nominal).

Note: Results were based on an ANCOVA model at Week 78, fitted with change from baseline as dependent variable and with treatment group, baseline biomarker value, baseline age, and laboratory ApoE $\epsilon 4$ status as independent variables.

Data source: ISE/CSR/T-CSF-ANCOVA-PC

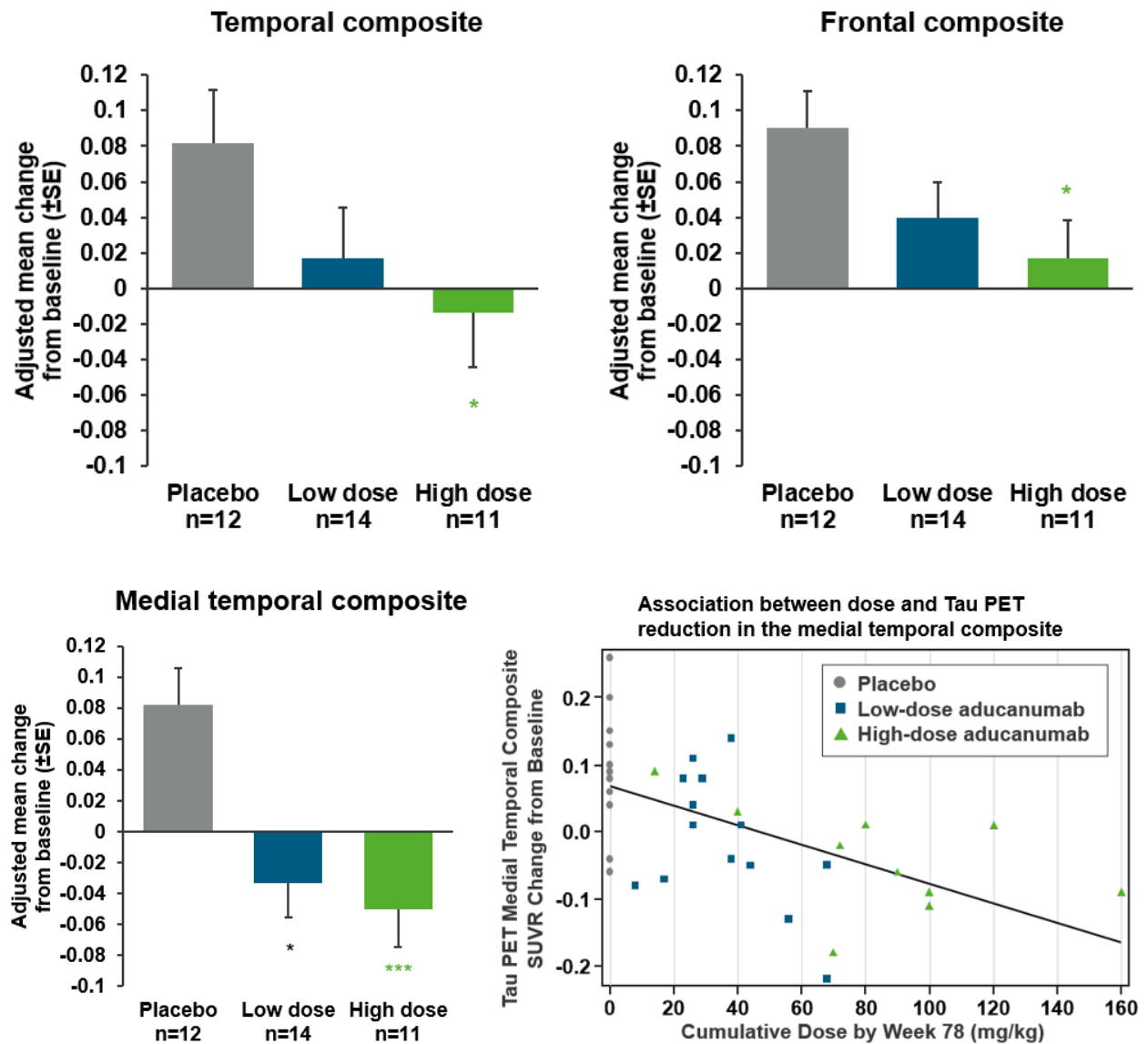
Effects on intracellular tau as measured by PET

Tau PET was incorporated later in the conduct of Studies 301 and 302, when the PET imaging agent [^{18}F]-MK-6240 became available. Therefore, the number of subjects enrolled in the tau PET substudy was small, and the tau PET analysis population comprised later-enrolled subjects from both studies. Furthermore, due to the early termination of the studies, the postbaseline visit was on average at 14 months and ranged approximately between 9 to 20 months. Due to the small sample size, tau PET analyses were performed on pooled data from Studies 301 and 302.

Aducanumab resulted in a dose-dependent reduction in brain tau levels in the frontal (high dose, $p < 0.05$), temporal (high dose, $p < 0.05$), and medial temporal (both doses, $p < 0.05$) composite brain regions, which have tau involvement at the early stages of Alzheimer's disease (Figure 8). As expected, tau levels increased in these regions in the placebo group, indicating accumulation of tau pathology with disease progression over time. No statistically significant differences were observed for the cingulate, parietal, or occipital composites, which may be attributed to the limited degree of tau involvement in these regions at the early stages of Alzheimer's disease.

The largest magnitude of tau reduction was observed in the medial temporal composite; both high-dose and low-dose aducanumab produced a greater reduction in tau than placebo ($p = 0.0005$ and $p = 0.0012$, respectively). In line with this observation of a dose response, a higher cumulative dose was associated with a greater magnitude of reduction in tau in the medial temporal composite region (Figure 8); this relationship was observed even though (as noted above) the postbaseline assessment was on average at 14 months. A consistent relationship between cumulative dose and change from baseline tau was also observed for the frontal and temporal regions as well.

Figure 8: Change from baseline in Tau PET in Studies 301 and 302 (Pooled Data): Frontal, Temporal, and Medial Temporal Composite Regions; Association between Tau PET Change in Medial Temporal Composite Region and Dose



* $p < 0.05$, *** $p < 0.001$ (nominal).

Note 1: Results were based on an ANCOVA model fitted with change from baseline as dependent variable and with treatment group, baseline tau PET value, and laboratory ApoE $\epsilon 4$ status as independent variables.

Note 2: Due to the early termination of the studies, all the Tau PET assessments performed in the placebo-controlled period were used as one postbaseline timepoint.

Data source: ISE/CSR/T-TAU-ANCOVA-PC-POOL, ISE/CSR/F-TAU-SCA-CUMDOS-WK78-PC

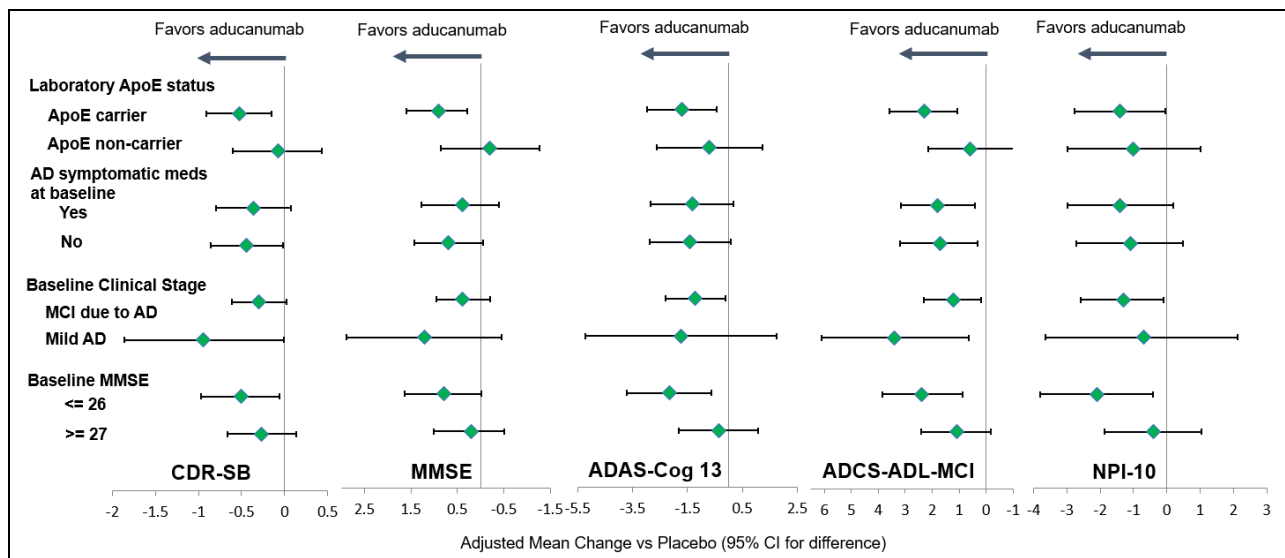
Subgroup analyses

Prespecified subgroup analyses were performed across demographic and disease characteristics. Seven subgroup factors were prespecified (region, age, sex, ApoE $\epsilon 4$ status, baseline clinical stage, baseline MMSE, and baseline use of symptomatic treatments for Alzheimer's disease), each with 2 or 3 subgroup categories, for a total of 16 subgroups.

Treatment contrasts favored aducanumab high dose in 79 of 80 comparisons across 5 clinical endpoints (CDR-SB, MMSE, ADAS-Cog13, ADCS-ADL-MCI, NPI-10) [Figure 9 and Figure 10].

Except for the region subgroups, these same prespecified subgroups listed above were compared regarding reductions in amyloid PET; region subgroups were not considered because no plausible basis existed for differences by region for this outcome. In all 13 of these subgroups, both doses of aducanumab had statistically significant reduction in amyloid PET versus placebo, with the reductions being larger in magnitude for high dose than low dose in each comparison.

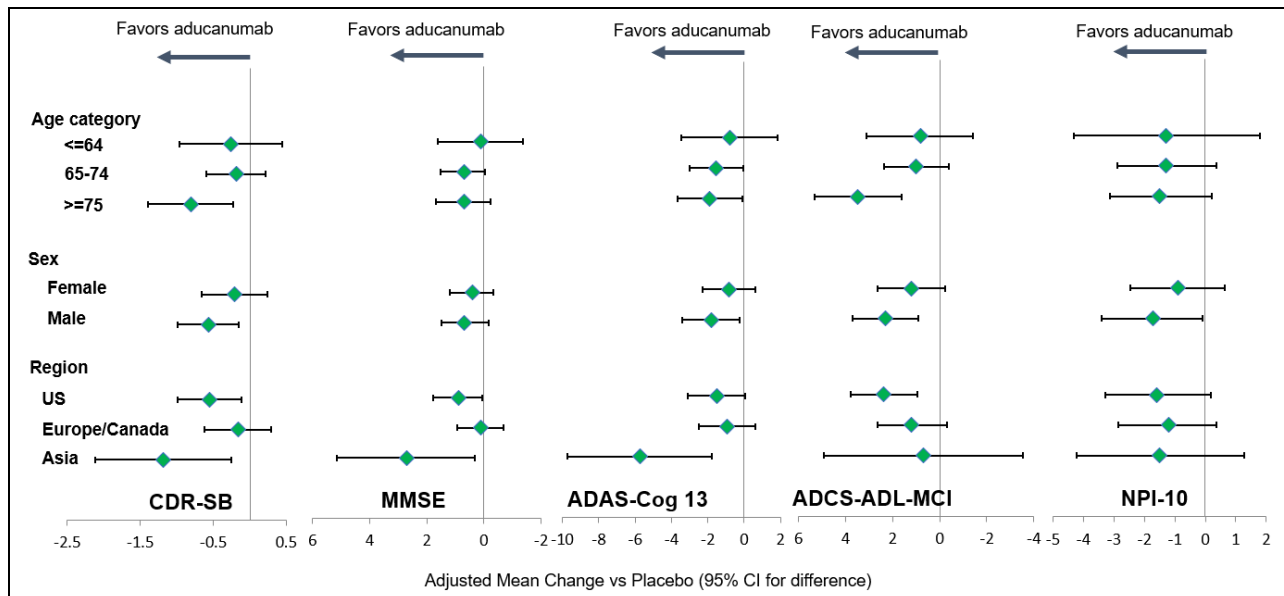
Figure 9: Subgroup Analyses (Disease Characteristics) for Clinical Endpoints, Study 302



Abbreviations: AD = Alzheimer's disease.

Data source: Adapted from the Integrated Summary of Efficacy, Section 4.1.3.5.

Figure 10: Subgroup Analyses (Demographics) for Clinical Endpoints, Study 302



Data source: Adapted from the Integrated Summary of Efficacy, Section 4.1.3.5.

Subgroup analysis - ApoE ε4

ApoE ε4 carriage is the most salient known genetic risk factor for development of Alzheimer's disease. The effects of ApoE ε4 on the risk of developing Alzheimer's disease and the age at onset are well established, with the ApoE ε4 allele increasing the risk of Alzheimer's disease and reducing the age at onset.

However, the relationship between ApoE ε4 carriage and clinical decline is less well defined. In biomarker-verified amyloid-positive subjects, longitudinal and cross-sectional observational studies have demonstrated that once symptoms are present, ApoE ε4 status does not have an impact on the rate of cognitive decline [Roberts 2018; Suzuki 2020].

Whether ApoE ε4 carriage has an impact on the treatment effect of an anti-Aβ therapy is unknown.

The pre-planned subgroup analyses for Study 302 suggested slowing of decline relative to placebo in both ApoE ε4 carriers and noncarriers (Figure 9). Although the effect in carriers was larger, the interaction test was not significant, indicating that the difference in the treatment effect between carriers and noncarriers could have been due to chance alone.

Additional analyses to assess the potential impact of ApoE ε4 included assessment of CDR-SB using the OTC and uncensored datasets. Based on the OTC dataset, the percent decline versus placebo was -23% for carriers and -19% for noncarriers. Based on the uncensored dataset, the percent decline versus placebo was -26% for carriers and -20% for noncarriers. In the PK/PD modeling of CDR-SB, ApoE ε4 status was not a clinically meaningful covariate, indicating that the exposure-response relationship did not differ between carriers and noncarriers (Section 3.4.3). In addition, mean changes on amyloid SUVR PET showed that in ApoE ε4 noncarriers the degree of reduction in amyloid is not lower than that in ApoE ε4 carriers.

The set of evidence suggests that the pharmacology of aducanumab is similar in ApoE ϵ 4 carriers and noncarriers, and that a treatment effect exists in both ApoE ϵ 4 carriers and noncarriers.

3.4.3. Clinical Pharmacology

A population PK model was developed using data from all clinical studies, comprising approximately 3000 participants and 50,000 concentration observations. This modeling established that the PK of aducanumab was typical for an IgG1 monoclonal antibody. The mean (model estimated) half-life was 24.8 days and the mean time to steady state with every 4 weeks dosing was approximately 4 months. Weight, age, gender, race, and baseline MMSE score had statistically significant influences on aducanumab PK but the impact of these covariate effects was not clinically meaningful.

Nonclinical and early clinical studies established a dose-response effect for aducanumab. Therefore, it is useful to further evaluate the exposure-response relationships and other PK/PD properties from the larger confirmatory dataset in Study 302. To that end, several population exposure-response models were developed.

A population PK/PD model was developed to describe the relationship between exposure to aducanumab and brain amyloid plaque removal, as measured by brain amyloid PET composite SUVR. An exposure-dependent reduction in brain amyloid was observed. The relationship between exposure and amyloid removal is illustrated in [Figure 11](#), where the mean amyloid profiles for the 3 mg/kg, 6 mg/kg, and 10 mg/kg titration regimens differed substantially.

Additionally, the consequences of dose interruptions due to ARIA were assessed. [Figure 12](#) shows the simulated amyloid profiles for participants titrated to 10 mg/kg with monthly dosing over 5 years without dose interruptions, and for participants who have a 12-week interruption in dosing at Month 5. The median number of doses missed by participants with ARIA was 3. Hence, this dosing pattern is representative, although many variations in dosing interruptions and suspensions were observed.

The interruption in dosing had a prolonged influence on brain amyloid plaque removal, including at the Week 78 primary endpoint and beyond. Therefore, in addition to magnitude and duration, the consistency of dosing is also an important factor in the degree of amyloid removal. Age, weight, ApoE ϵ 4 status, and MMSE baseline score had statistically significant influences on the relationship between exposure and amyloid plaque removal, but the impact of these covariate effects was not clinically meaningful.

Figure 11: Population Mean Simulations of Nominal Brain Amyloid PET-Time Profiles by Dose (3, 6, and 10 mg/kg) – Study 302

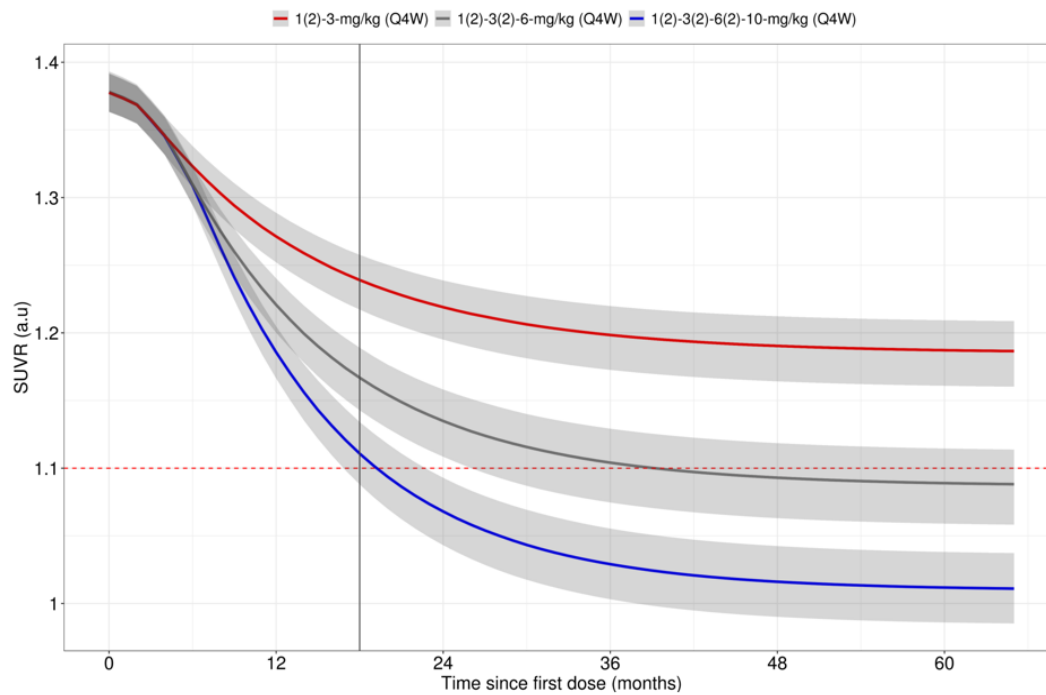
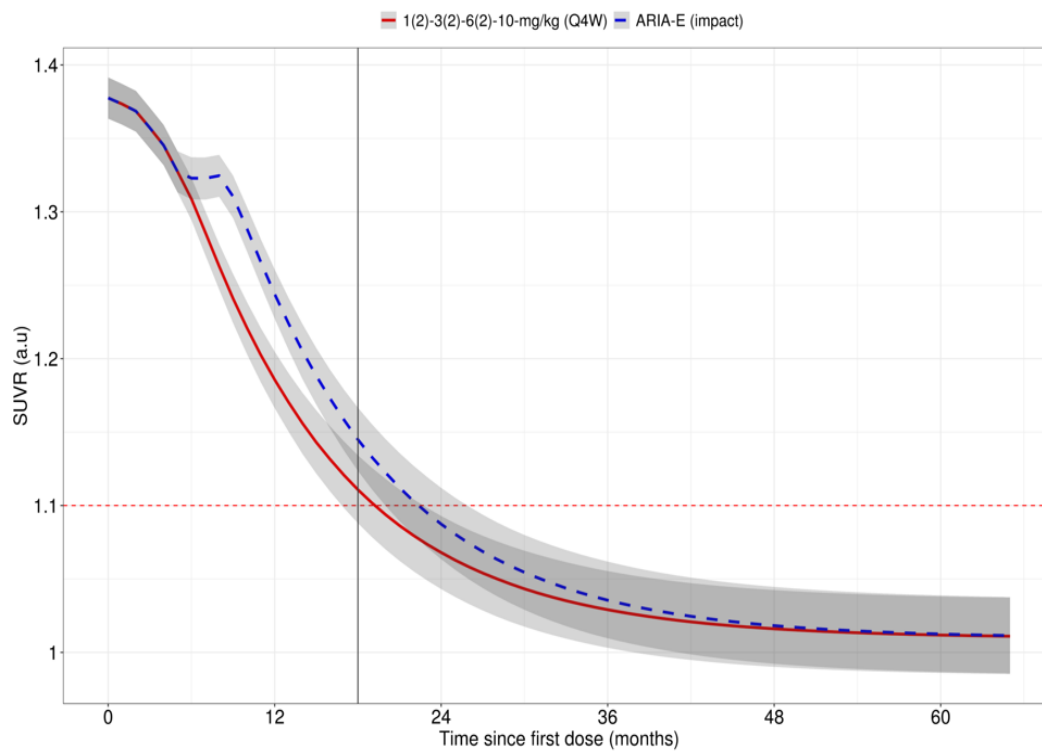


Figure 12: Effect of Dose Interruption due to ARIA on Nominal Brain Amyloid-Time Profile – Study 302



The exposure-response for amyloid PET is useful to understand how aducanumab works in the body.

Exposure-response evaluation for clinical outcomes is also needed to understand the consequences of these effects. Population exposure-response models were developed on the clinical outcomes (CDR-SB, ADCS-ADL-MCI, and ADAS-Cog13) to evaluate these relationships.

A logit-linear disease progression model best described the changes over time in CDR-SB in the placebo group. To account for the considerable variation in disease progression rates, following the logic of [Delor 2013], a mixture model was used that assigned subjects to 1 of 3 disease progression rates.

The exposure-CDR-SB response model, which includes a drug effect component slowing down the disease progression rate, detected a statistically significant response as indicated by the drug effect slope estimate. Simply put, greater exposure to aducanumab led to less clinical decline.

Similar to the exposure-CDR-SB response model, a model relating SUVR to CDR-SB also detected a significant drug effect, confirming that greater amyloid plaque removal leads to greater efficacy. Observed SUVR measurements were available in ~30% of participants. As such, a predicted SUVR value was generated for ~70% of participants (using the exposure-SUVR model) to use as an input for the exposure-response model linking SUVR to CDR-SB. In fact, the latter model has a higher resolution for signal detection than the exposure-response model, since it may better account for the existing pharmacodynamic delays in the system.

3.4.4. Conclusion

Study 302 is a positive study, providing the primary contribution to the substantial evidence of the effectiveness of aducanumab. Evidence of the robustness of these results is provided by the following observations. These conclusions are based on pre-specified analyses.

- Aducanumab high dose significantly slowed the progression of Alzheimer's disease as shown on the primary efficacy endpoint (CDR-SB), all 3 multiplicity-adjusted secondary efficacy endpoints (MMSE, ADAS-Cog13, and ADCS-ADL-MCI), and the tertiary efficacy endpoint (NPI-10). Collectively, on these 5 clinical measures, aducanumab high dose demonstrated a consistent, statistically significant reduction of decline as compared with placebo after 18 months, with treatment differences ranging from -18% (on the MMSE) to -87% (on the NPI-10).
- These treatment effects across 5 independent scales of cognitive, functional, and behavioral decline are clinically meaningful and address an unmet medical need.
- These benefits were observed in an Alzheimer's disease population of limited baseline severity, against a background of minimal expected disease progression over this timeframe.
- The primary and secondary endpoint results were robust to departures from missing data and normality assumptions.
- An intermediate effect was observed in the low-dose group, consistent with a dose-response relationship.

- The exposure-response model for CDR-SB quantifies the exposure-response relationship and detects a significant drug effect for aducanumab.
- The treatment benefit was observed across a broad range of pre-defined, relevant subgroups defined by demography and baseline disease-related characteristics.
- The clinical effects of aducanumab are supported by substantial treatment effects on objective measures of underlying Alzheimer's disease pathophysiology, including brain amyloid pathology as well as downstream markers of intracellular tau pathology and neurodegeneration. The effect of aducanumab on these objective markers was dose dependent and increased with a longer duration of treatment.

The FDA's Position:

FDA agrees that the results of Study 302 are highly persuasive and the study is capable of providing the primary contribution to a demonstration of substantial evidence of effectiveness of aducanumab.

Study 302 is a strongly positive study on multiple distinct and important clinical measures, robust to numerous sensitivity analyses, and supported by well-characterized biomarker data. Beneficial effects on clinical measures are supported by evidence suggesting a dose-response relationship on clinical outcomes and by evidence of a dose- and time-dependent relationship on biomarkers of fundamental Alzheimer's disease pathophysiology, including brain amyloid burden, the primary direct marker of aducanumab's intended mechanistic effect.

Further clinical support for a benefit of aducanumab is found in the presentation of the individual domains of CDR-SB, the primary outcome, which were all consistent with the overall result, and in the statistically significant exploratory analysis of NPI-10, which assesses clinical findings not directly evaluated by the other clinical efficacy outcomes.

Although there was no alpha adjustment for considering the sub-items of the CDR, it is routine to evaluate the movement of components of a scale when the overall scale itself demonstrates statistical significance according to the prespecified testing procedure. It is worth noting that the large absolute effect on NPI-10 may be influenced by non-normal and skewed data and may not be robust to outliers.

3.5. Summary of Study 301

3.5.1. Disposition, Demographics, and Baseline Disease Characteristics

The Applicant's Position

A total of 1647 participants were randomized and dosed in the study. Of these, 88.3% of participants were either still on study or had completed the study as of March 20, 2019. Among those who had the opportunity to complete the Month 18 visit by March 20, 2019, 85.5% completed the study.

Overall, 11.7% of participants had withdrawn from the study by March 20, 2019. Among those with the opportunity to complete the Week 78 visit, 14.5% withdrew from the study prematurely.

Similar to Study 302, the degree of patients prematurely withdrawing from the study was low. In general, Alzheimer's disease studies of this duration have reported a premature withdrawal rate of 20% to 30% [Doody 2014; Salloway 2014; Sevigny 2016]. To address this, the study power calculation included an assumption of a 30% dropout rate.

Demographics and baseline characteristics were similar across treatment groups in the ITT Population. Per protocol design, a majority of ITT participants had a diagnosis of MCI due to Alzheimer's disease (80.4%), while 19.6% of participants had mild Alzheimer's disease dementia (Table 9). Demographics and baseline characteristics were well balanced across treatment groups within the PET substudy.

Table 9: Demographics and Baseline Characteristics – ITT Population

	PBO (N=545)	Low dose (N=547)	High dose (N=555)	Total (N=1647)
Age in years, mean ± SD	69.8±7.72	70.4±6.96	70.0±7.65	70.1±7.45
Sex, Female n (%)	287 (52.7)	284 (51.9)	292 (52.6)	863 (52.4)
Race				
Asian n (%)	55 (10.1)	55 (10.1)	65 (11.7)	175 (10.6)
White n (%)	413 (75.8)	412 (75.3)	413 (74.4)	1238 (75.2)
Education years, mean ± SD	14.7±3.66	14.6±3.77	14.6±3.72	14.6±3.71
Alzheimer's disease medications used, n (%)	299 (54.9)	317 (58.0)	313 (56.4)	929 (56.4)
ApoE ε4, n (%)				
Carriers	376 (69.0)	391 (71.5)	378 (68.1)	1145 (69.5)
Noncarriers	167 (30.6)	156 (28.5)	176 (31.7)	499 (30.3)
Clinical stage, n (%)				
MCI due to Alzheimer's disease	443 (81.3)	440 (80.4)	442 (79.6)	1325 (80.4)
Mild Alzheimer's disease	102 (18.7)	107 (19.6)	113 (20.4)	322 (19.6)
PET SUVR, mean composite ± SD	1.376±0.1990	1.385±0.1859	1.407±0.1786	1.389±0.1885
(n) - PET substudy only	(204)	(198)	(183)	(585)

Abbreviations: PBO = placebo; PET = positron emission tomography; SUVR = standard uptake value ratio.
Data source: 221AD301/CSR/T-DM-PC; 221AD301/CSR/T-BL-CHAR-PC; 221AD301/CSR/T-BL-CHAR-PET-PC

The FDA's Position:

The FDA agrees with the Applicant's presentation of disposition, demographics, and baseline disease characteristics for Study 301.

3.5.2. Efficacy

The Applicant's Position

Study 301 failed to meet its primary and secondary objectives.

Neither treatment group of Study 301 had statistically significant differences from placebo on the primary efficacy endpoint or the secondary efficacy endpoints ([Table 10](#), [Figure 13](#)).

Numeric differences between the low-dose group and the placebo group favored aducanumab ([Table 10](#)) and were similar in magnitude to the numeric differences observed between low dose and placebo in Study 302 ([Table 3](#) and [Table 7](#)).

A statistically significant, dose-dependent reduction in brain amyloid plaque, as measured by PET, was seen in Study 301 for both dose groups at Weeks 26 and 78 ($p < 0.0001$ for all comparisons). Likewise, a statistically significant dose-dependent increase in CSF β -amyloid₁₋₄₂ levels was seen for the high dose ($p = 0.0006$) in Study 301. However, the magnitude of the treatment effect for the high dose on these biomarkers was smaller in Study 301 than in Study 302: 16.5% smaller for PET and 37.7% smaller for CSF β -amyloid₁₋₄₂. Unlike Study 302, the differences from placebo on CSF levels of p-Tau and t-Tau (biomarkers of downstream disease-related pathophysiology and neurodegeneration, respectively) were not significant.

In considering these results, it is notable that across Studies 301 and 302, results for the low dose are similar but results for the high dose are divergent. Analyses conducted to understand why the results of Study 301 are partially discordant with those of Study 302 are described in [Section 3.6](#). Some of the analyses applied to Study 302 to understand robustness of the results are not applicable to Study 301 because Study 301 did not show an effect.

Table 10: Primary and Secondary Endpoints at Week 78 in Study 301: ITT Population

Diff vs PBO ^a (%) p-value			
	PBO decline (N=545)	Low dose (N=547)	High dose (N=555)
CDR-SB	n=333 1.56	n=331 -0.18 (-12%) 0.2250	n=295 0.03 (2%) 0.8330
MMSE	n=332 -3.5	n=334 0.2 (-6%) 0.4795	n=297 -0.1 (3%) 0.8106
ADAS-Cog13	n=331 5.140	n=332 -0.583 (-11%) 0.2536	n=294 -0.588 (-11%) 0.2578
ADCS-ADL-MCI	n=331 -3.8	n=330 0.7 (-18%) 0.1225	n=298 0.7 (-18%) 0.1506

Abbreviations: ADAS-Cog13 = Alzheimer's Disease Assessment Scale - Cognitive Subscale (13 items); ADCS-ADL-MCI = Alzheimer's Disease Cooperative Study - Activities of Daily Living Inventory (Mild Cognitive Impairment version); CDR-SB = Clinical Dementia Rating – Sum of Boxes; MMRM = mixed model for repeated measures; MMSE = Mini-Mental State Examination; PBO = placebo

^aDifference vs placebo at Week 78. Negative percentage means less progression in the treated arm.

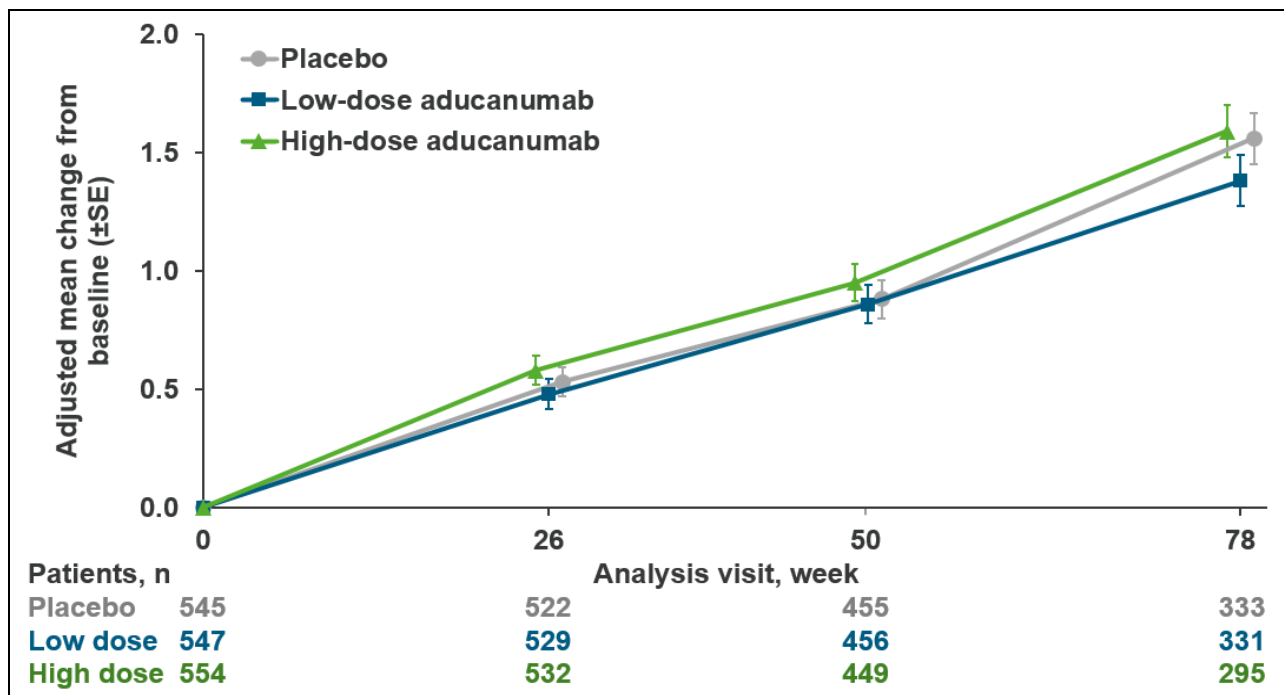
N: numbers of all randomized and dosed participants that were included in the analysis; n: numbers of randomized and dosed subjects with endpoint assessment at Week 78.

Note: The primary analysis was conducted on the ITT population excluding data collected after March 20, 2019.

Note: A MMRM was used as the primary analysis to analyze change from baseline CDR-SB using fixed effects of treatment group, time (categorical), treatment group-by-time interaction, baseline CDR-SB, baseline CDR-SB by time interaction, baseline MMSE, Alzheimer's disease symptomatic medication use at baseline, region, and laboratory ApoE ε4 status. The change from baseline in MMSE, ADAS-Cog13 and ADCS-ADL-MCI scores were also analyzed using MMRM.

Data source: 221AD301/CSR/T-CDR-MMRM-PC; 221AD301/CSR/T-MMSE-MMRM-PC; 221AD301/CSR/T-ADAS-MMRM-PC; 221AD301/CSR/T-ADL-MMRM-PC

Figure 13: Change From Baseline on the CDR-SB Over Time, Study 301



Abbreviations: SE = standard error.

Note: Results were based on an MMRM model, with change from baseline in CDR-SB as dependent variable and with fixed effects of treatment group, categorical visit, treatment-by-visit interaction, baseline CDR-SB, baseline CDR-SB by visit interaction, baseline MMSE, Alzheimer's disease symptomatic medication use at baseline, region, and laboratory ApoE ε4 status.

Data source: 221AD301/CSR/T-CDR-MMRM- PC

The FDA's Position:

The FDA agrees Study 301 is a negative study. At the October 21, 2019, Type C Meeting, the FDA stated that, "available data do not suggest the future use of Study 301 as an efficacy study providing independent evidence of effectiveness supporting the approval of aducanumab for the treatment of Alzheimer's disease."

Upon initial review, the one positive study (Study 302) and one negative study (Study 301) were given equal weight and consideration. Despite divergent outcomes in the primary endpoint, there were some key similarities between the two studies. The low-dose aducanumab treatment arms, while not statistically significant, demonstrated consistent numerical effects favoring aducanumab in the studies. Also, aducanumab produced a time- and dose-dependent reduction in brain amyloid burden. Upon closer review of the individual studies, Study 302 appeared to be a strongly positive study on many distinct clinical measures, robust to numerous sensitivity analyses, and supported by well-characterized biomarker data. In the context of a positive Study 302, the suggestion of a dose-response relationship observed in Study 103, and the numerically favorable results of similar magnitude in the low-dose groups in both studies, the high-dose group in Study 301 tends to stand apart not only for the negative outcome on the primary endpoint, but also the difference in biomarker profiles compared to Study 302.

3.6. Effect of Study 301 on the Persuasiveness of Study 302

The Applicant's Position

3.6.1. Introduction

Sections 3.4 and 3.5 describe the final results for Study 302 and Study 301. The divergence in outcomes was already apparent in the March ITT data, as described in Section 3.3.1. When the March ITT results were presented at the June 2019 Type C meeting, the FDA stated that if the study outcomes were shown to not be confounded by the early termination of the studies, then Study 302 “might also be considered exceptionally persuasive,” and “[t]he submission of a marketing application for aducanumab based primarily on the results of Study 302 as a single positive efficacy study may also be considered. It is possible that the results of Study 301 may have a role in supporting the results of Study 302 or may be understood well enough to be dismissible (i.e., to not represent evidence that the drug is ineffective).”

After the FDA and Biogen jointly concluded that the data from Studies 301 and 302 were suitable for interpretation and valid for further analysis (Sections 3.3.3 and 3.3.4), the FDA and Biogen began a collaborative investigation in order to gain a maximum understanding of the existing data (Section 3.3.2). The stated goals of this collaboration were to achieve “a detailed understanding, informed by plans for further analysis ... of the overall results.” This included both:

- the partially divergent results, specifically, the high dose clinical efficacy and biomarker results, in Studies 301 and 302; *and*
- the results that were consistent between the 2 studies, namely:
 - the low-dose arms in both studies had similar treatment effects across clinical outcomes that were intermediate in magnitude compared with the difference between Study 302 high dose and placebo.
 - the low-dose arms in both studies had similar treatment effects on proximal biomarkers (amyloid PET and CSF β -amyloid₁₋₄₂), and downstream biomarkers specific to Alzheimer's disease (tau PET and CSF p-Tau) and neurodegeneration (CSF t-Tau).
 - the PK and pharmacologic properties of aducanumab did not differ between studies.

3.6.2. Hypotheses Tested

To understand the similarities and differences between Studies 301 and 302 described above, with a view to assessing how Study 301 affected the persuasiveness of Study 302, Biogen and the FDA developed specific hypotheses that were specified in analysis plans before the analyses began. The analyses were exploratory (post hoc), but they were undertaken with rigor to the greatest extent possible.

The overarching hypothesis was that if aducanumab were an effective treatment for Alzheimer's disease, then there should be patients in Study 302, the positive study, that drive the overall effect. Additionally, in such a situation, a set of patients in Study 301, the failed study, with those

same characteristics should show a response similar to that seen in Study 302. Specific hypotheses within this framework were as follows:

1. Whether there were differences in dosing between the 2 studies, and whether such differences, if they existed, contributed to the differences in the results was a focus of investigation because:
 - a. Protocols were amended during the studies to enable more subjects to achieve the target dose of 10 mg/kg. Specifically, as described in Section 3.1.3, the protocols for Studies 301 and 302 were amended twice during the studies with the effect of increasing exposure to 10 mg/kg.
 - Subjects who suspended dosing due to ARIA could (after resolution of findings) resume dosing at the same dose and continue titration to the target dose (rather than resume at the next lower dose with no further increases in dose permitted).
 - The target dose for ApoE ε4 carriers in the high-dose regimen was increased from 6 mg/kg to 10 mg/kg.
 - b. Due to the differences in enrollment timing, the protocol amendments influenced more patients in Study 302 than Study 301 because Study 302 started later; therefore, more patients were enrolled after the amendment.
 - c. Nonclinical and previous clinical studies for aducanumab showed a clear dose-exposure response. For example, in Study 103, reduction in brain amyloid pathology was closely correlated with dose, and a dose response was evident also on the clinical measures (Sections 3.1.2 and 3.7).

It was also prospectively recognized that multiple factors could potentially influence efficacy outcomes. These additional factors, and hypotheses that were specified to address them, included:

2. Imbalances in prognostic factors or treatment effect modifiers, which are always a consideration in divergent results. Therefore, whether imbalances in demographic and disease characteristics contributed to the divergence in results was investigated.
3. The effects of ARIA on efficacy outcomes was another area of investigation, because ARIA is a frequent adverse event with the potential to impact outcomes due to (1) the management of ARIA through temporary dose suspension and (2) the potential for functional unblinding.
4. Routine statistical diagnostic tests of the primary analysis indicated the assumption of normal distributed outcomes had been substantially violated. Therefore, the consequences of this non-normality were investigated.

In summary, 4 areas were investigated to understand the partially conflicting results in Study 302 and 301: 1) dose, 2) baseline characteristics, 3) ARIA, and 4) non-normality of the clinical data. These are addressed in Section 3.6.3.

These investigations relied on post hoc analyses, for which there are well known limitations. To mitigate these limitations, hypotheses were specified before analysis, and multiple analytic approaches with differing strengths, limitations, and assumptions were used.

These post hoc analyses were conducted to understand the differences in results between Studies 301 and 302. None of the conclusions on primary and secondary efficacy outcomes in Studies 103, 301, and 302 were from post-hoc analyses.

The FDA's Position:

In general, the FDA agrees with the Applicant's characterization of the hypotheses tested. If aducanumab is effective, it follows that Study 302 is a positive study and that there would be patients in Study 301 who, based on certain characteristics, should show response of similar character to patients in Study 302. Given the possibility that aducanumab is an effective drug for a disease with an enormous unmet medical need, it was agreed by the Applicant and Division at the June 14, 2019, Type C Meeting that extensive resources should be brought to bear on achieving a maximum understanding of the existing data.

By their nature, these analyses were post hoc and exploratory and therefore carry with them the appropriate caveats and caution in their interpretation. To address these concerns, any exploration of the data was to be rigorous, limited in scope, and based on predetermined and well-defined hypotheses. To the maximum degree possible, the analyses were pre-specified. An important distinction is that these analyses were not aimed at obtaining independent support from Study 301. Study 301 was a negative study. The purpose of these analyses is to provide maximum understanding of the partially discordant results and to determine if this understanding precludes independent consideration of Study 302.

3.6.3. Results

Overview

Each of the areas investigated by Biogen in collaboration with FDA and the associated conclusion are listed in [Table 11](#). Brief summaries are provided for factors not having a meaningful impact on divergence in results. More detailed summaries are provided for factors found to be influential.

Table 11: Overview of Hypotheses and Outcomes for Why Results for the High-Dose Groups Were Discordant Between Studies 301 and 302

1	Did differences in dosing between the 2 studies contribute to the difference in results?	Was a contributing factor
2	Did the violation of statistical assumptions regarding non-normality have an impact on the results?	Was a contributing factor
3	Were there imbalances in demographic and disease characteristics that contributed to the difference in results?	Not a factor
4	Were there differences in the incidence, severity, association with symptoms, or management of ARIA with potential implications of functional unblinding that contributed to the difference in results?	Not a factor

Demographics and disease characteristics

For demographic and disease characteristics to contribute to the difference in results between the high-dose arms of Studies 301 and 302, these factors must both (1) have an appreciable influence on disease progression and (2) be substantially unbalanced between treatment arms.

Extensive statistical modeling was done to better understand factors that influenced disease progression. These analyses determined that baseline demographic and illness characteristics had a minimal influence on disease progression.

Furthermore, differences between treatment arms and studies in demographic and baseline disease characteristics were minimal. This was as expected in clinical trials with such large sample sizes and randomization as in Studies 301 and 302. The small differences found did not contribute to the divergence in results.

The FDA's Position:

Prior to the June 14, 2019, Type C Meeting, the FDA requested that the Applicant examine whether there were differences in demographics and baseline disease characteristics between the studies. The FDA agrees that any differences are minor and do not appear to have a meaningful impact on the outcome of the studies.

ARIA

Several considerations were addressed in assessing whether ARIA contributed to the differences between Studies 301 and 302:

- Symptoms of ARIA could theoretically negatively impact clinical outcomes.
- ARIA could be a marker of greater clinical effect if it were associated with greater amyloid reduction (which has not been established).
- Dose suspensions due to ARIA could result in reduced clinical effect.
- ARIA could influence clinical outcomes through functional unblinding of participants with ARIA.

Participants with ARIA were subject to additional follow-up MRIs and dose suspension. Given that the majority of ARIA cases were expected to be on active treatment, ARIA management could therefore potentially lead to functional unblinding. This potential was minimized by requiring efficacy raters to be different from those assessing and monitoring safety and by requiring that they remain blinded to all other clinical assessments.

The incidence, severity, and symptoms of ARIA were similar in Studies 301 and 302. Therefore, if ARIA influenced outcomes, that influence should be similar in the two studies and not contribute to the difference in results between studies. ARIA-E events typically occurred early in treatment, with over 70% occurring during the first 8 doses. Incidence of dose suspension, dose reduction, and discontinuation of study treatment due to first ARIA-E events was similar in Studies 301 and 302. For a complete discussion of ARIA incidence, detection, and management, see Section 4.3.

To evaluate the potential for functional unblinding due to ARIA, results based on all observations were compared with an otherwise identical analysis in which post-ARIA observations were removed. These analyses of the primary and secondary endpoints yielded outcomes similar to the primary analysis that included all data (Table 12).

The treatment contrast for the high-dose group in Study 302 was greater after excluding post-ARIA observations (33% slowing in decline vs. 22%). Therefore, the statistically significant difference between high dose and placebo on the primary endpoint in Study 302 was not influenced by systematic bias from functional unblinding associated with ARIA.

Table 12: Change from Baseline in CDR-SB at Week 78: All Participants With and Without Post-ARIA Observations Excluded, Studies 301 and 302, ITT Dataset

	Study 301				Study 302	
	PBO decline (N=545)	diff vs. PBO ^a (%)	diff vs. PBO ^a (%)	PBO decline (N=548)	diff vs. PBO ^a (%)	diff vs. PBO ^a (%)
		Low dose (N=547)	High dose (N=555)		Low dose (N=543)	High dose (N=547)
All participants – all observations	n=333 1.56	n=331 -0.18 (-12%)	n=295 0.03 (2%)	n=288 1.74	n=290 -0.26 (-15%)	n=299 -0.39 (-22%)
All participants – excluding post-ARIA observations	n=298 1.55	n=240 -0.11 (-7%)	n=181 0.00 (0%)	n=254 1.72	n=194 -0.19 (-11%)	n=172 -0.57 (-33%)

Abbreviations: diff = difference; PBO = placebo.

a: difference vs placebo at Week 78. Negative percentage means less progression in the treated arm.

N: number of randomized and dosed participants included in the analysis.

n: numbers of randomized and dosed participants with primary endpoint assessment at Week 78.

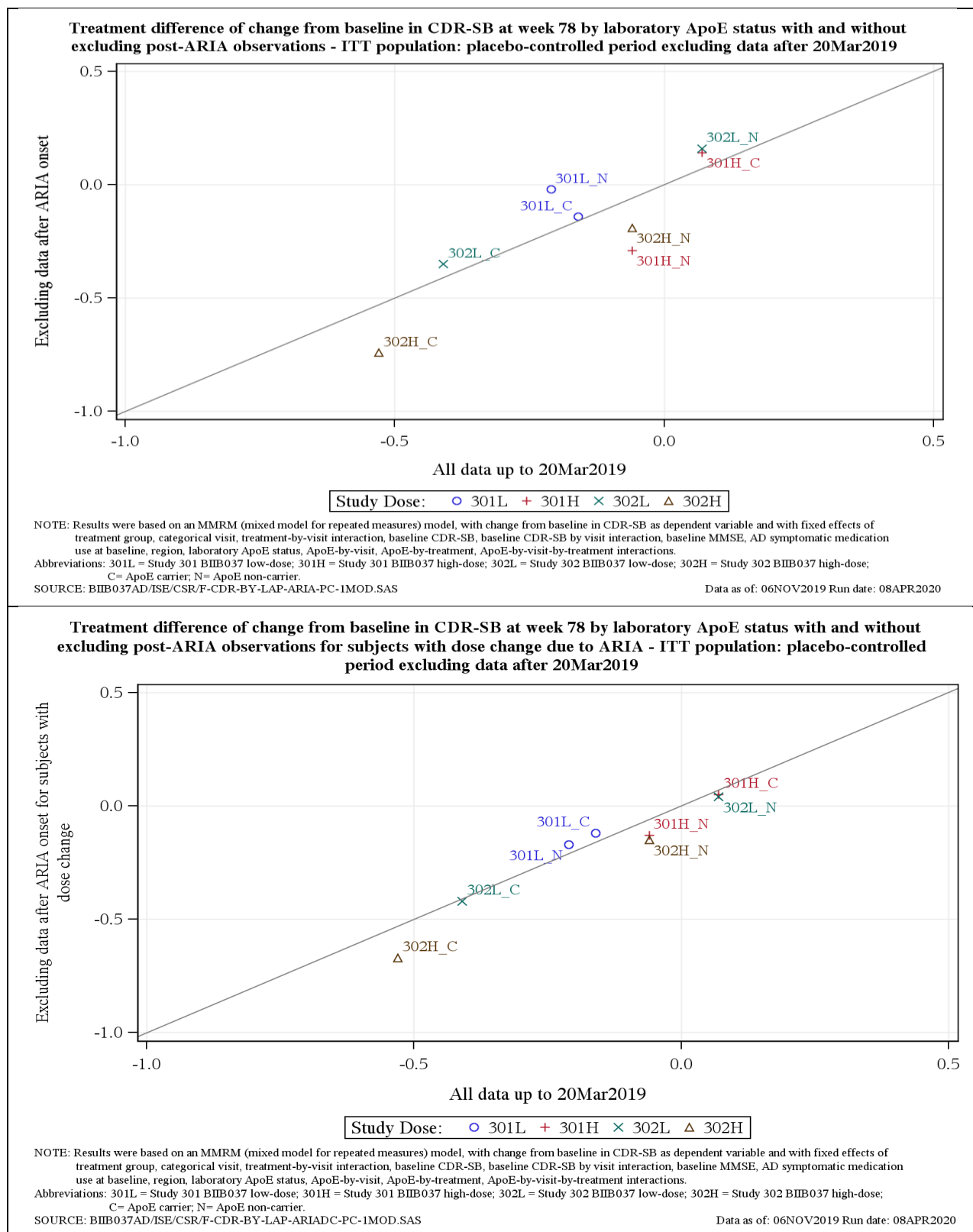
Data source: 221AD301/CSR/T-CDR-MMRM-PC, 221AD302/CSR/T-CDR-MMRM-PC ISE/CSR/T-CDR-CNSR-ARIA-PC

For a more granular comparison, prespecified groupings defined by dose, study, and ApoE ε4 status, with and without excluding all post-ARIA observations, and with and without excluding post-ARIA observations for participants with a dosing change due to ARIA were examined (Figure 14). The X axis is the treatment contrast from all data and the Y axis is the treatment contrast after excluding relevant observations. The data points are scattered evenly above and

below the line of unity, indicating random variability and not bias (data points reflecting bias would be consistently below or above the unity line).

Again, no evidence was found for systematic bias from functional unblinding due to ARIA or ARIA management.

Figure 14: Treatment Difference for Change on CDR-SB at Week 78 by ApoE ε4 Status with and Without Post-ARIA Observations



The FDA's Position:

FDA agrees a systematic bias due to functional unblinding caused by ARIA is not apparent. Some degree of functional unblinding was inevitable, but the Applicant took steps in the protocol to minimize its impact (i.e., independent and blinded raters). The analysis presented by the Applicant must be based on a post-randomization factor (i.e., occurrence of ARIA), as baseline factors do not reliably predict the occurrence of ARIA. The results of this analysis do not indicate a systematic bias introduced by ARIA.

Non-normal distribution of the data

Some degree of non-normality in change from baseline scores on CDR-SB was anticipated during study planning. However, standard statistical diagnostics showed that the statistical assumption for normality of residuals had been substantially violated. The data were “right skewed,” with a small number of participants having unusually rapid decline. This finding led to the investigation of the influence of rapid progressors on results.

The anticipated decline in an early symptomatic Alzheimer's disease population on CDR-SB in a 78-week period is 1-2 points [Coric 2012; Egan 2019; Ostrowitzki 2017]. Alzheimer's disease patients who, after diagnosis, have a very precipitous decline in their symptoms are also seen in clinical practice [Abu-Rumeileh 2018; Llorens 2016; Schmidt 2012]. However, factors prognostic of rapid decline have not been identified.

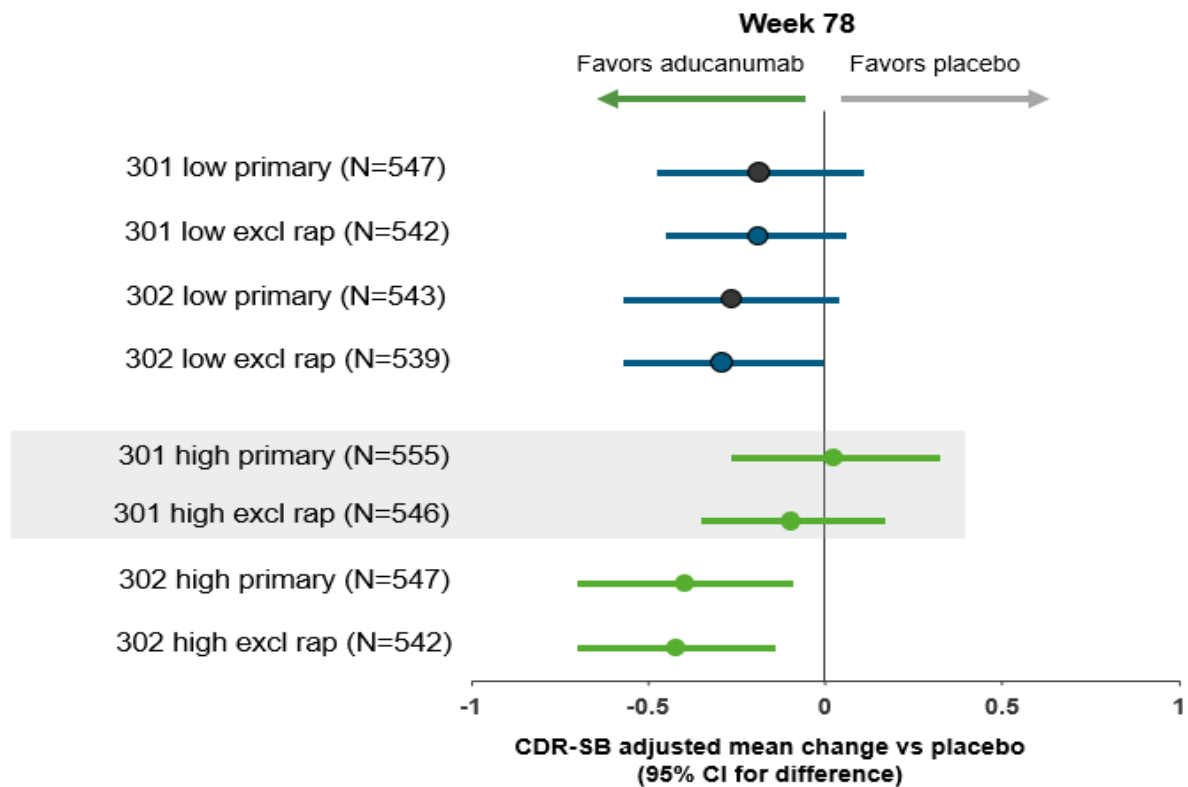
There is no precedent in clinical trials for defining rapid progressors. A cut-off of change > 8 points on CDR-SB was chosen as the primary definition of rapid progression. This is 4-fold the upper limit of the anticipated mean decline in this study population and, therefore, is an extreme change compared to typical progression. Additionally, a >8 point change on the CDR-SB is consistent with the definition for statistical outliers [Tukey 1977]. Thirty-one participants were classified as rapid progressors. Participants categorized as rapid progressors based on CDR-SB also had rapid declines in the other endpoints.

Except for the high-dose group in Study 301, each treatment group in each study (N > 500), including the placebo groups, had 4 or 5 participants with this degree of progression; the high-dose group in Study 301 had 9 (twice as many as the other groups). Given the divergence between the 2 studies in the results in the high-dose groups, the imbalance in the number of rapid progressing participants in the high-dose arms was a focus of the collaborative investigation.

To isolate the effect rapid progressors had on estimates of the treatment effect, results from the primary and secondary endpoints were compared before and after excluding observations from the rapid progressors. The intent of this analysis was to explain the difference in results between studies, not to find a more suitable primary analysis. Excluding rapid progressors had a notably greater impact on results in the high-dose group of Study 301 than in the other groups (Figure 15).

Results using other cut-offs to identify rapid progressors (6-, 7-, 9-, and 10-point changes in CDR-SB) yielded similar results, where excluding rapid progressors had a larger influence on the Study 301 high-dose group. Results were also consistent across secondary outcomes.

Figure 15: CDR-SB Results With and Without Rapid Progressors, Studies 301 and 302



Abbreviations: CI = confidence interval; excl = excluding; rap = rapid progressors.

N: numbers of all randomized and dosed participants that were included in the analysis.

Note: Results were based on an MMRM model, with change from baseline in CDR-SB as dependent variable and with fixed effects of treatment group, categorical visit, treatment-by-visit interaction, baseline CDR-SB, baseline CDR-SB by visit interaction, baseline MMSE, Alzheimer's disease symptomatic medication use at baseline, region, and laboratory ApoE status.

Data source: 221AD301/CSR/T-CDR-MMRM-PC; 221AD302/CSR/T-CDR-MMRM-PC; ISE/CSR/T-CDR-MMRM-CDR8-PC.

Characteristics of rapid progression

With only 31 participants in the Phase 3 studies classified as rapid progressors, there was limited statistical power to identify characteristics of rapid progressors that differ from other participants.

However, rapid progressors were not found to differ from the study sample based on any demographic or baseline disease characteristic, including ApoE ϵ 4 status, comorbidities, concomitant medications, or the incidence of AEs, including ARIA.

Neither randomized treatment group nor actual dosing was a factor in rapid progression. As noted above, except for the high-dose group in Study 301, each of the treatment arms – including placebo – had 4 or 5 rapid progressors (with the high-dose group of Study 301 having 9). Since rapid progression was evident early in treatment, with large declines already present at the Week 26 assessment, for those in the high-dose arms, participants were not yet at the highest dose of 10 mg/kg. Five of the 9 rapid progressors in the Study 301 high-dose arm had 6 or fewer doses of 10 mg/kg, and only 1 participant had >10 doses of 10 mg/kg throughout their treatment. Therefore, it is unlikely that rapid progression was related to aducanumab treatment.

Rapidly progressing participants were not differentiated by baseline clinical characteristics, and they did not differ from the study population as a whole in the prevalence of psychiatric disorders at study entry, the incidence of psychiatric-related AEs during the study, or in changes of medications that influence acute mental status during the trial.

Therefore, these findings are consistent with published studies in suggesting that there is no single clinical feature that consistently differentiates rapidly progressing participants from normally progressing Alzheimer's disease patients [[Abu-Rumeileh 2018](#)].

The FDA's Position:

FDA agrees that (1) small imbalances in the number of rapid progressors can have a relatively large impact on the magnitude of the primary and secondary endpoints using the primary analysis method and (2) the high dose arm in Study 301 was disproportionately affected by such an imbalance.

An additional consideration for investigating the distribution of the treatment effect was the finding from the futility analysis dataset in which the high dose arm in Study 301 appeared to show greater cognitive decline than the placebo arm. A credible indication that aducanumab treatment in Study 301 was accelerating cognitive decline would undermine the results of Study 302. An exploratory analysis based on the observation of skewed data revealed the existence of a small number of rapid progressors, consistent with reports in the literature and observations in other large datasets. When the small number of rapid progressors are removed, however, the point estimate of the treatment effect for the high dose in Study 301 favors aducanumab, a finding consistent with every other treatment arm across the aducanumab development program. This is not to say that the imbalance in rapid progressors that disproportionately affected the high dose arm of Study 301 is the cause for the failure of Study 301 on its primary outcome. As noted previously, this is an exploratory post hoc analysis that must be viewed with that limitation in mind. Instead of removing all the rapid progressors, removing only the single worst high dose rapid progressor does not result in the point estimate of the treatment effect favoring aducanumab, though this approach retains a residual imbalance in the remaining rapid progressors. These explorations suggest that a single outlier has less influence than a larger number, which is to be expected, but a relatively larger number requires consideration of whether a systemic issue contributed to the imbalance. There was no prespecified analysis robust to outliers on which to rely. An additional exploratory post hoc analysis that is robust to outliers results in an estimate of the treatment effect for the high dose that is essentially the same as placebo. There is also no indication from Study 301 that aducanumab is the cause of the rapid progression observed in subjects receiving the high dose. With no external data or data from the study itself able to differentiate rapid progressors from the remainder of the population, and no indication that aducanumab is the cause of rapid progression, a systemic issue does not appear responsible.

In sum, the high dose arm in Study 301 was disproportionately affected by an imbalance of rapid progressors, with essentially twice the number of such patients as were present in all other treatment groups. The total number of rapid progressors was small, such that removing them still leaves a large treatment population on which to base an exploratory analysis of treatment effect. This analysis resulted in a point estimate of the treatment effect for the high dose in Study 301 that favors aducanumab, indicating that small imbalances in the number of rapid progressors can have a relatively large impact on the magnitude of the primary and secondary endpoints using the primary analysis method.

Influence of Dosing

Basis for possible dosing differences between studies

Two amendments to the Phase 3 protocols influenced the ability of participants randomized to aducanumab high dose to have consistent exposure to the target 10 mg/kg dose:

- Protocol Version 3 changed dosing management of ARIA, such that participants who met prespecified criteria for dose suspension due to ARIA could now (1) resume treatment at the same dose (rather than the lower dose) and (2) continue titration to their target dose (rather than remain at the lower dose). This amendment increased the likelihood that such participants would reach and maintain their target dose.
- Protocol Version 4 changed the target dose for ApoE ϵ 4 carriers (approximately two-thirds of each study population) in the high-dose group from 6 mg/kg to 10 mg/kg (10 mg/kg was the target dose for high-dose noncarriers throughout the studies).

While Studies 301 and 302 were identical in design, they started one month apart, with Study 301 starting first and remaining ahead in enrollment. Therefore, Protocol Versions 3 and 4 influenced more participants in Study 302 (81.3% and 55.9%, respectively) than in Study 301 (73.4% and 49.1%, respectively).

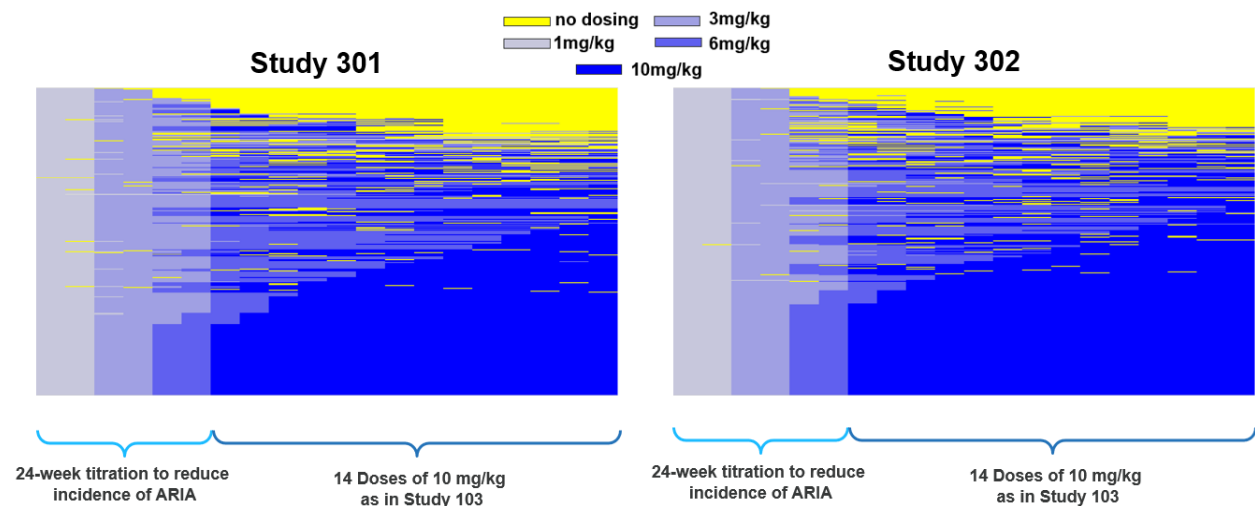
Differences in dosing between studies and within studies over time

Figure 16 shows dosing heat maps that illustrate individual dosing profiles for individuals in the high-dose group of each study who had the opportunity to reach Week 78 by March 20, 2019. On the heat maps, dark blue represents the 10 mg/kg dose, lower doses are light blue, and no dose is yellow.

The heat maps illustrate the following important factors:

- Even among the individuals who had the opportunity to complete the 78 weeks, both studies have a low level of participants who received the full quotient of intended target dose of 10 mg/kg; only 22.3% of participants in Study 301 and 28.8% of participants in Study 302 received all 14 doses of 10 mg/kg.
- Considerable heterogeneity in individual participant dosing was present.

Figure 16: Individual Dosing Profiles for High-Dose Participants in Studies 301 and 302



Data source: 221AD301/CSR/L-EX; 221AD302/CSR/L-EX

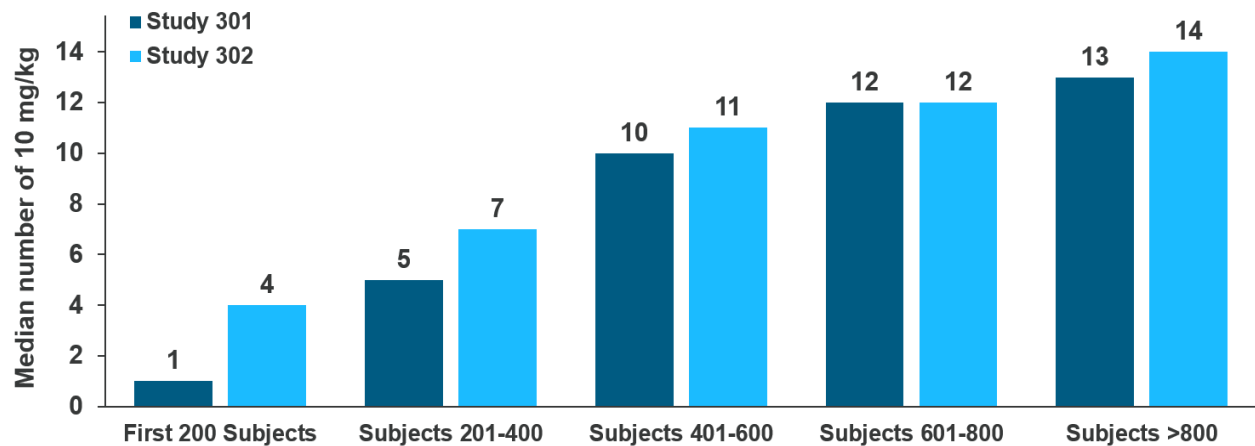
Some of the heterogeneity in individual dosing is due to ARIA, while considerably more is directly due to protocol changes. The expected cumulative doses at Week 78 with no interruption in dosing were 160 mg/kg and 98 mg/kg for participants randomized to groups with a maximum dose of 10 mg/kg and 6 mg/kg, respectively. Therefore, before implementation of Protocol Version 4, the maximum cumulative dose for ApoE ϵ 4 carriers (approximately two thirds of the population) randomized to the high dose (6 mg/kg) was 98 mg/kg. Therefore, mean and median cumulative dose for Studies 301 and 302 high-dose arms increased over time following implementation of the protocol amendment.

To further illustrate this change in dosing over the duration of the study, the median number of 10 mg/kg doses by enrollment cohorts of 200 participants is plotted in Figure 17. As can be seen in the figure, in the first 200-participant cohort, the median number of 10 mg/kg doses in Study 301 was 1, compared with a median of 4 in Study 302 (out of the 14 intended 10 mg/kg doses after the titration period). In the last 200 enrolled, the median number of 10 mg/kg doses was 13 and 14 in Studies 301 and 302, respectively.

Note that differences in dosing between the 2 studies are attributable to differences in timing of the implementation of Protocol Version 4 between the 2 studies as well as differences in stratification based on ApoE ϵ 4 carriage over time.

Overall, exposure to 10 mg/kg dosing continued to increase over time in both studies, while dosing was greater in Study 302 throughout the studies.

Figure 17: Median Number of 10 mg/kg Doses by Enrollment Time



Data source: ISE/ADHOC/T-EX-SUM-BY-ENROLL-PC-OTC

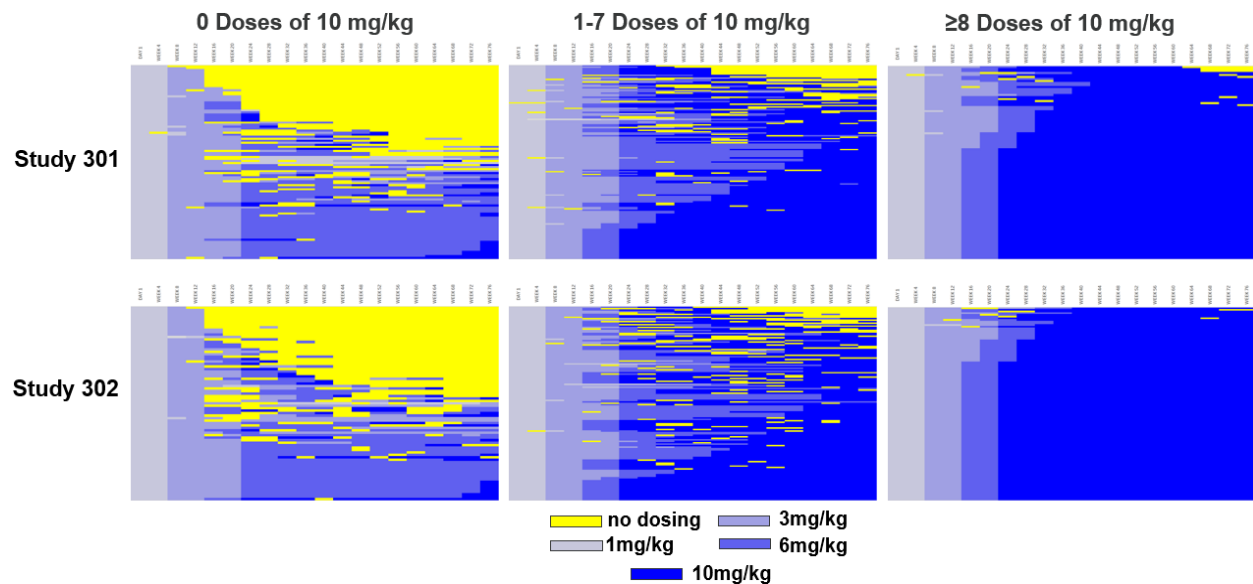
Another way to understand the heterogeneity in dosing in the high-dose groups is to assess individual participant dosing profiles. As detailed in a subsequent section, dosing was categorized in a number of ways by subsetting the high-dose arms into groups representing: 1) high exposure to 10 mg/kg dosing; 2) intermediate exposure to 10 mg/kg dosing, and 3) no (or very limited) exposure to 10 mg/kg dosing.

Individual participant dosing heat maps for three such categories in Studies 301 and 302 are presented in [Figure 18](#). The categorizations were (after titration):

- **High:** ≥ 8 uninterrupted doses of 10 mg/kg at steady state
- **Intermediate:** 1-7 uninterrupted doses of 10 mg/kg at steady state
- **Zero:** no doses of 10 mg/kg at steady state

In Study 302, the proportion of participants in the high subset of the high-dose arm (≥ 8 uninterrupted doses of 10 mg/kg at steady state) (35.0%) was higher than in Study 301 (28.8%), while in Study 301, the proportion of participants with zero doses of 10 mg/kg at steady state was higher (28.8%) than in Study 302 (24.2%). In the intermediate subset of the high-dose arm (1-7 uninterrupted doses of 10 mg/kg at steady state), the heat maps in [Figure 18](#) show considerable heterogeneity in dosing, i.e., there are many dosing patterns present in this category. Greater heterogeneity in dosing profiles is expected to increase potential for heterogeneity in results.

Figure 18: Dosing Profiles (Heat Maps) for Zero, Intermediate, and High/Consistent Dosing Subgroups Within the High-Dose Groups of Studies 301 and 302



Data source: 221AD301/CSR/L-EX; 221AD302/CSR/L-EX, ISE/ADHOC/L-SUBJ-SS10-PSMB-PC-OTC

Similarities and differences in outcomes between studies

To assess the impact of differences in dosing over time on clinical outcomes in the high-dose arms between studies and within studies, two approaches were used to create various subsets:

- Subsets of the high-dose arm defined by randomization (ApoE ε4 carriers / noncarriers) and pre-Protocol Version 4 / post-Protocol Version 4.
- Subsets of the high-dose arm formed by the total mg/kg of drug taken and by how many 10 mg/kg doses each participant had taken during the 78-week treatment period.

Additional details on the analyses are summarized in [Table 13](#).

An advantage of using groups formed by randomization (Approach 1) was that it preserved randomization, and therefore standard statistical analyses could be used. A disadvantage of the randomized group approach was that it did not consider the actual dosing that each participant received.

Forming subgroups of the high-dose group based on actual dosing (Approach 2) enables consideration of actual dosing but does not preserve randomization because groups are formed based on post-randomization outcomes. Therefore, for the subsets of the high-dose group based on actual dosing, analytic approaches commonly used in observational studies were used to select appropriately matched subgroups of placebo participants as the comparison group.

Table 13: Analyses to Investigate Efficacy Outcomes in the High-Dose Group by Dosing Subgroups

<p>Approach 1: Subgroups defined by randomization (OTC data)</p> <ul style="list-style-type: none"> 8 groups per study created by stratified randomization (further subdivided by pre- or post-Protocol Version 4): <ul style="list-style-type: none"> Dose: Low/High ApoE ϵ4: Carrier / Noncarrier Differences versus placebo on the 4 primary and secondary endpoints were compared, with and without rapid progressors excluded. Because groups are created by randomization, placebo subgroups are appropriately matched to the corresponding aducanumab group.
<p>Approach 2: Subgroups based on actual dosing profiles (OTC data)</p> <ul style="list-style-type: none"> 3 methods to categorize individual participant dosing: <ul style="list-style-type: none"> Number of uninterrupted steady-state doses of 10 mg/kg Number of doses of 10 mg/kg without regard for interruption Cumulative dose (sum of all doses taken) Propensity score matching was used for each method of categorizing dosing to find appropriately matched placebo groups (primary approach, emphasized controlling bias). <ul style="list-style-type: none"> Analyses were conducted on the primary and 3 secondary endpoints, with and without rapid progressors excluded. Two additional approaches were applied on selected dosing subgroups to evaluate the robustness of the primary approach: <ul style="list-style-type: none"> Unmatched: included all randomized placebo participants Matching based on ApoE ϵ4 status and enrollment timing

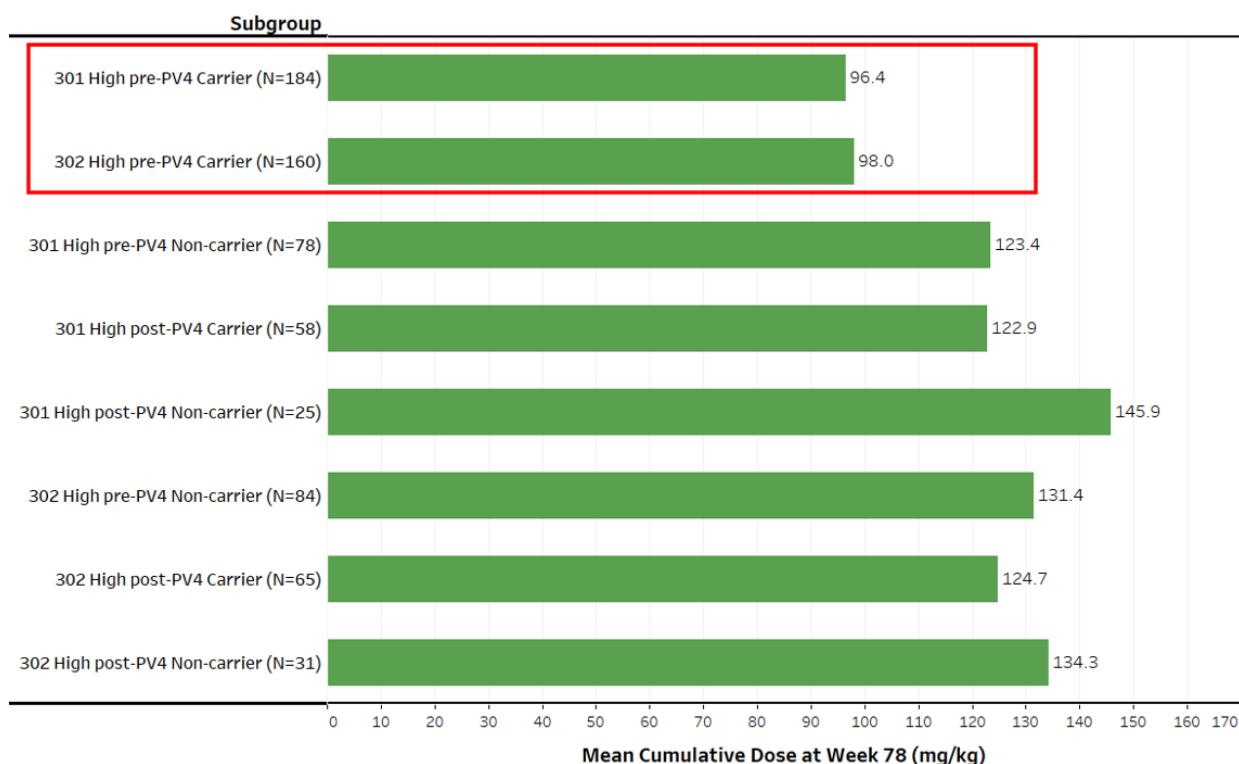
Representative results of the analyses conducted in Approach 1 and Approach 2 are presented below.

Approach 1: Subgroups defined by randomization

The expected cumulative doses at Week 78 with no interruption in dosing were 160 mg/kg and 98 mg/kg for participants randomized to groups with a maximum dose of 10 mg/kg and 6 mg/kg, respectively.

ApoE ϵ 4 carriers in both studies in the high-dose group enrolled prior to Protocol Version 4 did not have the opportunity to receive all 14 doses of 10 mg/kg and had exposures of 96.4 and 98.0 mg/kg in Study 301 and Study 302, respectively. These exposures were at or close to the expected cumulative doses of 98 mg/kg for participants randomized to a maximum dose of 6 mg/kg, but are appreciably lower (by approximately 20% to 34%) than the mean cumulative dose of the other high-dose subgroups (Figure 19, top 2 bars). Two thirds of the study sample are ApoE ϵ 4 carriers; therefore, approximately two-thirds of participants randomized to the high-dose group prior to the protocol amendments had lower cumulative doses.

Figure 19: Heterogeneity in Dosing Subsets of the High-Dose Group by Mean Cumulative Dose, Studies 301 and 302



Abbreviations: PV4 = protocol version 4.

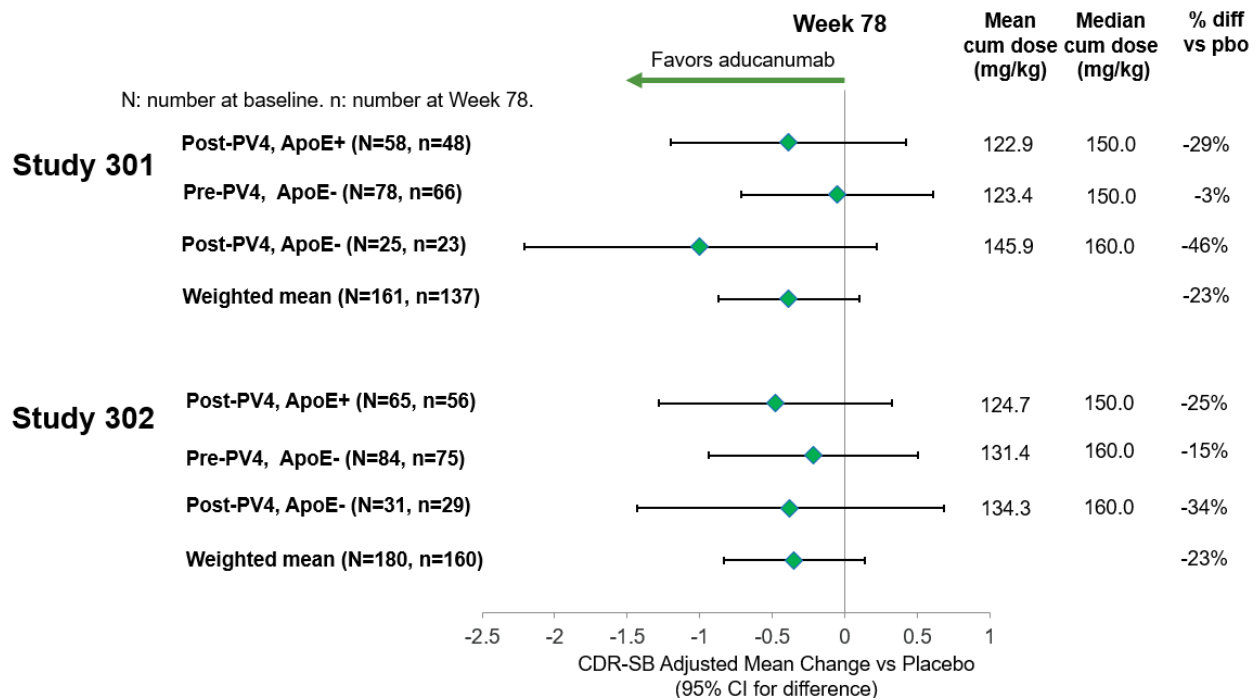
Note: High-dose pre-PV4 carriers had the opportunity to receive 0-13 doses of 10 mg/kg (after titration). High-dose post-PV4 carriers had the opportunity to receive all 14 doses of 10 mg/kg (after titration). High-dose noncarriers had the opportunity to receive all 14 doses of 10 mg/kg under all protocol versions.

Data source: ISE/CSR/T-EX-SUM-BY-PV4LAP-PC-OTC

Results from the three subgroups in each study that did have the opportunity for 14 doses of 10 mg/kg are summarized in [Figure 20](#). The percent weighted mean treatment differences between high dose and placebo based on combining the three subgroups within each study were similar in Studies 301 and 302 (-23% in both studies). Moreover, results for each of the individual subgroups were also similar for the two studies.

In conclusion, in Study 301, participants randomized with the opportunity for a full 14 doses of 10 mg/kg had clinical outcomes similar to those in Study 302.

Figure 20: Comparison of CDR-SB in Subsets with Opportunity to Receive 14 Doses of 10 mg/kg – OTC Population



Abbreviations: CI = confidence interval; cum = cumulative; diff = difference; pbo = placebo; MMRM = mixed model for repeated measures; PV4 = Protocol Version 4

Note: Results were based on an MMRM model, with change from baseline in CDR-SB as dependent variable and with fixed effects of treatment group, categorical visit, laboratory ApoE ε4 status, PV4 consenting status, and two-way, three-way and four-way interactions of these four covariates, and additional fixed effects of baseline CDR-SB, baseline CDR-SB by visit interaction, baseline MMSE, Alzheimer's disease symptomatic medication use at baseline and region.

Data source: ISE/CSR/T-CDR-BY-PV4LAP-PC-OTC, ISE/CSR/T-CDR-BY-PV4LAP-PC-OTC-HEXP, ISE/CSR/T-EX-SUM-BY-PV4LAP-PC-OTC

Approach 2: Subgroups based on actual dosing

In this approach, dosing *within* the high-dose groups was categorized by creating subsets. In the results representative of this approach presented below, the categorizations were (after titration):

- **High/Consistent:** ≥ 8 uninterrupted doses of 10 mg/kg at steady state
- **Intermediate:** 1-7 uninterrupted doses of 10 mg/kg at steady state
- **Zero:** no doses of 10 mg/kg at steady state

Individual participant dosing heat maps for these three dosing categories in Studies 301 and 302 were presented above in [Figure 18](#).

Results on the CDR-SB for these dosing categories are presented in [Figure 21](#).

- Participants in the high-dose subset with ≥ 8 uninterrupted doses of 10 mg/kg at steady state had reductions versus placebo of approximately 35% in both Study 301 and Study 302. i.e., in Study 301 despite the overall high-dose group failing to meet clinical outcomes, a subset of participants with high and consistent exposure to 10 mg/kg aducanumab had outcomes similar to those in Study 302.

- Participants in the subset of the high-dose groups with no exposure to 10 mg/kg at steady state had minimal difference from placebo in both Study 301 and Study 302.

Both subsets demonstrate that in Study 301 and Study 302 similar exposures produce similar outcomes, and that exposure to 10 mg/kg is important for efficacy.

In each of these subsets, there are different proportions of participants in both studies. On balance, these favored higher exposure to 10 mg/kg in Study 302: In the high dosing subset with efficacy outcomes of 35% reduction, Study 302 had more participants than Study 301 (35.0% vs. 28.8%). Additionally, Study 302 had fewer participants in the no exposure to 10 mg/kg dosing category than Study 301 (24.2% vs. 28.8%).

In contrast to the similarity in outcomes seen in the high and zero exposure subsets of 10 mg/kg, in the intermediate subset of the high-dose group (1-7 uninterrupted doses of 10 mg/kg at steady state), participants had reduced decline versus placebo in Study 302 but not in Study 301. In this intermediate subset of the high-dose group, both the mean cumulative dose and the number of participants are similar in both Study 301 and Study 302. Excluding rapid progressors reduces the differences between studies in the results for the intermediate 10 mg/kg dosing subset in Study 301; however, some difference on CDR-SB remain ([Figure 22](#)).

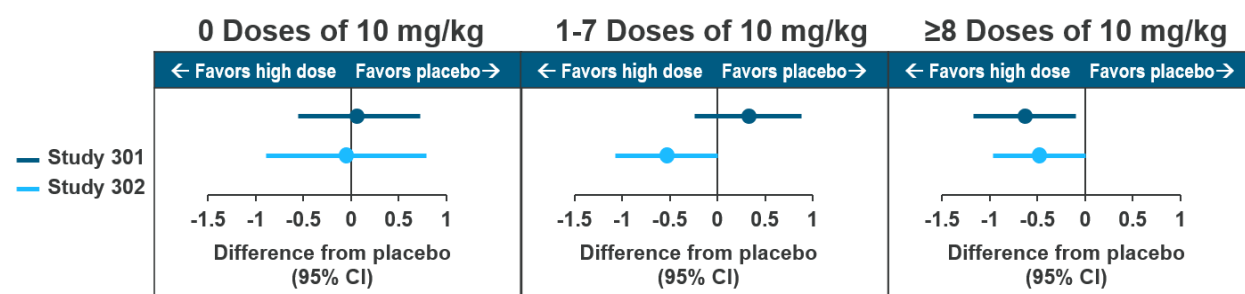
On amyloid PET SUVR, the results for this intermediate exposure subset did not differ markedly between studies. Therefore, and as also shown in the PK/PD model, the dose and biological activity of aducanumab was similar in the two studies at this intermediate exposure to 10 mg/kg doses. What is discrepant in the intermediate dosing subset of the high-dose group is the clinical outcome.

In summary, the discrepancy between results for the high-dose groups of Study 302 and Study 301 was due at least in part to the participants with intermediate dosing, in conjunction with the proportion of participants in the zero and high subsets, favoring a lower overall exposure to 10 mg/kg in Study 301.

However, results for the intermediate exposure to 10 mg/kg dosing subset remain somewhat discrepant between Study 301 and Study 302, even with the rapid progressors removed.

Given that the clinical outcomes are highly consistent in both studies across the low-dose groups, as well as across the high and low exposure subsets of the high-dose groups, and that the biological activity of aducanumab as measured by amyloid PET is similar in the two studies at the intermediate dosing subset, it is reasonable to conclude that the remaining difference in the clinical outcome in the Study 301 intermediate exposure subset may be due to several factors, including greater heterogeneity in dosing patterns ([Figure 18](#)) and measurement noise in the context of low sample sizes and modest measurement sensitivity in an 18-month study [[Evans 2018](#)].

Figure 21: Consistent CDR-SB Outcomes in Participants Receiving Sufficient Doses of 10 mg/kg



Abbreviations: CI = confidence interval.; MMRM = mixed model for repeated measures

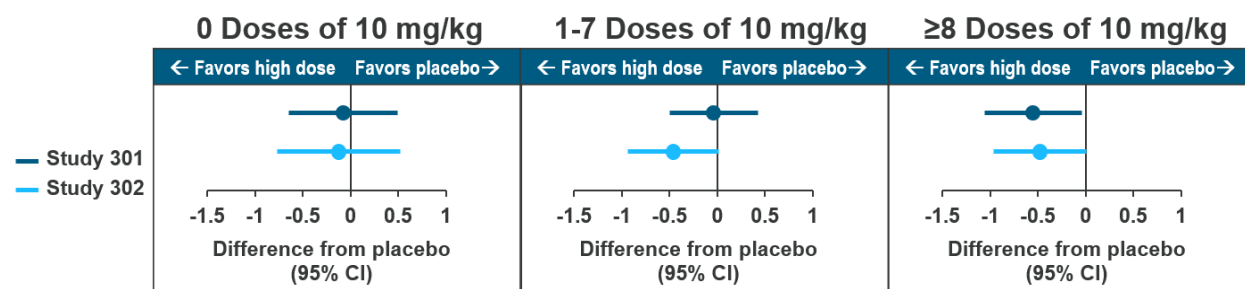
Note 1: The analysis was conducted on the OTC population, i.e., ITT population who had the opportunity to complete the Week 78 visit by March 20, 2019.

Note 2: Placebo participants were selected to match the baseline characteristics of the treated participants in each dosing category based on a propensity score model that includes covariates of laboratory ApoE ε4 status, age, sex, baseline clinical stage, baseline scores of CDR-SB, MMSE, ADAS-Cog13, ADCS-ADL-MCI, years of education, years since first Alzheimer's disease symptom, Alzheimer's disease symptomatic medication use at baseline, US/non-US and enrollment window of every 200 subjects. Placebo and treated subjects matched exactly on laboratory ApoE ε4 status.

Note 3: An MMRM model was used to analyze change from baseline CDR-SB using fixed effects of treatment group, time (categorical), treatment group-by-time interaction, baseline CDR-SB, baseline CDR-SB by time interaction, baseline MMSE, Alzheimer's disease symptomatic medication use at baseline, region, and laboratory ApoE ε4 status.

Data source: ISE/CSR/T-CDR-SS10-PSMB-PC

Figure 22: Consistent CDR-SB Outcomes in Participants Receiving Sufficient Doses of 10 mg/kg – Excluding Rapid Progressors



Abbreviations: CI = confidence interval.; MMRM = mixed model for repeated measures.

Note 1: The analysis was conducted on the OTC population, i.e., ITT population who had the opportunity to complete the Week 78 visit by March 20, 2019.

Note 2: Placebo participants were selected to match the baseline characteristics of the treated participants in each dosing category based on a propensity score model that includes covariates of laboratory ApoE ε4 status, age, sex, baseline clinical stage, baseline scores of CDR-SB, MMSE, ADAS-Cog13, ADCS-ADL-MCI, years of education, years since first Alzheimer's disease symptom, Alzheimer's disease symptomatic medication use at baseline, US/non-US and enrollment window of every 200 subjects. Placebo and treated subjects matched exactly on laboratory ApoE ε4 status.

Note 3: An MMRM model was used to analyze change from baseline CDR-SB using fixed effects of treatment group, time (categorical), treatment group-by-time interaction, baseline CDR-SB, baseline CDR-SB by time interaction, baseline MMSE, Alzheimer's disease symptomatic medication use at baseline, region, and laboratory ApoE ε4 status.

Data source: ISE/CSR/T-CDR-SS10-PSMB-CDR8-PC.

Exposure Response

A summary of exposure-response modeling results for Study 302 is provided in Section 3.4.3. Similar exposure-response modeling was applied to pooled data from Studies 301 and 302, with

data from > 3000 participants and 50,000 PK samples. Using pooled data allowed for objective tests of whether the exposure-response relationship on pharmacodynamic (brain amyloid) and efficacy differed between studies.

This modeling did not find statistically significant differences between Studies 301 and 302 in the exposure-response relationship for effects on brain amyloid or most of the clinical outcomes (ADAS-Cog13 and ADCS-ADL-MCI, and SUVR-CDR-SB response models). Results of exposure-response analyses relating Cavg to CDR-SB, when including all participants in the pooled 301 and 302 studies, indicated that the drug effect in Study 301 is not statistically significant. However, when considering only the population of “typical progressors” (which represent 86% of the Study 301 and 302 population), the drug effect for Study 301 is statistically significant for the exposure-response analyses relating Cavg to CDR-SB. Given the substantial heterogeneity in the CDR-SB outcomes, the exposure-CDR-SB model is less precise than the exposure-amyloid PET SUVR model.

The lack of a between-study difference in the exposure-response relationship for amyloid removal suggests that the divergent clinical results for high dose were not due to the underlying actions of aducanumab.

The exposure-response relationships on the clinical outcomes support that, with greater exposure, greater clinical effects are seen in both Study 301 and Study 302. These results support the observed data where participants in Study 301 with high and consistent exposure to aducanumab had outcomes similar to those in Study 302 ([Figure 21](#) and [Figure 22](#)).

The FDA's Position:

FDA agrees with the importance of dose as an area of investigation. The lack of adequate dosing has been cited as a contributing factor to the failure of previous clinical trials of amyloid-targeting therapies and protocols of notable clinical trials have been amended in recent years to significantly increase dose levels. The importance of dose was also directly established in the aducanumab program by Study 103 which demonstrated a dose-dependent reduction in brain amyloid and reduction of decline on clinical outcome measures. Through the implementation of protocol amendments, exposure to aducanumab increased over the course of Studies 301 and 302. To a lesser extent aducanumab exposure also differed between the studies. It is worth noting that the Division asked the Applicant about the role of dosing in preliminary comments to the June 14, 2019, Type C Meeting, before any investigation of the results of Studies 301 and 302 began.

A guiding principle of the hypothesis was that if aducanumab is effective and the effect is dose-related as in Study 302, it follows that patients in Study 301 with adequate and consistent dosing should also demonstrate an effect on clinical endpoints. An absence of an effect in this subgroup of patients in Study 301 would diminish the persuasiveness of Study 302. Although it is impossible to fully account for all factors that may contribute to findings in subgroups formed by post-randomization factors (i.e., dosing), a variety of approaches, each with strengths and limitations, appears to show that consistent exposure to high doses of aducanumab does lead to similar treatment effects in the two studies.

Evaluation of dosing is complicated by many factors, including ARIA, ApoE ϵ 4 carrier status, and the observation that the placebo response worsened later in the studies when more patients also happened to have the opportunity to receive the 10 mg/kg dose. The analysis used in Approach 1 preserved randomization and thus placebo subgroups are matched to the corresponding aducanumab treatment group and to some degree accounts for potential changes over time. The analysis used in Approach 2 considered region (US vs. non-US), ApoE ϵ 4 carrier status, and enrollment window as well as other factors, in an attempt to account for relevant baseline characteristics. Outcomes in non-US countries, however, were heterogeneous and it may not be appropriate to pool them into a single non-US category. On the other hand, the number of patients enrolled in each country was insufficient to ensure adequate matching on this variable. Before embarking on Approach 2, it was recognized that matching may not account for all possible confounders.

Although the precise relationship between dosing and treatment effect is unknown, the difference in various measures of aducanumab exposure between the studies is modest and dosing alone does not explain the negative finding for the high dose in Study 301. Also, there remains a subset of high dose assigned patients in Study 301 who received intermediate exposure to 10 mg/kg yet failed to show a treatment effect of similar character to high dose assigned patients who received intermediate exposure to 10 mg/kg in Study 302 or even subjects who received the low dose in either study. FDA also acknowledges the inherent variability in clinical measures and challenges measuring clinical decline in this patient population. For these reasons, these analyses do not provide independent evidence of the effectiveness of aducanumab, but rather contribute to the overall understanding of Study 301.

3.6.4. Conclusions

Study 301 and Study 302 have partially discordant results, specifically, the clinical and biomarker results in the high-dose groups. Based on the post hoc analyses presented in this section, together with the results reported in previous sections for each of the studies (Sections 3.4 and 3.5), Biogen and the FDA have jointly concluded that Study 301 does not represent evidence that aducanumab is ineffective. Indeed, using multiple methods, participants in Study 301 with sufficient exposure to 10 mg/kg have clinical outcomes similar to those in Study 302. Furthermore, it was concluded that the differences between studies in clinical and biomarker outcomes in the high-dose group are sufficiently well understood in Study 301 so as to allow independent consideration of the persuasiveness of Study 302.

What is consistent between Studies 301 and 302 is:

- Consistent exposure to 10 mg/kg in the high-dose groups of both studies results in similar treatment effects.
- The low-dose groups in both studies had similar treatment effects across clinical outcomes.
- The low-dose groups in both studies had similar treatment effects on proximal biomarkers and downstream biomarkers of disease.
- The clinical and biomarker outcomes are supported by exposure-response models of the pooled Study 301 and Study 302 data, in which the pharmacological properties of aducanumab on brain amyloid and clinical outcomes did not differ between studies. These outcomes indicate that the discordant results were not due inherently to aducanumab.
- Demographic and disease characteristics did not differ between studies and were not attributable to the differences between Studies 301 and 302.
- The frequency, severity, and management of ARIA did not differ between the studies. ARIA occurs early most frequently during the titration phase and is dose dependent, although ARIA occurs at all dose levels. No systematic bias due to potential functional unblinding due to ARIA could be detected and did not differ between studies.

What differentially impacts the Study 301 high-dose results and contributes to the divergence in outcomes between the 2 studies is:

- In the Study 301 high-dose group, clinical outcomes were substantially impacted by an imbalance in a very small number of rapidly progressing participants.
- In the Study 301 high-dose group, clinical and biomarker outcomes were impacted by lower exposures to the target dose of 10 mg/kg. Fewer participants in Study 301 had high exposure to 10 mg/kg and more participants had no exposure to 10 mg/kg than in Study 302.
- In the setting of lower exposure and a greater number of rapid progressors in the Study 301 high-dose group than all other groups, the inherent variability in the

clinical outcome measures and potentially other factors may have additionally contributed to the discordant results between studies in the high-dose group.

It is reasonable to conclude that the partially discrepant results between Studies 301 and 302 are sufficiently well understood to allow for independent consideration of the persuasiveness of Study 302.

The work was presented at a Type C meeting with the FDA (October 2019), where it was concluded by the FDA, “The analyses conducted since the June 14, 2019, Type C meeting, have established not only that the results of Studies 301 and 302 are interpretable, but on face, suggest an understanding of the discordant results of Studies 301 and 302 sufficient to allow for independent consideration of whether Study 302 might provide evidence adequate to establish the effectiveness of aducanumab for the treatment of Alzheimer's disease.”

Study 302 is further supported by the prior study, Study 103, which is discussed in Section 3.7.

The FDA's Position:

FDA agrees that the understanding of the discordant results in the high dose arms of Studies 301 and 302, while not complete, is sufficient to allow for independent consideration of Study 302.

3.7. Summary of Study 103

Results for Study 103 are summarized below. The focus is on results from the aducanumab 10 mg/kg group, since the Month 12 efficacy results in the 10 mg/kg group in Study 103 are the relevant findings for comparison with the Month 18 results in the high dose group in Study 302. To contextualize these results, we note the following aspects of the studies:

- Study 103 had a 12-month placebo-controlled period, while Study 302 had an 18-month placebo-controlled period.
- In Study 302, the high-dose regimen consisted of 14 doses of 10 mg/kg over 1 year, following the 6-month titration period.
- In Study 103, the 10 mg/kg dose regimen consisted of 14 doses of 10 mg/kg over the 12-month placebo-controlled period.

3.7.1. Disposition, Demographics, and Baseline Characteristics

The Applicant's Position

A total of 196 participants were randomized and dosed in the study. Of these, 154 (79%) completed the study and 42 (21%) withdrew from the study. In the 10 mg/kg group, 11 (34%) withdrew from study.

Treatment groups were generally balanced with respect to demographics (Table 14). Baseline disease characteristics were generally balanced across groups and were consistent with enrollment criteria (Table 14).

Table 14: Demographics and Baseline Characteristics—Placebo-Controlled Period—Study 103

Characteristic	Placebo (n = 48)	1 mg/kg (n = 31)	3 mg/kg (n = 32)	6 mg/kg (n = 30)	10 mg/kg (n = 32)	Titration (n = 23)	Total (n = 196)
Age in years, mean ± SD	73.3±6.82	72.6±7.75	70.5±8.17	73.3±9.27	73.7±8.33	73.1±7.83	72.8±7.93
Sex, Female n (%)	28 (58)	13 (42)	17 (53)	15 (50)	15 (47)	10 (43)	98 (50)
Race							
Asian n (%)	0 (0)	0 (0)	0 (0)	0 (0)	1 (3)	0 (0)	1 (<1)
White n (%)	48 (100)	31 (100)	31 (97)	28 (93)	30 (94)	23 (100)	191 (97)
Education years, mean ± SD	15.5±2.98	15.5±3.15	15.5±2.42	16.1±2.76	15.2±2.35	14.5±3.33	15.4±2.84
AD medications used, n (%)	32 (67)	21 (68)	28 (88)	20 (67)	17 (53)	12 (52)	130 (66)
ApoE ε4, n (%)							
Carriers	34 (71)	19 (61)	21 (66)	21 (70)	20 (63)	23 (100)	138 (70)
Noncarriers	14 (29)	12 (39)	11 (34)	9 (30)	12 (38)	0 (0)	58 (30)
Clinical stage, n (%)							
Prodromal AD	22 (46)	10 (32)	14 (44)	12 (40)	13 (41)	13 (57)	84(43)
Mild AD	26 (54)	21 (68)	18 (56)	18 (60)	19 (59)	10 (43)	112 (57)
PET SUVR, mean composite ± SD	1.435±0.1719	1.441±0.1461	1.464±0.1494	1.429±0.1989	1.441±0.1917	1.325±0.1796	1.430±0.1751

Data source: 221AD103/CSR/T-BL-DEMOG-PC-ARM19; 221AD103/CSR/T-BL-ALZHER-PC-ARM19

The FDA's Position:

The FDA has verified the Applicant's presentation of disposition, demographics, and baseline disease characteristics for Study 103. The disease characteristics represent a population overlapping with Studies 301 and 302 but with more patients later in the disease continuum. Due to the relatively smaller sample sizes in each arm of Study 103, there are small imbalances in some of the baseline demographic characteristics. For example, the percentage of female subjects in the 10 mg/kg arm (47%) is lower than in the pooled placebo arm (58%). Also, there are more ApoE ϵ 4 carriers in the pooled placebo arm (71%) than the 10 mg/kg arm (63%). This imbalance may be due in part to the pooling of placebo subjects across arms including Arm 9 which only enrolled ApoE ϵ 4 carriers. Subjects in the pooled placebo arms also had lower mean baseline CDR-SB (2.69) than subjects in the 10 mg/kg arm (3.14). Modeling analyses indicated that baseline demographic and disease characteristics explain a relatively small proportion of the overall variability in progression of CDR-SB. Therefore, minor baseline imbalances in Study 103 are unlikely to fundamentally alter the interpretation of the results. As noted previously, the separating of the randomization by cohorts means that any comparison with pooled placebo or between doses is not supported by randomization.

The FDA agrees the 10 mg/kg fixed-dose in Study 103 is the relevant dose to compare to the high dose in Study 302 for the reasons listed by the Applicant.

3.7.2. Efficacy

The Applicant's Position

Aducanumab resulted in a dose-dependent reduction in clinical decline, as measured by the CDR-SB and MMSE, in comparison with placebo at Month 12.

- Based on the pre-specified analysis method (ANCOVA), efficacy was seen in the 10 mg/kg fixed-dose group as compared with the placebo group on both the CDR-SB ($p = 0.0246$) and the MMSE ($p = 0.0430$) at Month 12 ([Table 15](#)).
- Consistent with the pre-specified analysis, sensitivity analyses using an MMRM model showed treatment differences favoring the aducanumab 10 mg/kg fixed-dose group vs placebo on both the CDR-SB (-1.08 [-57%], $p=0.0464$) and MMSE (1.9 [-76%], $p=0.0356$).

These results were achieved even though the study was not powered to detect changes in clinical measures of efficacy.

The dose-dependent effect of aducanumab on clinical outcomes as compared with placebo at Month 12 (as assessed by CDR-SB and MMSE scores) was statistically significant ($p = 0.0264$ and 0.0387 , respectively), based on a linear contrast test in ANCOVA based on data from the 1, 3, 6, and 10 mg/kg fixed-dose groups.

Additional analyses of CDR-SB data were performed to address a limitation inherent in the cohort design of this proof-of-concept study, namely, the pooling of the placebo group from each cohort due to the 3:1 randomization (Section 3.1.2, [Figure 3](#)). Comparison of the aducanumab 10 mg/kg fixed-dose group with the concurrent placebo group using ANCOVA showed a

treatment difference of -1.29 (p=0.0756); the corresponding MMRM result was -1.12 (p=0.1254). Supplementary post hoc analyses (1) excluding data after intercurrent events (change in AD medication or early discontinuation from treatment) and (2) excluding data from the titration placebo arm also yielded directionally consistent results.

The clinical efficacy results in Study 103 were supported by a statistically significant, dose-dependent reduction in brain amyloid plaque, as measured by PET, in comparison with placebo at Month 12.

- Maximal reduction in brain amyloid plaque was observed in the 10 mg/kg fixed-dose group as compared with placebo (p < 0.0001) [Figure 23].
- The magnitude of the reduction in brain amyloid plaques in the 10 mg/kg group as compared with placebo in Study 103 (-0.277) was similar to that observed in the aducanumab high-dose group in Study 302 (-0.278).

Table 15: CDR-SB and MMSE Change from baseline at 12 Months—Placebo-Controlled Period--Study 103

Difference vs PBO ^a (%) p-value						
	PBO decline (N=46)	Fixed dose Groups				Titration ^b (N=22)
		1 mg/kg (N=30)	3 mg/kg (N=32)	6 mg/kg (N=29)	10 mg/kg (N=30)	
CDR-SB	n = 39 1.89	n = 23 -0.20 (-11%) 0.7249	n = 27 -0.56 (-30%) 0.2995	n = 26 -0.80 (-42%) 0.1398	n = 23 -1.26 (-67%) 0.0246	n = 21 -1.19 (-63%) 0.0432
MMSE	n = 40 -2.5	n = 25 0.2 (-8%) 0.7932	n = 26 1.7 (-68%) 0.0700	n = 26 0.5 (-20%) 0.6133	n = 25 1.9 (-76%) 0.0430	n = 21 1.5 (-60%) 0.1496

Abbreviations: PBO = placebo.

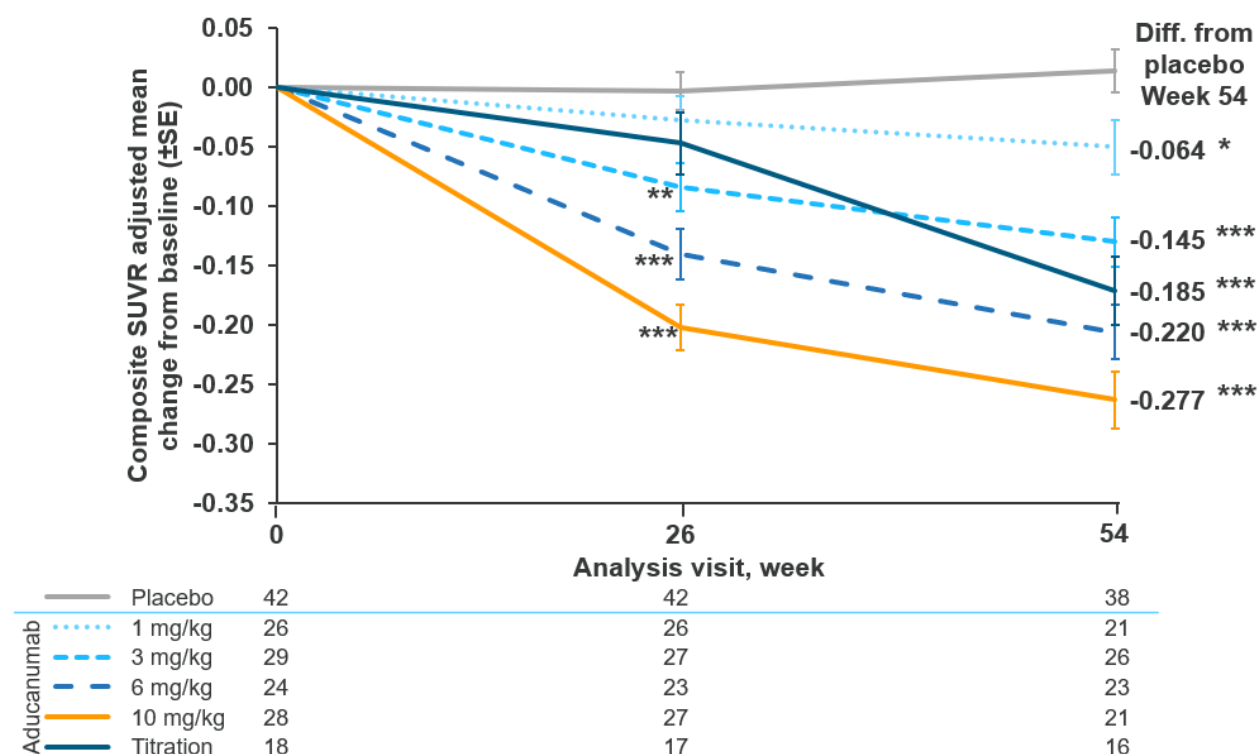
^a Difference vs placebo at Week 54 for CDR-SB and at Week 52 for MMSE.

^b Titration regimen consisted of 2 doses of 1 mg/kg, then 4 doses of 3 mg/kg, 5 doses of 6 mg/kg, and 3 doses of 10 mg/kg. N: number of participants in the efficacy analysis population, defined as all participants who were randomized, received at least one dose of study treatment, and have both baseline and at least one postbaseline questionnaire assessment; n: numbers of participants in the efficacy analysis population with endpoint assessment at Week 54/52.

Note: Change from baseline in CDR-SB and MMSE scores during the placebo-controlled period was analyzed using an ANCOVA model, the primary analysis specified in Study 103 Statistical Analysis Plan, with change from baseline as dependent variable, and treatment group, baseline score, and laboratory ApoE status as independent variables.

Data sources: 221AD103/CSR/T-CDR-CHG-ANCOV-PC-ARM19; 221AD103/CSR/T-MMSE-CHG-ANCOV-PC-ARM19

Figure 23: Brain amyloid PET Change from Baseline at 6 and 12 Months—Placebo-Controlled Period—Study 103



Abbreviations: Diff = difference; SE = standard error.

* p < 0.05, ** p < 0.01, *** p < 0.001 (nominal).

Note: Results were based on an ANCOVA model, the primary analysis specified in Study 103 Statistical Analysis Plan, with change from baseline as dependent variable, and treatment group, baseline value, and laboratory ApoE ε4 status as independent variables.

Data sources: 221AD103/CSR/T-CDR-CHG-ANCOV-PC-ARM19

3.7.3. Conclusion

In Study 103, the clinical and pharmacodynamic effects in aducanumab indicated a dose-dependent effect, with the greatest statistically significant amyloid plaque reduction and reduction in clinical decline observed for aducanumab 10 mg/kg given over 12 months. Sensitivity analyses using MMRM support the primary analysis of the clinical endpoints. Other post hoc analyses using concurrent placebo or excluding data after intercurrent events showed directionally consistent results.

The CDR-SB, MMSE, and brain amyloid plaque results in Study 103 are consistent with the corresponding CDR-SB, MMSE, and brain amyloid plaque results from Study 302.

Therefore, based on the above, Study 103 provides an additional contribution to the substantial evidence of the effectiveness of aducanumab.

The FDA's Position:

The results presented by the Applicant are an accurate representation of the prespecified analyses for Study 103. An MMRM model (the same primary analysis method that was used in Studies 301 and 302), which more appropriately handles dropouts, with the same covariates as the ANCOVA model specified in the SAP was used as a supplementary analysis and provided similar estimates of the treatment effect for the 10 mg/kg treatment arm for both CDR-SB and MMSE.

Of note, the titration arm of Study 103 has a nominal p-value of 0.0432 using ANCOVA. However, the MMRM analysis of the titration group is not significant ($p=0.2044$) likely due to the inclusion of a titration group dropout with a 6.5 point worsening at Week 26 which is excluded from the Applicant's by-Visit ANCOVA analysis of Week 54.

The titration arm in Study 103 is not the relevant dose to compare to the high dose arm in Study 302 because the titration period was much longer in Study 103 (44 weeks) compared to Study 302 (24 weeks). Subjects in the titration arm in Study 103 received 3 doses of 10 mg/kg compared to 14 doses of 10 mg/kg in the high dose arm in Study 302. In fact, the reduction of brain β -amyloid for the titration arm in Study 103 at Week 54 was more similar to the lower dose in Study 302.

Because Study 103 was designed as a safety and tolerability study, the following must be considered when interpreting the clinical efficacy results:

- Although there was a statistical analysis plan for analysis of clinical endpoints, there was no control for multiplicity.
- The placebo group is pooled across the different arms enrolled in the study and is therefore not entirely contemporaneous with the 10 mg/kg treatment arm and the analysis does not reflect the randomization.
- The statistical significance of the 10 mg/kg treatment arm demonstrated with both ANCOVA and MMRM analyses was not robust to the exclusion of post-baseline starting of AD medications or exclusion of the titration placebo arm.

Despite the smaller sample size, the 10 mg/kg dose arm was able to achieve statistical significance according to the pre-specified analysis plan. Also, the dose-response relationships for $A\beta$ reduction provide support for the positive finding in the 10 mg/kg treatment arm and are consistent with the dose-response relationship observed for CDR-SB and MMSE.

While the FDA agrees the 10 mg/kg fixed-dose in Study 103 is the relevant dose to compare to the high dose in Study 302, there are notable differences in the time course of $A\beta$ reduction in the two treatment arms which should be considered. Specifically, the fixed-dose 10 mg/kg treatment reduces β -amyloid plaque levels earlier than the titration regimen used in Studies 301 and 302 (as seen in [Figure 23](#) vs. [Figure 6](#)). The precise relationship between reduction in β -amyloid plaque levels and cognitive decline is unknown but may depend on the time to achieve a certain level of $A\beta$ reduction.

3.8. Efficacy Summary and Conclusions

The Applicant's Position

Study 302 is a positive study, with statistically significant differences on the prespecified primary, secondary, and tertiary clinical efficacy outcomes that are internally consistent, robust to departures from assumptions, and provides the primary contribution to the substantial evidence of the effectiveness of aducanumab.

- Study 302 results showed statistically significant differences from placebo in the high-dose group on the primary endpoint, all three secondary endpoints, and the tertiary efficacy endpoint. These 5 endpoints cover a broad range of symptoms, with little overlap between them, and therefore each endpoint provides important and independent evidence of the broad impact of aducanumab on the range of symptoms of Alzheimer's disease (Table 3, Table 7, Table 8).
- Internal consistency of the efficacy results was demonstrated in that 79 of 80 subgroup comparisons across 5 clinical endpoints and all 13 subgroup comparisons for amyloid PET showed the treatment effect favored aducanumab high dose over placebo (Figure 9, Figure 10).
- Statistical significance of the high-dose group in Study 302 was robust to departures from assumptions regarding missing data and non-normality (Table 5, Table 6).
- The clinical effects of aducanumab are supported by substantial treatment effects on multiple objective measures of underlying Alzheimer's disease pathophysiology, including brain amyloid levels as well as biomarkers of downstream tau pathology and neurodegeneration (Figure 6, Figure 7, Figure 8).

Study 103 also contributes to the substantial evidence for the effectiveness of aducanumab. Study 103 results for the 10 mg/kg dose group showed a statistically significant clinical benefit, as measured on the CDR-SB and MMSE, as well as dose- and time- dependent effects on brain amyloid plaque levels (Table 15, Figure 23). Therefore, results of the 10 mg/kg group in Study 103 were consistent with and support the high-dose results from Study 302.

Study 301 failed to meet its primary and secondary objectives. Post hoc analyses concluded that Study 301 is sufficiently well understood so as not to detract from the persuasiveness of Study 302 and not to preclude the independent consideration of Study 302. Study 301 does not represent evidence that aducanumab is ineffective, and indeed, there is evidence of effectiveness in participants with consistent exposure to 10 mg/kg in Study 301, where effects on clinical endpoints are consistent with those observed in Study 302.

Presented below is an integrated summary of efficacy results from Studies 302 and 103, noting that CDR-SB, MMSE, and brain amyloid plaque PET levels were common to the 2 studies. The consistency of the results for these 3 endpoints in the high-dose group of Study 302 and the 10 mg/kg group of Study 103 is evident (Table 16, Table 17).

In summary, the totality of the data demonstrate that aducanumab is effective in removing a core pathology of Alzheimer's disease and thereby slows clinical decline, including decline in memory, language, executive and visuospatial function, personality, and behavior, all of which cause loss of abilities to perform instrumental and/or basic activities of daily living.

Table 16: Study 302 (18 Months) and Study 103 (12 Months): Change from baseline on the CDR-SB, MMSE, ADAS-Cog13, ADCS-ADL-MCI, and NPI-10

	Study 302			Study 103	
	Diff vs PBO (%) p-value			Diff vs PBO (%) p-value	
	PBO decline N=548	LOW N=543	HIGH N=547	PBO decline N=48	10 mg/kg N=32
CDR-SB	n=288 1.74	n=290 -0.26 (-15%) 0.0901	n=299 -0.39 (-22%) 0.0120	n=39 1.89	n=23 -1.26 (-67%) 0.0246
MMSE	n=288 -3.3	n=293 -0.1 (3%) 0.7578	n=299 0.6 (-18%) 0.0493	n=40 -2.5	n=25 1.9 (-76%) 0.0430
ADAS-Cog13	n=287 5.162	n=289 -0.701 (-14%) 0.1962	n=293 -1.400 (-27%) 0.0097		
ADCS-ADL-MCI	n=283 -4.3	n=286 0.7 (-16%) 0.1515	n=295 1.7 (-40%) 0.0006		
NPI-10	n=282 1.5	n=283 -0.5 (-33%) 0.3921	n=291 -1.3 (-87%) 0.0215		

Abbreviations: Diff = difference; PBO = placebo.

N: number of subjects in ITT population; n: number of ITT subjects with endpoint assessment at 18 months (Study 302) or 12 months (Study 103)

Note: Study 103 results for placebo and the 10 mg/kg fixed dose arm are shown. Refer to Section 3.7 for results for all groups.

Note: Study 103 results were based on an ANCOVA model at 12 months, which is the primary model specified in Study 103 Statistical Analysis Plan. Study 302 results were based on an MMRM model up to 18 months, which is the primary model specified in Study 302 Statistical Analysis Plan.

Note: Study 301/302 efficacy data after March 20, 2019 were censored.

Data source: Table 8 and Table 15.

Table 17: Studies 301 and 302 (18 Months) and Study 103 (12 Months): Change from Baseline in Biomarkers

	Study 302		Study 103	
		Diff vs PBO p-value		Diff vs PBO p-value
Amyloid PET	PBO change from baseline N=159	LOW N=159	HIGH N=170	PBO decline N=42 10 mg/kg N=28
	n=93 0.014	n=100 -0.179 <0.0001	n=109 -0.278 <0.0001	n=21 -0.277 <0.0001
CSF analyte	PBO change from baseline	Diff vs PBO p-value		
		LOW	HIGH	
β-amyloid ₁₋₄₂ CSF	n=28 -30.69	n=33 179.57 < 0.0001	n=17 318.88 < 0.0001	
p-Tau CSF	n=28 -0.49	n=33 -15.64 0.0035	n=17 -22.44 0.0005	
t-Tau CSF	n=28 -0.39	n=33 -86.74 0.0148	n=17 -112.05 0.0008	
Tau PET composite region *	PBO change from baseline	Diff vs PBO p-value		
		LOW	HIGH	
Frontal	n=12 0.090	n=14 -0.049 0.0876	n=11 -0.073 0.0212	
Medial Temporal	n=12 0.082	n=14 -0.115 0.0012	n=11 -0.132 0.0005	
Temporal	n=12 0.082	n=14 -0.065 0.1174	n=11 -0.096 0.0304	

Abbreviations: Diff = difference; PBO = placebo.

N: number of subjects in the [¹⁸F]-florbetapir amyloid PET analysis population; n: number of subjects in the analysis population with an assessment at 18 months (Study 302) or 12 months (Study 103)

Note: Study 103 results for placebo and the 10 mg/kg fixed dose arm are shown. Refer to Section 3.7 for results for all groups.

Note: Study 103 results were based on an ANCOVA model at 12 months, which is the primary model specified in Study103 Statistical Analysis Plan. Study 302 results were based on an MMRM model up to 18 months, which is the primary model specified in Study 302 Statistical Analysis Plan.

*Due to the early termination of the studies, all Tau PET assessments performed in the placebo-controlled period were pooled from Study 301 and 302 and used as one postbaseline timepoint.

Data source: 221AD302/CSR/T-PET-MMRM-COMP-CERE-PC; 221AD103/CSR/T-PETCOMP-CHG-ANCOV-PC-ARM19; ISE/CSR/T-CSF-ANCOVA-PC; ISE/CSR/T-TAU-ANCOVA-PC-POOL

The FDA's Position:

Study 302 provides the primary evidence of effectiveness of aducanumab. The effect of aducanumab in Study 302 is robust and exceptionally persuasive on several of the instruments used to evaluate efficacy. In fact, the effects observed for the primary and secondary endpoints encompass two acceptable approaches to establish effectiveness: (1) primary endpoint of CDR-SB and (2) co-primary endpoints of ADAS-Cog 13 and ADCS-ADL-MCI. The estimate of the treatment effect in the low dose arm was numerically favorable and was consistent with a dose-response relationship. The treatment effect for the high dose arm is supported by consistently favorable results across subgroups of interest. Biomarker results demonstrate target engagement and treatment effects on markers of Alzheimer's disease pathophysiology.

The results of Study 103 are appropriately viewed as supportive evidence of the effectiveness of aducanumab. Despite the limitations of a trial designed to assess safety and tolerability rather than effectiveness, the 10 mg/kg dose arm was able to achieve statistical significance according to the prespecified analysis plan. Also, the dose-response relationship for A β reduction provides support for the positive finding in the 10 mg/kg treatment arm and is consistent with the dose-response relationship observed for CDR-SB and MMSE.

Study 301 is a negative study and does not contribute to the evidence of effectiveness of aducanumab. The results presented in Section 3.6 should not be interpreted as "explaining why Study 301 was negative." This was not the intention of this work and such a complete explanation is not necessary to establish the effectiveness of aducanumab. At the June 14, 2019, Type C Meeting, the Division clearly stated that, "available data do not suggest the future use of Study 301 as an efficacy study providing independent evidence of effectiveness supporting the approval of aducanumab." Rather, the Division noted the possibility that analyses "may be understood well enough... to not represent evidence that the drug is ineffective." The analyses presented in Section 3.6 are exploratory by design but limited in scope and focused on pre-defined areas of interest. The rapid progressor analysis indicated that a small imbalance in the number of rapid progressing patients in the high-dose arm in Study 301 had a disproportionate impact on the estimate of the treatment effect using the primary analysis method. An examination of dosing in Study 301 indicates that patients with higher exposure to the 10 mg/kg dose in Study 301 had similar responses to patients in Study 302. These two factors contribute to the overall understanding of Study 301 and together do not meaningfully detract from the persuasiveness of Study 302. There were no findings from the exploration that represented evidence that aducanumab is not effective.

4. SAFETY

The Applicant's Position

The safety profile of aducanumab is well characterized in patients with Alzheimer's disease based on the clinical development program, which includes 2 large, identically designed, global Phase 3 studies, Study 301 and Study 302. Collectively the 2 studies provide an adequate safety database that exceeds the criteria in published guidance regarding the extent of exposure for adequate safety evaluation of a therapy intended for long-term treatment of a non-life-threatening condition [ICH 1994]. In the Phase 3 studies, 2757 participants with Alzheimer's disease were dosed with aducanumab (3, 6, or 10 mg/kg), totaling 4736 person-years of follow-up. Of these, 1347 participants received 10 mg/kg aducanumab, accounting for 2300 person-years of follow-up.

The focus of the safety evaluation of aducanumab presented in this briefing book is on the analyses of pooled data from the Phase 3 studies, including the placebo-controlled experience (Pool A1) and the aducanumab-treated experience (Pool A2, which includes placebo-controlled and long-term extension data). Pool A1 allows for a direct comparison between aducanumab and placebo. This pool enables the assessment of potential adverse drug reactions by indicating which AEs occur more frequently with aducanumab than with placebo. Pool A2 allows for an assessment of the overall safety profile, including longer-term safety in the Phase 3 studies regardless of study period.

Other studies of aducanumab in Alzheimer's disease (Studies 101, 103, 104, 205) had comparatively smaller sample sizes, shorter treatment period durations, and/or tested different aducanumab doses and dose regimens as compared with the Phase 3 program. Nonetheless, the safety results of these smaller studies were also considered in assessing the safety profile of aducanumab.

The safety profile of aducanumab was consistent between the two Phase 3 studies when analyzed by randomized treatment group (placebo, low dose, and high dose), including the incidence of serious adverse events (SAEs) and treatment discontinuation due to an AE, as well as the incidence of ARIA-E which is the most common AE identified among aducanumab-treated participants.

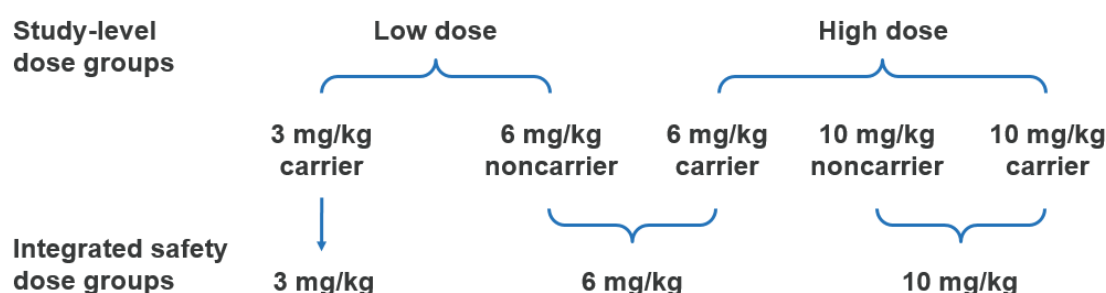
In the Phase 3 studies, participants were randomized to receive low or high doses of aducanumab or placebo and randomization was stratified by ApoE ϵ 4 status (carrier or noncarrier) (Figure 24). The low dose was either 3 mg/kg (titrated from 1 to 3 mg/kg over 8 weeks) in ApoE ϵ 4 carriers, or 6 mg/kg (titrated from 1 to 3 to 6 mg/kg over 24 weeks) in noncarriers. Initially, the target dose in the high-dose groups also differed based on the participant's ApoE ϵ 4 status (6 mg/kg for carriers and 10 mg/kg for noncarriers); starting with Protocol Version 4, the high-dose target was 10 mg/kg (titrated from 1 to 3 to 6 to 10 mg/kg over 24 weeks) regardless of ApoE ϵ 4 status.

To assess the impact of a specific aducanumab dose in the pooled analyses, safety data for Pools A1 and A2 have been analyzed by treatment group based on the target dose in the titration regimen (rather than by study-level dose group) using the following categories: placebo (Pool A1 only), aducanumab 3 mg/kg, aducanumab 6 mg/kg, aducanumab 10 mg/kg, and total aducanumab (total aducanumab combines all the different doses, 3 mg/kg, 6 mg/kg, and 10 mg/kg). Participants randomized to the low-dose carrier, low-dose noncarrier, or high-dose

noncarrier dose groups in the Phase 3 studies are assigned to the aducanumab 3 mg/kg, 6 mg/kg, and 10 mg/kg categories, respectively (Figure 24). Participants randomized to the high-dose carrier group were assigned to either the 6 mg/kg or 10 mg/kg category, based on date of consent to Protocol Version 4 or higher and actual dose received. Over 90% of participants randomized to the high-dose carrier group are analyzed in the 10 mg/kg pooled treatment group. It should be noted that sensitivity analyses were also performed by study-level treatment group and the results were similar to those observed by target dose group.

The integrated safety analyses from the Phase 3 studies presented in this briefing book provide robust characterization of the safety of aducanumab in participants with Alzheimer's disease and are consistent with the safety profile of Study 302. The information presented in the following sections will focus on a description of the safety profile of the 10 mg/kg dose.

Figure 24: Pooling Strategy for the Integrated Safety Analysis



Carrier and noncarrier status refer to the ApoE ε4 allele.

The FDA's Position:

FDA acknowledges the Applicant's pooling strategy. The FDA primary safety review focuses on Study 302, which will be considered the pivotal efficacy study. Safety findings from Pool A1, Pool A2, Studies 301 and 103, and other studies in Alzheimer's disease are used as supportive information in the FDA review.

4.1. Overall Extent of Exposure

The Applicant's Position

The extent and duration of treatment with aducanumab in the clinical development program provides a robust safety database for evaluation of aducanumab, which is intended for long-term treatment of Alzheimer's disease.

Assessment of the safety of aducanumab in the 2 pooled analyses presented in this briefing book (Pool A1 and Pool A2) is based on the large number of participants exposed in the Phase 3 Studies 301 and 302 and the duration of exposure. These pooled analyses are supported by selected analyses performed using integrated safety data from the 2 Phase 3 studies and Studies 103 and 104 (Pool B). Because of the large number of participants in the Phase 3 studies as compared with Studies 103 and 104, the aggregate safety data in Pool B are very similar to Pool A2. All results were taken into account when assessing the aducanumab safety profile.

In the placebo-controlled period of the Phase 3 studies (Pool A1), 1087 participants were exposed to placebo and 2198 participants were exposed to any dose of aducanumab, including 1033 participants exposed to 10mg/kg ([Table 18](#)). Total person-years of exposure was 1416.1 and 2751.7, and total person-years of follow-up was 1556.7 and 3136.4, in the placebo and total aducanumab groups, respectively. Total person-years of exposure for 1033 participants in the 10 mg/kg dose group was 1294.5 and total person-years of follow-up was 1478.9.

A total of 2757 participants from the Phase 3 studies were exposed to aducanumab in the aducanumab-treated period (Pool A2) [[Table 19](#)]. Total person-years of exposure was 3983.5 and total person-years of follow-up was 4736.1. Total person-years of exposure for 1347 participants in the 10 mg/kg dose group in Pool A2 was 1933.3 and total person-years of follow-up was 2300.5.

In the 4 multiple-dose studies included in Pool B, 2959 participants were exposed to aducanumab. Participants were followed for a mean duration of 1.8 years, representing 5319.1 person-years of follow-up.

An overall summary of exposure for each pooled analysis by treatment and for total placebo and aducanumab treatment is presented in [Table 20](#).

Table 18: Overall Extent of Exposure – Pool A1

Overall extent of exposure - Pool A1
Page 1 of 5

	Placebo (N=1087)	BIIB037 3 mg/kg (N=760)	BIIB037 6 mg/kg (N=405)	BIIB037 10 mg/kg (N=1033)	BIIB037 total (N=2198)
Number of infusions received (categorical)					
1-5	34 (3.1)	41 (5.4)	48 (11.9)	48 (4.6)	137 (6.2)
6-10	96 (8.8)	84 (11.1)	43 (10.6)	148 (14.3)	275 (12.5)
11-15	207 (19.0)	151 (19.9)	57 (14.1)	211 (20.4)	419 (19.1)
16-20	750 (69.0)	484 (63.7)	257 (63.5)	626 (60.6)	1367 (62.2)
Number of infusions received (continuous)					
n	1087	760	405	1033	2198
Mean	16.7	15.9	15.3	15.4	15.6
SD	4.53	4.92	5.83	4.91	5.10
Median	19.0	18.0	18.0	17.0	18.0
Q1, Q3	14.0, 20.0	13.0, 20.0	11.0, 20.0	12.0, 20.0	12.0, 20.0
Min, Max	1, 20	1, 20	1, 20	1, 20	1, 20

NOTE: Numbers in parentheses are percentages.

- (a) Based on a schedule of 1 dose every 4 weeks, duration of exposure to study treatment in years is defined as (Date of last dose in in placebo-controlled period - Date of first dose of placebo or aducanumab +29)/365.25.
- (b) Duration of follow-up in years is defined as (Date of last day of placebo-controlled period - Date of first dose of placebo or aducanumab +1)/365.25.
- (c) Duration of follow-up after treatment discontinuation in years is defined as (Date of last day of placebo-controlled period - Date of last dose in the placebo-controlled period + 29)/365.25.
- (d) Time on study after last dose in weeks is defined as (Date of last day of the placebo-controlled period - Date of last exposure to study treatment + 1)/7.

SOURCE: BIIB037AD/ISS/CSR/T-EX-A1.SAS

Data as of: 06NOV2019 Run date: 27MAY2020

Overall extent of exposure - Pool A1
Page 2 of 5

	Placebo (N=1087)	BIIB037 3 mg/kg (N=760)	BIIB037 6 mg/kg (N=405)	BIIB037 10 mg/kg (N=1033)	BIIB037 total (N=2198)
Duration of exposure to study treatment (categorical) (a)					
>=12 weeks	1069 (98.3)	754 (99.2)	391 (96.5)	1023 (99.0)	2168 (98.6)
>=24 weeks	1051 (96.7)	719 (94.6)	358 (88.4)	982 (95.1)	2059 (93.7)
>=36 weeks	1027 (94.5)	692 (91.1)	341 (84.2)	935 (90.5)	1968 (89.5)
>=48 weeks	909 (83.6)	619 (81.4)	305 (75.3)	823 (79.7)	1747 (79.5)
>=60 weeks	801 (73.7)	546 (71.8)	279 (68.9)	721 (69.8)	1546 (70.3)
Duration of exposure to study treatment (years) (continuous) (a)					
n	1087	760	405	1033	2198
Mean	1.3	1.3	1.2	1.3	1.3
SD	0.35	0.38	0.45	0.37	0.39
Median	1.5	1.5	1.5	1.5	1.5
Q1, Q3	1.1, 1.5	1.1, 1.5	0.9, 1.5	1.0, 1.5	1.0, 1.5
Min, Max	0, 2	0, 2	0, 2	0, 2	0, 2
Sum	1416.1	967.3	489.9	1294.5	2751.7

NOTE: Numbers in parentheses are percentages.

- (a) Based on a schedule of 1 dose every 4 weeks, duration of exposure to study treatment in years is defined as (Date of last dose in in placebo-controlled period - Date of first dose of placebo or aducanumab +29)/365.25.
- (b) Duration of follow-up in years is defined as (Date of last day of placebo-controlled period - Date of first dose of placebo or aducanumab +1)/365.25.
- (c) Duration of follow-up after treatment discontinuation in years is defined as (Date of last day of placebo-controlled period - Date of last dose in the placebo-controlled period + 29)/365.25.
- (d) Time on study after last dose in weeks is defined as (Date of last day of the placebo-controlled period - Date of last exposure to study treatment + 1)/7.

SOURCE: BIIB037AD/ISS/CSR/T-EX-A1.SAS

Data as of: 06NOV2019 Run date: 27MAY2020

Overall extent of exposure - Pool A1
Page 3 of 5

	Placebo (N=1087)	BIIB037 3 mg/kg (N=760)	BIIB037 6 mg/kg (N=405)	BIIB037 10 mg/kg (N=1033)	BIIB037 total (N=2198)
Duration of follow-up (years) (continuous) (b)					
n	1087	760	405	1033	2198
Mean	1.4	1.4	1.4	1.4	1.4
SD	0.30	0.29	0.40	0.29	0.31
Median	1.5	1.5	1.5	1.5	1.5
Q1, Q3	1.4, 1.5	1.4, 1.5	1.2, 1.5	1.3, 1.5	1.3, 1.5
Min, Max	0, 3	0, 3	0, 3	0, 3	0, 3
Sum	1556.7	1091.3	566.3	1478.9	3136.4
Duration of follow-up (years) after treatment discontinuation (c)					
n	528	372	207	555	1134
Mean	0.3	0.3	0.4	0.3	0.3
SD	0.26	0.30	0.44	0.28	0.32
Median	0.3	0.3	0.3	0.3	0.3
Q1, Q3	0.1, 0.3	0.2, 0.3	0.2, 0.3	0.2, 0.3	0.2, 0.3
Min, Max	0, 3	0, 2	0, 3	0, 3	0, 3
Sum	138.9	120.4	75.4	179.3	375.2

NOTE: Numbers in parentheses are percentages.

- (a) Based on a schedule of 1 dose every 4 weeks, duration of exposure to study treatment in years is defined as (Date of last dose in in placebo-controlled period - Date of first dose of placebo or aducanumab +29)/365.25.
- (b) Duration of follow-up in years is defined as (Date of last day of placebo-controlled period - Date of first dose of placebo or aducanumab +1)/365.25.
- (c) Duration of follow-up after treatment discontinuation in years is defined as (Date of last day of placebo-controlled period - Date of last dose in the placebo-controlled period + 29)/365.25.
- (d) Time on study after last dose in weeks is defined as (Date of last day of the placebo-controlled period - Date of last exposure to study treatment + 1)/7.

SOURCE: BIIB037AD/ISS/CSR/T-EX-A1.SAS

Data as of: 06NOV2019 Run date: 27MAY2020

Overall extent of exposure - Pool A1
Page 4 of 5

	Placebo (N=1087)	BIIB037 3 mg/kg (N=760)	BIIB037 6 mg/kg (N=405)	BIIB037 10 mg/kg (N=1033)	BIIB037 total (N=2198)
Time on study after the last dose (weeks) (d)					
n	528	372	207	555	1134
Mean	17.8	21.0	23.1	20.9	21.3
SD	13.77	15.94	23.02	14.48	16.81
Median	18.1	18.3	18.3	18.3	18.3
Q1, Q3	11.7, 19.1	16.9, 20.1	15.9, 21.1	17.0, 20.7	16.7, 20.6
Min, Max	0, 158	0, 126	1, 147	0, 153	0, 153
Sum	9392.6	7803.1	4772.6	11622.7	24198.4
Maximum dose of BIIB037 received					
0 mg/kg (placebo)	1087 (100)	0	0	0	0
1 mg/kg	0	6 (0.8)	14 (3.5)	9 (0.9)	29 (1.3)
3 mg/kg	0	753 (99.1)	46 (11.4)	45 (4.4)	844 (38.4)
6 mg/kg	0	0	345 (85.2)	58 (5.6)	403 (18.3)
10 mg/kg	0	1 (0.1)	0	921 (89.2)	922 (41.9)

NOTE: Numbers in parentheses are percentages.

- (a) Based on a schedule of 1 dose every 4 weeks, duration of exposure to study treatment in years is defined as (Date of last dose in in placebo-controlled period - Date of first dose of placebo or aducanumab +29)/365.25.
- (b) Duration of follow-up in years is defined as (Date of last day of placebo-controlled period - Date of first dose of placebo or aducanumab +1)/365.25.
- (c) Duration of follow-up after treatment discontinuation in years is defined as (Date of last day of placebo-controlled period - Date of last dose in the placebo-controlled period + 29)/365.25.
- (d) Time on study after last dose in weeks is defined as (Date of last day of the placebo-controlled period - Date of last exposure to study treatment + 1)/7.

SOURCE: BIIB037AD/ISS/CSR/T-EX-A1.SAS

Data as of: 06NOV2019 Run date: 27MAY2020

Overall extent of exposure - Pool A1
Page 5 of 5

	Placebo (N=1087)	BIIB037 3 mg/kg (N=760)	BIIB037 6 mg/kg (N=405)	BIIB037 10 mg/kg (N=1033)	BIIB037 total (N=2198)
Safety treatment group in individual study					
Placebo	1087 (100)	0	0	0	0
BIIB037 low dose	0	760 (100)	333 (82.2)	0	1093 (49.7)
BIIB037 high dose	0	0	72 (17.8)	1033 (100)	1105 (50.3)

NOTE: Numbers in parentheses are percentages.

- (a) Based on a schedule of 1 dose every 4 weeks, duration of exposure to study treatment in years is defined as (Date of last dose in in placebo-controlled period - Date of first dose of placebo or aducanumab +29)/365.25.
- (b) Duration of follow-up in years is defined as (Date of last day of placebo-controlled period - Date of first dose of placebo or aducanumab +1)/365.25.
- (c) Duration of follow-up after treatment discontinuation in years is defined as (Date of last day of placebo-controlled period - Date of last dose in the placebo-controlled period + 29)/365.25.
- (d) Time on study after last dose in weeks is defined as (Date of last day of the placebo-controlled period - Date of last exposure to study treatment + 1)/7.

SOURCE: BIIB037AD/ISS/CSR/T-EX-A1.SAS

Data as of: 06NOV2019 Run date: 27MAY2020

Table 19: Overall Extent of Exposure – Pool A2

Overall extent of exposure - Pool A2
Page 1 of 6

	BIIB037 3 mg/kg (N=953)	BIIB037 6 mg/kg (N=457)	BIIB037 10 mg/kg (N=1347)	BIIB037 total (N=2757)
Number of infusions received (categorical)				
1-5	92 (9.7)	70 (15.3)	157 (11.7)	319 (11.6)
6-10	140 (14.7)	75 (16.4)	228 (16.9)	443 (16.1)
11-15	192 (20.1)	69 (15.1)	248 (18.4)	509 (18.5)
16-20	149 (15.6)	80 (17.5)	231 (17.1)	460 (16.7)
21-26	162 (17.0)	64 (14.0)	198 (14.7)	424 (15.4)
27-39	182 (19.1)	85 (18.6)	238 (17.7)	505 (18.3)
40-52	36 (3.8)	14 (3.1)	47 (3.5)	97 (3.5)
Number of infusions received (continuous)				
n	953	457	1347	2757
Mean	18.6	17.5	17.8	18.0
SD	10.41	10.76	10.46	10.50
Median	17.0	16.0	16.0	17.0
Q1, Q3	11.0, 25.0	9.0, 25.0	10.0, 25.0	10.0, 25.0
Min, Max	1, 46	1, 45	1, 46	1, 46

NOTE: Numbers in parentheses are percentages.

- (a) Based on a schedule of 1 dose every 4 weeks, duration of exposure to study treatment in years is defined as (Date of last dose on study- Date of first dose of aducanumab +29)/365.25.
- (b) Duration of follow-up in years is defined as (Date of last day of placebo-controlled period - Date of first dose of placebo or aducanumab +1)/365.25.
- (c) Duration of follow-up after treatment discontinuation in years is defined as (Date of last day on study - Date of last dose of aducanumab + 29)/365.25.
- (d) Time on study after last dose in weeks is defined as (Date of last day on study - Date of last exposure to study treatment + 1)/7.

SOURCE: BIIB037AD/ISS/CSR/T-EX-A2.SAS

Data as of: 06NOV2019 Run date: 01JUN2020

Overall extent of exposure - Pool A2
Page 2 of 6

	BIIB037 3 mg/kg (N=953)	BIIB037 6 mg/kg (N=457)	BIIB037 10 mg/kg (N=1347)	BIIB037 total (N=2757)
Duration of exposure to study treatment (categorical) (a)				
>=0.5 years	832 (87.3)	374 (81.8)	1151 (85.4)	2357 (85.5)
>=1 years	647 (67.9)	287 (62.8)	872 (64.7)	1806 (65.5)
>=1.5 years	448 (47.0)	192 (42.0)	582 (43.2)	1222 (44.3)
>=2 years	240 (25.2)	107 (23.4)	319 (23.7)	666 (24.2)
>=3 years	47 (4.9)	17 (3.7)	65 (4.8)	129 (4.7)
Duration of exposure to study treatment (years) (continuous) (a)				
n	953	457	1347	2757
Mean	1.49	1.38	1.44	1.44
SD	0.816	0.840	0.820	0.823
Median	1.38	1.31	1.31	1.32
Q1, Q3	0.85, 2.00	0.75, 1.94	0.84, 2.00	0.84, 2.00
Min, Max	0.1, 3.6	0.1, 3.5	0.1, 3.5	0.1, 3.6
Sum	1419.4	630.9	1933.3	3983.5

NOTE: Numbers in parentheses are percentages.

- (a) Based on a schedule of 1 dose every 4 weeks, duration of exposure to study treatment in years is defined as (Date of last dose on study- Date of first dose of aducanumab +29)/365.25.
- (b) Duration of follow-up in years is defined as (Date of last day of placebo-controlled period - Date of first dose of placebo or aducanumab +1)/365.25.
- (c) Duration of follow-up after treatment discontinuation in years is defined as (Date of last day on study - Date of last dose of aducanumab + 29)/365.25.
- (d) Time on study after last dose in weeks is defined as (Date of last day on study - Date of last exposure to study treatment + 1)/7.

SOURCE: BIIB037AD/ISS/CSR/T-EX-A2.SAS

Data as of: 06NOV2019 Run date: 01JUN2020

Overall extent of exposure - Pool A2
Page 3 of 6

	BIIB037 3 mg/kg (N=953)	BIIB037 6 mg/kg (N=457)	BIIB037 10 mg/kg (N=1347)	BIIB037 total (N=2757)
<hr/>				
Duration of follow-up (years) (continuous) (b)				
n	953	457	1347	2757
Mean	1.76	1.66	1.71	1.72
SD	0.795	0.826	0.798	0.802
Median	1.65	1.58	1.59	1.61
Q1, Q3	1.15, 2.29	1.05, 2.22	1.12, 2.26	1.12, 2.26
Min, Max	0.1, 3.9	0.0, 3.8	0.0, 3.8	0.0, 3.9
Sum	1674.7	760.9	2300.5	4736.1
<hr/>				
Duration of follow-up (years) after treatment discontinuation (c)				
n	953	457	1347	2757
Mean	0.27	0.29	0.27	0.27
SD	0.220	0.318	0.216	0.237
Median	0.27	0.27	0.27	0.27
Q1, Q3	0.22, 0.29	0.17, 0.29	0.22, 0.29	0.22, 0.29
Min, Max	0.0, 2.3	0.0, 2.7	0.0, 2.9	0.0, 2.9
Sum	257.1	130.8	369.6	757.5

NOTE: Numbers in parentheses are percentages.

- (a) Based on a schedule of 1 dose every 4 weeks, duration of exposure to study treatment in years is defined as (Date of last dose on study - Date of first dose of aducanumab + 29)/365.25.
- (b) Duration of follow-up in years is defined as (Date of last day of placebo-controlled period - Date of first dose of placebo or aducanumab + 1)/365.25.
- (c) Duration of follow-up after treatment discontinuation in years is defined as (Date of last day on study - Date of last dose of aducanumab + 29)/365.25.
- (d) Time on study after last dose in weeks is defined as (Date of last day on study - Date of last exposure to study treatment + 1)/7.

SOURCE: BIIB037AD/ISS/CSR/T-EX-A2.SAS

Data as of: 06NOV2019 Run date: 01JUN2020

Overall extent of exposure - Pool A2
Page 4 of 6

	BIIB037 3 mg/kg (N=953)	BIIB037 6 mg/kg (N=457)	BIIB037 10 mg/kg (N=1347)	BIIB037 total (N=2757)
Time on study after last dose (weeks) (continuous) (d)				
n	953	457	1347	2757
Mean	18.1	19.0	18.4	18.4
SD	11.63	16.67	11.41	12.50
Median	18.1	18.1	18.1	18.1
Q1, Q3	15.7, 19.3	13.1, 19.4	15.7, 19.4	15.4, 19.4
Min, Max	0, 126	1, 147	0, 153	0, 153
Sum	17272.0	8678.4	24741.3	50691.7
Maximum dose of BIIB037 received				
1 mg/kg	25 (2.6)	22 (4.8)	43 (3.2)	90 (3.3)
3 mg/kg	927 (97.3)	69 (15.1)	100 (7.4)	1096 (39.8)
6 mg/kg	0	366 (80.1)	107 (7.9)	473 (17.2)
10 mg/kg	1 (0.1)	0	1097 (81.4)	1098 (39.8)

NOTE: Numbers in parentheses are percentages.

- (a) Based on a schedule of 1 dose every 4 weeks, duration of exposure to study treatment in years is defined as (Date of last dose on study - Date of first dose of aducanumab +29)/365.25.
- (b) Duration of follow-up in years is defined as (Date of last day of placebo-controlled period - Date of first dose of placebo or aducanumab +1)/365.25.
- (c) Duration of follow-up after treatment discontinuation in years is defined as (Date of last day on study - Date of last dose of aducanumab + 29)/365.25.
- (d) Time on study after last dose in weeks is defined as (Date of last day on study - Date of last exposure to study treatment + 1)/7.

SOURCE: BIIB037AD/ISS/CSR/T-EX-A2.SAS

Data as of: 06NOV2019 Run date: 01JUN2020

Overall extent of exposure - Pool A2
Page 5 of 6

	BIIB037 3 mg/kg (N=953)	BIIB037 6 mg/kg (N=457)	BIIB037 10 mg/kg (N=1347)	BIIB037 total (N=2757)
At least one dose in long-term extension	581 (61.0)	250 (54.7)	792 (58.8)	1623 (58.9)
Study 221AD301	317 (33.3)	132 (28.9)	403 (29.9)	852 (30.9)
Study 221AD302	264 (27.7)	118 (25.8)	389 (28.9)	771 (28.0)
First dose of aducanumab was in long-term extension	194 (20.4)	85 (18.6)	280 (20.8)	559 (20.3)
Study 221AD301	103 (10.8)	46 (10.1)	148 (11.0)	297 (10.8)
Study 221AD302	91 (9.5)	39 (8.5)	132 (9.8)	262 (9.5)

NOTE: Numbers in parentheses are percentages.

- (a) Based on a schedule of 1 dose every 4 weeks, duration of exposure to study treatment in years is defined as (Date of last dose on study - Date of first dose of aducanumab + 29)/365.25.
- (b) Duration of follow-up in years is defined as (Date of last day of placebo-controlled period - Date of first dose of placebo or aducanumab + 1)/365.25.
- (c) Duration of follow-up after treatment discontinuation in years is defined as (Date of last day on study - Date of last dose of aducanumab + 29)/365.25.
- (d) Time on study after last dose in weeks is defined as (Date of last day on study - Date of last exposure to study treatment + 1)/7.

SOURCE: BIIB037AD/ISS/CSR/T-EX-A2.SAS

Data as of: 06NOV2019 Run date: 01JUN2020

Overall extent of exposure - Pool A2
Page 6 of 6

	BIIB037 3 mg/kg (N=953)	BIIB037 6 mg/kg (N=457)	BIIB037 10 mg/kg (N=1347)	BIIB037 total (N=2757)
Safety treatment group in individual study				
BIIB037 late start low dose	195 (20.5)	86 (18.8)	0	281 (10.2)
BIIB037 late start high dose	0	0	284 (21.1)	284 (10.3)
BIIB037 early start low dose	758 (79.5)	332 (72.6)	0	1090 (39.5)
BIIB037 early start high dose	0	39 (8.5)	1063 (78.9)	1102 (40.0)

NOTE: Numbers in parentheses are percentages.

- (a) Based on a schedule of 1 dose every 4 weeks, duration of exposure to study treatment in years is defined as (Date of last dose on study - Date of first dose of aducanumab +29)/365.25.
- (b) Duration of follow-up in years is defined as (Date of last day of placebo-controlled period - Date of first dose of placebo or aducanumab +1)/365.25.
- (c) Duration of follow-up after treatment discontinuation in years is defined as (Date of last day on study - Date of last dose of aducanumab + 29)/365.25.
- (d) Time on study after last dose in weeks is defined as (Date of last day on study - Date of last exposure to study treatment + 1)/7.

SOURCE: BIIB037AD/ISS/CSR/T-EX-A2.SAS

Data as of: 06NOV2019 Run date: 01JUN2020

Table 20: Overall Summary of Exposure – Safety Population – Aducanumab Summary of Clinical Safety

Overall summary of exposure - safety population - aducanumab summary of clinical safety

	Placebo	BIIB037 1-6 mg/kg	BIIB037 10 mg/kg	Total
Number of subjects (total follow up time in pool) (years)				
Pool A1	1087 (1556.7)	1165 (1657.6)	1033 (1478.9)	3285 (4693.1)
221AD301	540 (773.0)	593 (850.7)	514 (734.9)	1647 (2358.7)
221AD302	547 (783.7)	572 (806.9)	519 (744.0)	1638 (2334.5)
Pool A2		1410 (2435.6)	1347 (2300.5)	2757 (4736.1)
221AD301		720 (1292.9)	684 (1179.6)	1404 (2472.6)
221AD302		690 (1142.7)	663 (1120.9)	1353 (2263.5)
Pool B		1545 (2827.0)	1414 (2492.0)	2959 (5319.1)
221AD301		720 (1292.9)	684 (1179.6)	1404 (2472.6)
221AD302		690 (1142.7)	663 (1120.9)	1353 (2263.5)
221AD103		122 (384.7)	63 (188.3)	185 (573.0)
221AD104		13 (6.7)	4 (3.2)	17 (9.9)
Proportion of Pool B from Pool A2				
Proportion of subjects		91.3	95.3	93.2
Proportion of follow up time (years)		86.2	92.3	89.0

NOTE: An additional 119 subjects received BIIB037 in non-pooled studies 221AD101 (N=39), 221HV102 (N=28) and 221AD205 (N=52). These studies differed in population, route of administration, and duration of exposure compared to the pooled studies. These studies were not pooled as their inclusion in the denominator had the potential to dilute the incidence of adverse events.

SOURCE: BIIB037AD/ISS/CSR/T-EX-TOTAL-SCS.SAS

DATE: 06JAN2020

The FDA's Position:

FDA agrees that exposure overall appears to be adequate.

4.2. Summary of Adverse Events

The Applicant's Position

The overall incidence of AEs in Pool A1 was comparable to placebo at all doses of aducanumab (Table 21). Most participants in the placebo and 10 mg/kg aducanumab group had AEs with a maximum severity of mild or moderate. The incidence of severe AEs was higher in the 10 mg/kg group (14.5%) than the placebo group (8.5%). The most common AEs (incidence $\geq 5\%$) in the 10 mg/kg aducanumab group that also occurred at an incidence $\geq 2\%$ higher than placebo were ARIA-E, headache, ARIA-H microhemorrhage, fall, ARIA-H superficial siderosis, and diarrhea (Table 22). AEs associated with ARIA are described in this section and ARIA is discussed in detail in Section 4.3.

The overall incidence of SAEs in Pool A1 was also comparable to placebo (13.9%) at all doses of aducanumab, including 10 mg/kg (13.6%) [Table 21]. The majority of SAEs reported during the placebo-controlled period of the Phase 3 studies were generally consistent with the underlying diagnosis of Alzheimer's disease and the expected comorbidities for the age of the study population. The most common SAE (incidence $\geq 1\%$) in the 10 mg/kg aducanumab group that occurred at a higher incidence with aducanumab was ARIA-E ($<0.1\%$ placebo; 1.3% 10 mg/kg). In addition to ARIA-E, the incidence of serious events of ARIA-H microhemorrhage and superficial siderosis was higher in the 10 mg/kg aducanumab group (0.3% and 0.5%, respectively) than the placebo group (0%, both events).

There were 31 deaths reported in the aducanumab clinical development program during clinical studies. All deaths were reported in Studies 301, 302, and 103. One fatal event of intracranial hemorrhage in Study 103 was assessed as related to study treatment; none of the other fatal events was assessed by the investigator as related to study treatment. No deaths were considered related to AEs of ARIA. Reported causes of death were consistent with those that are common in patients with Alzheimer's disease and with the expected comorbidities for the age of the study population. Eighteen of the deaths were reported in the placebo-controlled period of the 3 studies, and the incidence of deaths between placebo and aducanumab was balanced (16 in the Phase 3 studies: 5 [0.5%] placebo versus 11 [0.5%] total aducanumab, with 8 [0.8%] in the 10 mg/kg group (Table 21); 2 in Study 103: 1 placebo and 1 aducanumab). The other 13 deaths were reported while participants were enrolled in the long-term extension period of the studies, when all participants received aducanumab (8 in the Phase 3 studies; 5 in Study 103).

In Pool A1 there was a higher incidence of AEs that led to discontinuation of study treatment among aducanumab-treated participants (8.8% in the 10 mg/kg group; 4.1% in the placebo group) [Table 21]. The higher incidence of treatment discontinuation was due to the increased incidence of ARIA events, including ARIA-H microhemorrhage (0% placebo; 2.0% 10 mg/kg), ARIA-H superficial siderosis (0.2% placebo; 3.0% 10 mg/kg), and ARIA-E ($<0.1\%$ placebo; 1.3% 10 mg/kg). Specific criteria in the Phase 3 protocols required permanent discontinuation of study treatment due to ARIA based on the type of ARIA, the radiographic severity of the event, and the presence and severity of clinical symptoms. These criteria evolved over the course of the studies and are described in Table 24. No other AEs leading to discontinuation of treatment in

the 10 mg/kg group were reported for more than 2 participants (0.2%) or the incidence was higher in the placebo group (cerebral hemorrhage: 0.4% placebo; 0.3% 10 mg/kg).

In the aducanumab-treated period of the Phase 3 studies (Pool A2, placebo-controlled and long-term extension), the AE profile with longer-term exposure to aducanumab was consistent with that seen in Pool A1. There was no notable increase in the overall incidence or incidence rates of AEs or in the incidence of common AEs (either by SOC, High Level Group Term, or preferred term). The overall incidence of AEs was 88.6% (Pool A1, 91.6%) in the 10 mg/kg aducanumab group, with an incidence rate of 51.9 per 100 person-years (Pool A1, 64.0). The type and severity of AEs with longer-term exposure to aducanumab was also similar to that seen in Pool A1. The incidence of severe AEs in the 10 mg/kg group was 16.1% (Pool A1, 14.5%), with an incidence rate of 9.4 per 100 person-years (Pool A1, 10.1). Review of SAE data revealed no notable increases in the incidence of all or related SAEs with longer-term exposure to aducanumab than those already identified for Pool A1. The overall incidence of SAEs was 15.6% (Pool A1, 13.6%) in the 10 mg/kg group, with an incidence rate of 9.1 per 100 person-years (Pool A1, 9.5). There were no notable increases in the incidence of AEs leading to discontinuation of study treatment with longer-term exposure to aducanumab. The overall incidence of AEs that led to discontinuation of treatment was 9.2% (Pool A1, 8.8%) in the 10 mg/kg group, with an incidence rate of 5.4 per 100 person-years (Pool A1, 6.2).

Table 21: Overall Summary of Adverse Events – Pool A1

Overall summary of adverse events - Pool A1					
	Placebo (N=1087) n (%)	BIIB037 3 mg/kg (N=760) n (%)	BIIB037 6 mg/kg (N=405) n (%)	BIIB037 10 mg/kg (N=1033) n (%)	BIIB037 total (N=2198) n (%)
Number of subjects with any event	945 (86.9)	700 (92.1)	347 (85.7)	946 (91.6)	1993 (90.7)
Severity (a)					
Mild	445 (40.9)	252 (33.2)	122 (30.1)	331 (32.0)	705 (32.1)
Moderate	408 (37.5)	328 (43.2)	177 (43.7)	465 (45.0)	970 (44.1)
Severe	92 (8.5)	120 (15.8)	48 (11.9)	150 (14.5)	318 (14.5)
Related event (b)	273 (25.1)	373 (49.1)	148 (36.5)	530 (51.3)	1051 (47.8)
Serious event	151 (13.9)	105 (13.8)	54 (13.3)	141 (13.6)	300 (13.6)
Related serious event	8 (0.7)	9 (1.2)	7 (1.7)	21 (2.0)	37 (1.7)
Events leading to study drug discontinuation	45 (4.1)	65 (8.6)	45 (11.1)	91 (8.8)	201 (9.1)
Events leading to study withdrawal	31 (2.9)	32 (4.2)	27 (6.7)	38 (3.7)	97 (4.4)
Number of deaths	5 (0.5)	3 (0.4)	0	8 (0.8)	11 (0.5)

NOTE 1: A subject can appear in more than one category.
(a) Each subject was counted once at maximum severity.
(b) Related to study drug as assessed by the investigator.

SOURCE: BIIB037AD/ISS/CSR/T-AE-SUM-A1.SAS

Data as of: 06NOV2019 Run date: 08JAN2020

Table 22: Adverse Events With at Least 5% Incidence in Aducanumab 10 mg/kg and 2% Higher Incidence Than Placebo – Pool A1

Adverse events with at least 5% incidence in BIIB037 10 mg/kg and 2% higher incidence than placebo - Pool A1

	Placebo (N=1087) n (%)	BIIB037 3 mg/kg (N=760) n (%)	BIIB037 6 mg/kg (N=405) n (%)	BIIB037 10 mg/kg (N=1033) n (%)	BIIB037 total (N=2198) n (%)
Number of subjects with any event	945 (86.9)	700 (92.1)	347 (85.7)	946 (91.6)	1993 (90.7)
Amyloid related imaging abnormality-oedema/effusion	29 (2.7)	223 (29.3)	83 (20.5)	362 (35.0)	668 (30.4)
Headache	165 (15.2)	161 (21.2)	58 (14.3)	212 (20.5)	431 (19.6)
Amyloid related imaging abnormality-microhaemorrhages and haemosiderin deposits	71 (6.5)	141 (18.6)	50 (12.3)	197 (19.1)	388 (17.7)
Fall	128 (11.8)	105 (13.8)	50 (12.3)	155 (15.0)	310 (14.1)
Superficial siderosis of central nervous system	24 (2.2)	91 (12.0)	23 (5.7)	151 (14.6)	265 (12.1)
Diarrhoea	74 (6.8)	62 (8.2)	27 (6.7)	92 (8.9)	181 (8.2)

NOTE 1: A subject was counted only once within each preferred term (MedDRA version 22.0).

NOTE 2: Preferred terms are presented in decreasing frequency of the table's BIIB037 10 mg/kg column.

NOTE 3: Preferred terms are displayed if the incidence is at least 5% in the BIIB037 10 mg/kg and the incidence difference compared to placebo is at least 2%.

SOURCE: BIIB037AD/ISS/CSR/T-AE-PT-ADR-10MG-A1.SAS

Data as of: 06NOV2019 Run date: 13JAN2020

The FDA's Position:

FDA agrees that the incidence of SAEs at all doses of aducanumab in Pool A1 is comparable to placebo and that the most common SAEs were related to ARIA-E. We note that SAEs of intracranial hemorrhage (including cerebral hemorrhage, cerebral hematoma, subarachnoid hemorrhage, and intracranial hemorrhage) occurred in fewer than 1% of patients and were balanced between aducanumab 10 mg/kg patients and placebo patients in Pool A1. FDA analysis resulted in small differences in event numbers compared to the findings reported by the applicant. For example, FDA analysis shows SAEs of ARIA-E in 1.4% of aducanumab 10 mg/kg and 0.1% of placebo subjects in Pool A1. The differences result from FDA reassignment to the 10 mg/kg group of 72 subjects that the Applicant evaluated in the 6 mg/kg assigned treatment group for safety despite a protocol change but included in the 10 mg/kg treatment group for evaluation of efficacy. FDA reassignment was performed to be consistent with the efficacy analysis. These differences do not lead to a difference in the conclusions. FDA analysis of Study 302 alone shows results comparable to Pool A1. FDA agrees with conclusions regarding the long-term extension studies.

FDA agrees that there were 31 deaths across the development program, that there was not an excess of deaths in the aducanumab group compared to placebo in Pool A1, and that the deaths that occurred were not related to ARIA.

FDA agrees with conclusions regarding discontinuations.

FDA notes that among all patients in Pool A1, in addition to ARIA TEAEs, the most common TEAEs were headache (20% for 10 mg/kg vs 15% for placebo), fall (15% for 10 mg/kg vs 12% for placebo), dizziness (10% for 10 mg/kg vs 9% for placebo). Findings were similar for Study 302 alone. FDA notes that among patients who did not have TEAEs related to ARIA in Pool A1, the most common TEAEs included fall (15% for 10 mg/kg vs 12% for placebo), headache (16% for 10 mg/kg vs 14% for placebo), and diarrhea (10% for 10 mg/kg vs 7% for placebo), with similar findings in Study 302 alone.

4.3. Summary of Adverse Events of Special Interest

The Applicant's Position

Amyloid related imaging abnormalities (ARIA) represent a spectrum of imaging findings detected on brain MRI. These findings include brain edema or sulcal effusion defined as ARIA-E, and brain microhemorrhage or localized superficial siderosis defined as ARIA-H. In the aducanumab clinical studies, some events of brain hemorrhage >1 cm termed macro-hemorrhage were also considered ARIA-H at the judgment of the central radiologist and the Investigator. ARIA-E and ARIA-H have been detected at an increased incidence during treatment with monoclonal antibodies that target A β , including aducanumab [Greenberg 2020; Penninkilampi 2017; Sevigny 2016; Sperling 2011b], and ARIA-E has been identified as an AE of special interest for aducanumab.

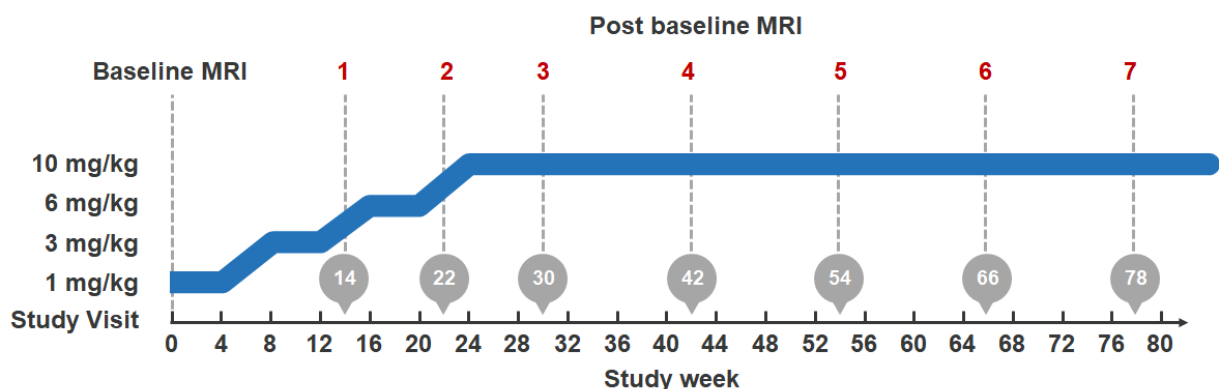
The biological mechanism of ARIA-E is not well understood. Published hypotheses suggest that increased cerebrovascular permeability could be caused either by increased A β clearance from the brain parenchyma, leading to saturation of the perivascular drainage system, and/or by direct

antibody interaction with deposited vascular amyloid, leading to its clearance and weakening of the vessel walls [Greenberg 2020; Zago 2013]. Published data also suggest that ARIA-H (microhemorrhages or superficial siderosis) events tend to accompany ARIA-E [Greenberg 2020]. In addition, brain microhemorrhages and localized superficial siderosis can occur in patients with Alzheimer's disease who have not received anti-A β monoclonal antibodies, suggesting that there is a background rate of such events in patients with Alzheimer's disease [Doody 2014; Salloway 2014].

ARIA risk minimization in the Phase 3 studies consisted of protocol-defined dose titration that occurred over the first 24 weeks of treatment, and a robust plan for monitoring for the potential occurrence of ARIA (Figure 25). ARIA monitoring was complemented by detailed guidance on the clinical management of participants in whom ARIA was detected, including drug disposition rules and follow-up MRI scans.

ARIA monitoring methods and data collection throughout the aducanumab clinical program included routine brain MRI scans performed for all participants at protocol-specified timepoints, follow-up MRI scans performed for participants in whom ARIA was detected, and a centralized MRI reader staffed with expert radiologists highly experienced with ARIA. In the Phase 3 Studies 301 and 302, all study participants had regularly scheduled brain MRIs performed at 5 postbaseline visits during approximately the first year of the placebo-controlled period and at 2 visits during the last 6 months of the placebo-controlled period (Figure 25). Likewise, there were 5 scheduled brain MRI visits during approximately the first year of the long-term extension to monitor for the potential occurrence in participants switching from placebo to aducanumab and performed in all participants to maintain the blind. Subsequently, the frequency of scheduled brain MRIs decreased to biannually and then starting at Year 3.5 to approximately once per year. Follow-up brain MRIs for participants who developed ARIA were performed every 4 weeks until ARIA resolved (ARIA-E) or stabilized (ARIA-H).

Figure 25: MRI Monitoring Frequency in the Placebo-Controlled Treatment Period of the Phase 3 Studies 301 and 302



In all aducanumab clinical studies in Alzheimer's disease, all ARIA findings detected on MRI were reported as AEs by Investigators. Guidelines on the management of participants who developed ARIA included criteria for dosing to be continued, suspended and then resumed, or permanently discontinued due to ARIA findings, depending on the type(s) of ARIA detected, the radiographic severity of the ARIA findings, whether a participant had any clinical symptoms,

and, if symptoms were present, their clinical severity. These criteria evolved over the course of the studies, as experience with ARIA in aducanumab-treated clinical trial participants accumulated. A detailed summary of how ARIA was classified by MRI severity and managed in Studies 301 and 302, including the evolution of ARIA management over the course of the studies is shown in [Table 23](#) and [Table 24](#).

ARIA monitoring and management procedures allowed for a detailed characterization of ARIA as observed in the Phase 3 studies, including the incidence, time course, radiographic severity, symptomatic status, type of symptoms if present, and clinical severity of symptoms if present.

Table 23: Severity of ARIA-E and ARIA-H on MRI

ARIA Type	Radiographic Severity		
	Mild	Moderate	Severe
ARIA-E	FLAIR hyperintensity confined to sulcus and or cortex/subcortical white matter in one location <5 cm	FLAIR hyperintensity 5 to 10 cm, or more than 1 site of involvement, each measuring <10 cm	FLAIR hyperintensity measuring >10 cm, often with significant subcortical white matter and/or sulcal involvement. One or more separate sites of involvement may be noted.
ARIA-H microhemorrhage	≤ 4 new incident microhemorrhages	5 to 9 new incident microhemorrhages	10 or more new incident microhemorrhages
ARIA-H superficial siderosis	1 focal area of superficial siderosis	2 focal areas of superficial siderosis	>2 focal areas of superficial siderosis

Table 24: ARIA Management by Protocol Version - Studies 301 and 302

Version 4 (March 2017): ApoE ε4 Carriers Up Titrate to 10 mg/kg			
Version 1 (April 2015)	Version 3-5 (July 2016-September 2017)	Version 6 (June 2018)	
- Asymptomatic, radiographically mild ARIA continue dosing at current dose and schedule ¹	- No change	- No change	
- Participants who suspend dosing after an episode of ARIA restart and remain at the next lower dose	- Participants who suspend dosing after 1 st ARIA episode to restart at the same dose ²	- No change	
- Participants with severe symptomatic ARIA permanently discontinue treatment	- Participants with severe symptomatic ARIA suspend dosing and resume at the same dose ²	- No change	
- Participants with serious symptomatic ARIA permanently discontinue treatment	- Participants with serious symptomatic ARIA (except other medically important ³) permanently discontinue treatment. Participants with serious (other medically important³) ARIA suspend dosing and resume at the same dose ²	- No change	
- Recurrent ARIA not specifically addressed, managed as the initial episode	- Participants who suspend dosing after a 2 nd ARIA episode to resume at the next lower dose ² - Participants who suspend dosing after a 3 rd ARIA episode discontinue treatment	- Participants who suspend dosing after a 2 nd ARIA episode continue treatment at the same dose ² - Participants who suspend dosing after a 3 rd ARIA episode continue treatment at the same dose ²	
- Participant with radiographically severe ARIA-H (≥10 new incident microhemorrhages or >2 superficial siderosis areas permanently discontinue	- Participants with radiographically severe ARIA-H (≥10 new incident microhemorrhages or >2 superficial siderosis areas), or macro-hemorrhage permanently discontinue	- No change	

¹Suspend dosing only if ARIA becomes symptomatic or radiographically moderate. Radiographically severe ARIA-E also suspends dosing, while severe ARIA-H permanently discontinues.

²Following resumption of dosing, titration to continue (if applicable).

³"Other medically important events" requiring dose suspension include SAEs that are not life-threatening (in the opinion of the Investigator), do not require inpatient hospitalization or prolongation of existing hospitalization, and do not result in significant/permanent disability or congenital anomalies/fetal defects, but may (in the opinion of the Investigator) jeopardize the participant or may require intervention.

The FDA's Position:

FDA agrees that the Applicant approached an anticipated adverse event in a structured, systematic way with scheduled MRIs and specified protocols for ARIA management.

4.3.1. ARIA Incidence

Analyses specific to ARIA are based on the safety MRI population, consisting of those dosed participants with at least 1 postbaseline MRI scan.

ARIA-E

During the placebo-controlled period of the Phase 3 studies (Pool A1), treatment with aducanumab (30.7%) increased the overall incidence of ARIA-E relative to placebo (2.7%) (Table 25). The incidence of ARIA-E was higher in the 10 mg/kg group (35.2%) compared to lower dose groups (3 mg/kg: 29.5%; 6 mg/kg: 21.2%). In addition to aducanumab dose, the incidence of ARIA-E was related to ApoE ϵ 4 status, with a higher incidence in ApoE ϵ 4 carriers compared to noncarriers, most notably in the 10 mg/kg dose group (carriers: 43.0%; noncarriers: 20.3%).

In Pool A1, 10.6% of participants in the 10 mg/kg group had more than 1 ARIA-E event. In the majority of participants in the 10 mg/kg group with a second ARIA-E event (any radiographic severity), the second episode occurred within 24 weeks of resolution of the first episode (30.6% within 0-12 weeks; 53.7% >12-24 weeks).

ARIA-H

Among aducanumab-treated participants in Pool A1 the overall incidence of ARIA-H was also higher than placebo (8.7%) and like ARIA-E, the incidence of ARIA-H was highest in the 10 mg/kg group (28.3%) compared to lower dose groups (3 mg/kg: 25.5%; 6 mg/kg: 16.1%) [Table 25]. Likewise, the incidence of ARIA-H microhemorrhage (placebo: 6.6%; 10 mg/kg: 19.1%) and ARIA-H superficial siderosis (placebo: 2.2%; 10 mg/kg: 14.7%) was higher in participants treated with aducanumab compared to placebo. ARIA-H macro-hemorrhages occurred rarely in the placebo group and in all aducanumab dose groups (placebo: 0.4%; 10 mg/kg: 0.3%).

Further analyses of the incidence of ARIA-H stratified by ARIA-E status showed that the increased incidence of ARIA-H with aducanumab treatment relative to placebo was due to the excess ARIA-H microhemorrhage and ARIA-H superficial siderosis in aducanumab-treated participants who also had ARIA-E (Table 26). Specifically, whereas the incidence of ARIA-H microhemorrhage (10 mg/kg: 40.3%) and ARIA-H superficial siderosis (10 mg/kg: 38.7%) was elevated for aducanumab-treated participants who also had ARIA-E, the incidence of ARIA-H microhemorrhage (10 mg/kg: 7.6%) and ARIA-H superficial siderosis (10 mg/kg: 1.6%) in those without ARIA-E at any time was similar to placebo-treated participants (ARIA-H microhemorrhage: 6.6%, ARIA-H superficial siderosis: 2.2%) [Table 25].

Furthermore, most occurrences of ARIA-H microhemorrhage or superficial siderosis were concurrent with episodes of ARIA-E (i.e., the events temporally overlapped) in all aducanumab dose groups. In the 10 mg/kg group, most events of radiographically mild (57.4%), moderate (84.1%), and severe (82.1%) ARIA-H microhemorrhages, as well as mild (84.8%), moderate (91.7%), and severe (91.4%) ARIA-H superficial siderosis events were concurrent with ARIA-E. In addition, in most participants the events were initially detected either the same day as ARIA-E or during a subsequent MRI performed to follow up on the same ARIA-E episode.

Table 25: ARIA Events Based on AE eCRF - Pool A1 Safety MRI Population

ARIA events based on AE eCRF - Pool A1 safety MRI population					
	Placebo (N=1076) n (%)	BIIB037 3 mg/kg (N=756) n (%)	BIIB037 6 mg/kg (N=392) n (%)	BIIB037 10 mg/kg (N=1029) n (%)	BIIB037 total (N=2177) n (%)
Number of subjects with					
ARIA-E	29 (2.7)	223 (29.5)	83 (21.2)	362 (35.2)	668 (30.7)
ARIA-H	94 (8.7)	193 (25.5)	63 (16.1)	291 (28.3)	547 (25.1)
ARIA-H microhemorrhage	71 (6.6)	141 (18.7)	50 (12.8)	197 (19.1)	388 (17.8)
ARIA-H macrohemorrhage	4 (0.4)	1 (0.1)	3 (0.8)	3 (0.3)	7 (0.3)
ARIA-H superficial siderosis	24 (2.2)	91 (12.0)	23 (5.9)	151 (14.7)	265 (12.2)
ARIA-E and ARIA-H	12 (1.1)	142 (18.8)	42 (10.7)	228 (22.2)	412 (18.9)
Concurrent ARIA-E and ARIA-H	12 (1.1)	135 (17.9)	39 (9.9)	216 (21.0)	390 (17.9)
ARIA-E or ARIA-H	111 (10.3)	274 (36.2)	104 (26.5)	425 (41.3)	803 (36.9)
Isolated ARIA-E	17 (1.6)	81 (10.7)	41 (10.5)	134 (13.0)	256 (11.8)
Isolated ARIA-H	82 (7.6)	51 (6.7)	21 (5.4)	63 (6.1)	135 (6.2)

NOTE 1: ARIA-E, ARIA-H microhemorrhage and ARIA-H superficial siderosis are identified based on preferred term (MedDRA 22.0). ARIA-H macrohemorrhage is identified based on eCRF reported term.

NOTE 2: Concurrent is defined as overlapping in MRI duration of 2 ARIA events.

NOTE 3: ARIA-E is defined as isolated if the subject experienced no ARIA-H during the placebo-controlled period. ARIA-H is isolated if subject experienced no ARIA-E during the placebo-controlled period.

SOURCE: BIIB037AD/ISS/CSR/T-ARIA-AE-A1.SAS

Data as of: 06NOV2019 Run date: 02APR2020

Table 26: Summary of ARIA-H Events Based on Adverse Event eCRF by ARIA-E Status – Pool A1 Safety MRI Population

Summary of ARIA-H events based on adverse event eCRF by ARIA-E status - Pool A1 safety MRI population

	Placebo (N=1076) n (%)	BIIB037 3 mg/kg (N=756) n (%)	BIIB037 6 mg/kg (N=392) n (%)	BIIB037 10 mg/kg (N=1029) n (%)	BIIB037 total (N=2177) n (%)
Number of subjects with ARIA-E	29 (100)	223 (100)	83 (100)	362 (100)	668 (100)
Number of subjects with ARIA-H (a)	12 (41.4)	142 (63.7)	42 (50.6)	228 (63.0)	412 (61.7)
ARIA-H microhemorrhage	4 (13.8)	103 (46.2)	32 (38.6)	146 (40.3)	281 (42.1)
ARIA-H macrohemorrhage	2 (6.9)	1 (0.4)	2 (2.4)	2 (0.6)	5 (0.7)
ARIA-H superficial siderosis	9 (31.0)	76 (34.1)	21 (25.3)	140 (38.7)	237 (35.5)
Number of subjects with no ARIA-E	1047 (100)	533 (100)	309 (100)	667 (100)	1509 (100)
Number of subjects with ARIA-H	82 (7.8)	51 (9.6)	21 (6.8)	63 (9.4)	135 (8.9)
ARIA-H microhemorrhage	67 (6.4)	38 (7.1)	18 (5.8)	51 (7.6)	107 (7.1)
ARIA-H macrohemorrhage	2 (0.2)	0	1 (0.3)	1 (0.1)	2 (0.1)
ARIA-H superficial siderosis	15 (1.4)	15 (2.8)	2 (0.6)	11 (1.6)	28 (1.9)

NOTE 1: ARIA-E, ARIA-H microhemorrhage and ARIA-H superficial siderosis are identified based on preferred term (MedDRA version 22.0). ARIA-H macrohemorrhage is identified based on eCRF reported term.

(a) May or may not have overlapped with ARIA-E.

SOURCE: BIIB037AD/ISS/CSR/T-ARIA-AE-BYARIAE-A1.SAS

Data as of: 06NOV2019 Run date: 13JAN2020

4.3.2. Characteristics of ARIA

Time Course

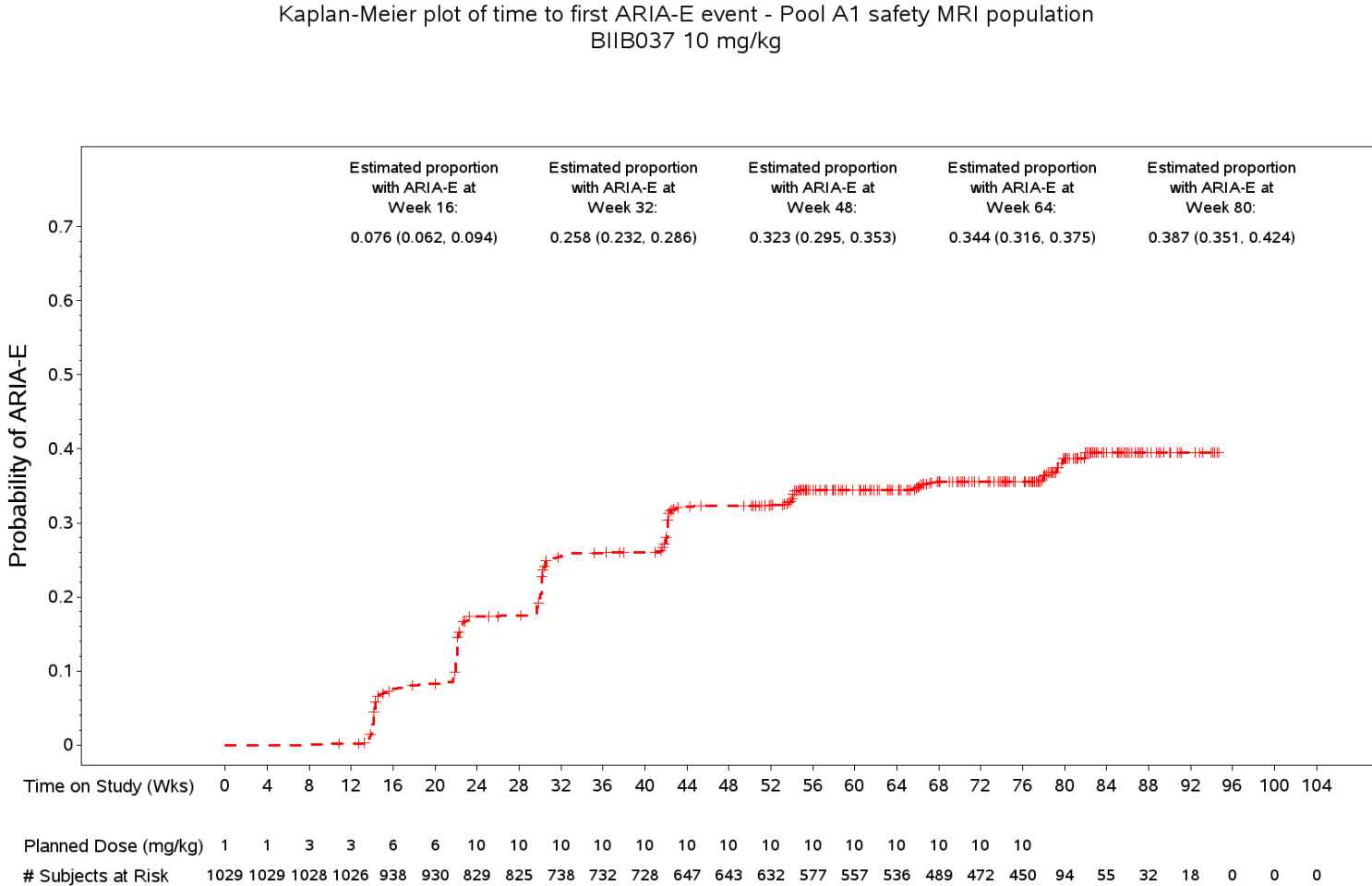
Onset of first ARIA-E events was evaluated by number of doses of aducanumab received prior to the first event based on data from 815 participants with ARIA-E included in Pool A2 (the aducanumab-treated experience which includes both the placebo-controlled and long-term extension). First ARIA-E events typically occurred early in treatment, with over 70% occurring during the first 8 doses (10 mg/kg: 73.1%). Similarly, the first symptomatic ARIA-E events (10 mg/kg: 70.5%), and the first radiographically moderate or severe ARIA-E events (10 mg/kg: 77.7%) had onset most frequently during the first 8 doses.

Analysis of the last dose of aducanumab received prior to onset of the first ARIA-E event showed that of 453 participants with ARIA-E in the aducanumab 10 mg/kg group, 29.4% of participants received 3 mg/kg, 29.8% received 6 mg/kg, and 40.6% received 10 mg/kg prior to onset of the first event. Similar results were observed for the 362 participants with ARIA-E in the 10 mg/kg group in Pool A1.

A Kaplan-Meier plot of time to first ARIA-E events for 1029 participants in Pool A1 is visually depicted for the 10 mg/kg group ([Figure 26](#)) and shows that onset of the first event primarily occurred within 32 weeks of starting treatment, corresponding to the first 8 doses. The probability of a first ARIA-E event decreased between Weeks 32 and 44 and after Week 44, occurrences of first events for these participants were infrequent through Week 192 of the study, as shown in [Figure 27](#). As described under “ARIA-H incidence” in Section [4.3.1](#), most ARIA-H events were concurrent with ARIA-E.

Further analyses of first ARIA-E events in Pool A1 by ApoE ϵ 4 carrier status showed that in 10 mg/kg noncarriers, onset of the first event occurred most notably between approximately Weeks 12 and 32, during titration to 10 mg/kg and while receiving 10 mg/kg ([Figure 28](#)). In 10 mg/kg carriers, onset of first ARIA-E events occurred most notably between approximately Weeks 12 and 44 during titration to 10 mg/kg and while receiving 10 mg/kg. ([Figure 29](#)).

**Figure 26: Kaplan-Meier Plot of Time to First ARIA-E Event – Pool A1 Safety MRI Population
Aducanumab 10 mg/kg**

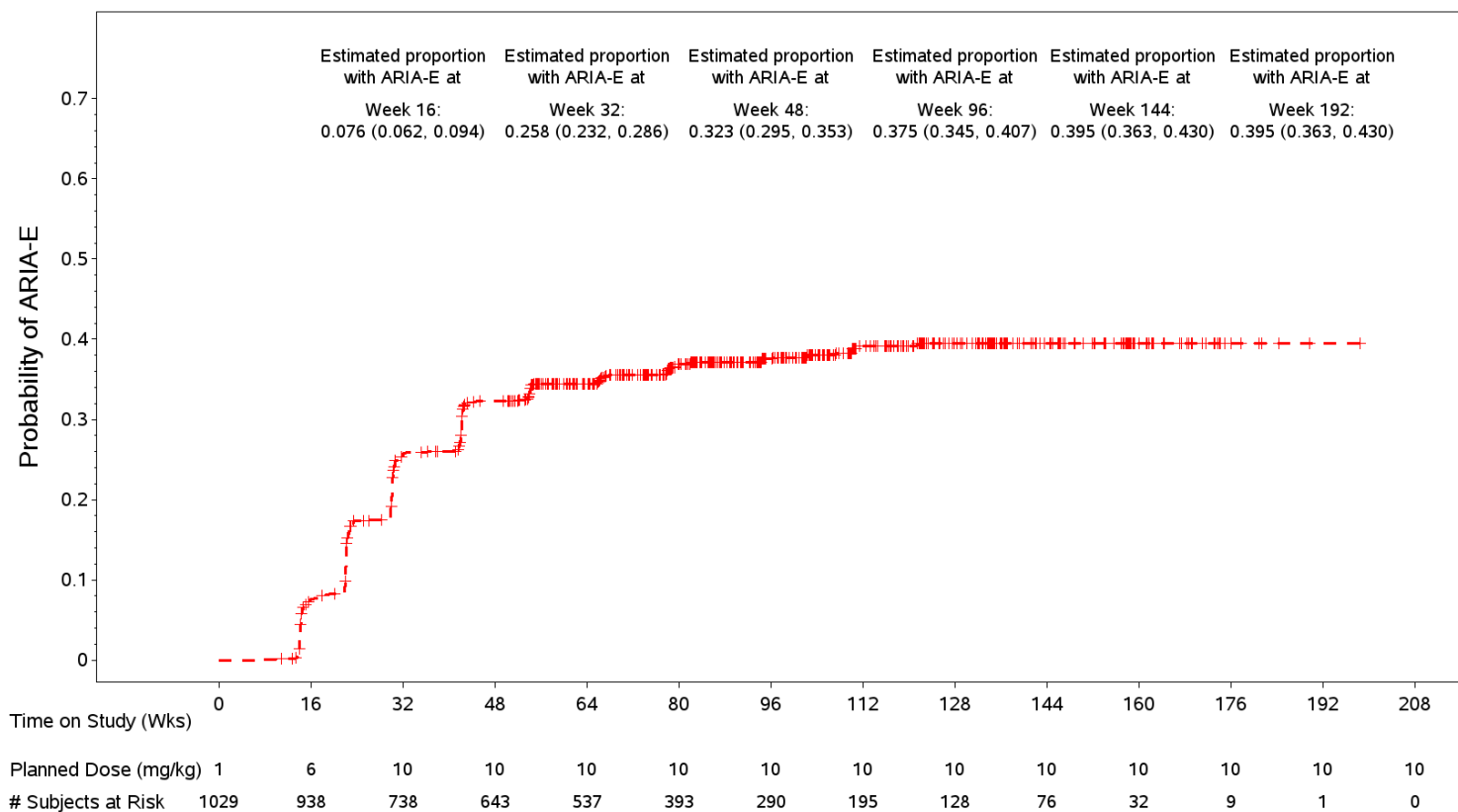


NOTE: "+" indicates censored data.
SOURCE: BIIB037AD/ISS/CSR/F-ARIAE-KM-A1.SAS

Data as of: 06NOV2019 Run date: 22JUN2020

Figure 27: Kaplan-Meier Plot of Time to First ARIA-E Event in Pool A2 For Those in the Pool A1 Aducanumab 10 mg/kg Group – Pool A2 Safety MRI Population

Kaplan-Meier plot of time to first ARIA-E event in Pool A2 for those in the Pool A1 BIIB037 10 mg/kg group
- Pool A2 safety MRI population

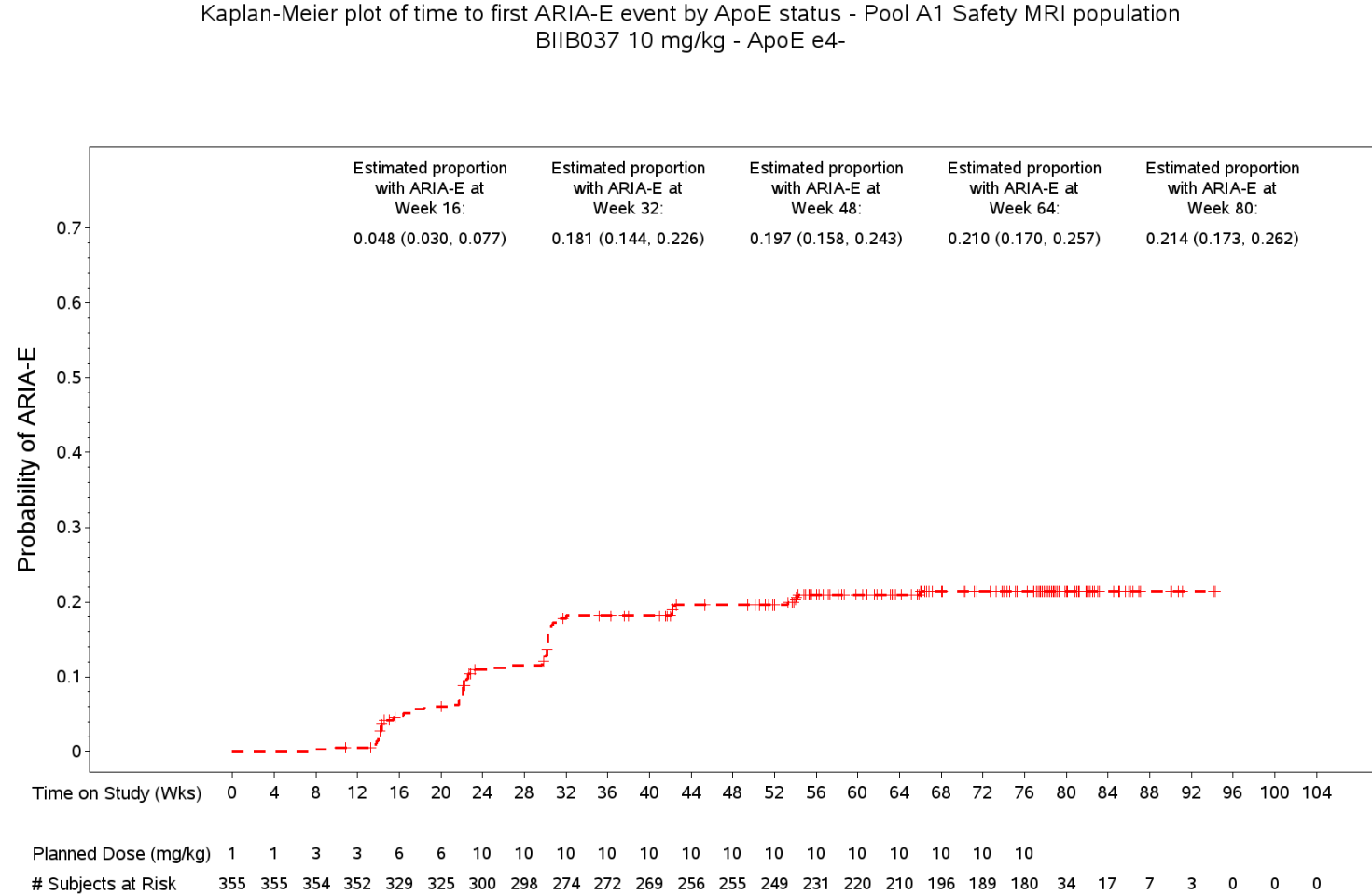


NOTE: "+" indicates censored data.

SOURCE: BIIB037AD/ISS/ADHOC/F-ARIAE-KM-EARLY10-A2.SAS

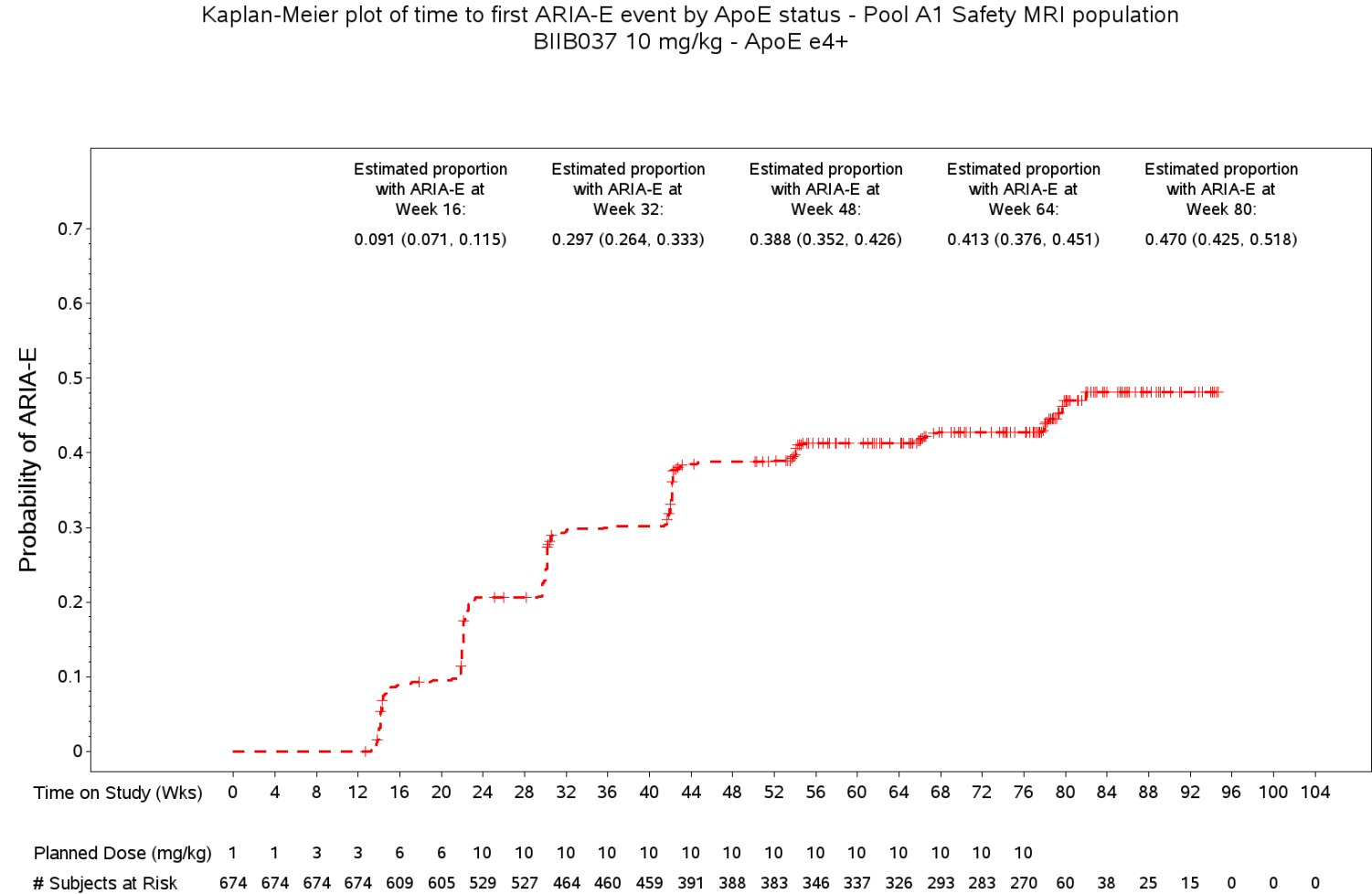
Data as of: 06NOV2019 Run date: 09SEP2020

**Figure 28: Kaplan-Meier Plot of Time to First ARIA-E Event by ApoE Status – Pool A1 Safety MRI Population
Aducanumab 10 mg/kg ApoE e4-**



NOTE: "+" indicates censored data.
SOURCE: BIIB037AD/ISS/CSR/F-ARIAE-KM-RAP-A1.SAS
Data as of: 06NOV2019 Run date: 22JUN2020

Figure 29: Kaplan-Meier Plot of Time to First ARIA-E Event by ApoE Status – Pool A1 Safety MRI Population
Aducanumab 10 mg/kg ApoE e4+



NOTE: "+" indicates censored data.
SOURCE: BIIB037AD/ISS/CSR/F-ARIAE-KM-RAP-A1.SAS

Radiographic Severity and Resolution of ARIA

During the placebo-controlled period (Pool A1), most participants in the aducanumab 10 mg/kg group with ARIA-E (N=362) had first events with a maximum radiographic severity of mild (33.4%) or moderate (56.6%). Approximately 10% of participants with ARIA-E had radiographically severe ARIA-E events.

The majority of first ARIA-E events resolved within 12 weeks in both the 10 mg/kg (68.5%) and total aducanumab (67.7%) groups. By 16 weeks, approximately 83% of events in both groups had resolved, and by 20 weeks about 92% of events in both groups had resolved.

Almost all ARIA-E events in both aducanumab dose groups in Pool A1 resolved while on study (10 mg/kg: 98.2%; total: 98.4%).

In the 10 mg/kg group (Pool A1), for most events of ARIA-H microhemorrhage (radiographically mild 63.8%, moderate 90.9%, and severe 64.3%) and of ARIA-H superficial siderosis (radiographically mild 94.3%, moderate 91.7%, and severe 88.6%), the ARIA-H radiographic findings stabilized within 8 weeks.

Symptomatic ARIA

If a participant experienced an AE considered related to an ARIA event in the judgment of the Investigator, the ARIA event was considered symptomatic ARIA and the ARIA-related AE was recorded as a symptom. In Pool A1, 103 of 1029 participants (10.0%) in the 10 mg/kg group and 176 of 2177 (8.1%) in the total aducanumab group experienced symptomatic ARIA (either ARIA-E or ARIA-H). Most participants who experienced ARIA were asymptomatic (10 mg/kg: 75.8%).

Among aducanumab-treated participants, the most frequent type of symptomatic ARIA was ARIA-E (10 mg/kg: 9.1%; Pool A1), followed by ARIA-H microhemorrhage (10 mg/kg: 2.7%), ARIA-H superficial siderosis (10 mg/kg: 2.4%), and ARIA-H macro-hemorrhage (10 mg/kg: 0.2%). The incidence of symptomatic ARIA-E among participants who had concurrent ARIA-E and ARIA-H (10 mg/kg: 25.5%) was similar to those with only ARIA-E (10 mg/kg: 22.3%).

Among participants with ARIA-E, the incidence of symptomatic ARIA-E was similar in ApoE ε4 carriers (10 mg/kg: 26.6%) and noncarriers (10 mg/kg: 23.6%).

The most common symptoms among participants in the 10 mg/kg group with symptomatic ARIA (ARIA-E or ARIA-H) were headache (46.6%), confusion (14.6%) and dizziness (10.7%). Other reported symptoms included nausea (7.8%), fatigue (4.9%), and blurred vision (4.9%).

The symptomatic severity of a symptomatic ARIA event is defined as the reported maximum clinical severity of the symptoms that temporally overlapped with the radiographic duration of the ARIA episode. When symptoms were reported during an ongoing ARIA episode (ARIA-E or ARIA-H), most were mild or moderate in the reported clinical severity (10 mg/kg, greatest severity among 425 participants with ARIA: 15.8% mild, 6.6% moderate, 0.9% severe symptoms).

Severe symptoms during an ARIA episode in Pool A1 were reported for a total of 10 of 2177 aducanumab-treated participants (0.5%), including 4 of 1029 participants (0.4%) in the 10 mg/kg group. Four of these 10 participants had a severe headache reported as a symptom during an ARIA episode; all other severe symptoms that occurred during an ARIA episode were each

reported by a single participant and included, but were not limited to, seizure, confusional state, muscular weakness, and dizziness. The symptomatic ARIA events with severe symptoms occurred early during treatment, within the first 8 aducanumab doses; all symptomatic ARIA episodes in Pool A1 aducanumab-treated participants with onset after the first 8 doses had ARIA-related AEs that were mild or moderate in clinical severity.

ARIA symptoms were typically transient, and the majority resolved (10 mg/kg: 86.9%). Approximately 70% were documented to have resolved within 12 weeks and approximately 79% resolved within 16 weeks.

The incidence and symptomatic severity of symptomatic ARIA events in Pool A2 was similar to that seen in Pool A1.

Given that seizure events and CNS hemorrhages (including ARIA-H macro-hemorrhages) were rarely reported in participants with ARIA-E, analyses were performed utilizing the Standardized MedDRA Queries (SMQs) for Convulsions and for Haemorrhagic CNS Vascular Conditions in order to perform a comprehensive assessment of these events reported in aducanumab and placebo groups.

The overall incidence rate of seizures was balanced between the placebo and aducanumab groups in the Phase 3 studies (all aducanumab doses combined: 0.2 per 100 person-years; placebo group: 0.6 per 100 person-years). Seizure events were reported overall in 22 of 2959 aducanumab-treated participants (Pool B). In 6 of the 22 participants with seizure, seizures coincided with ARIA-E. Given the temporal overlap of these seizure events with ARIA-E episodes, it is at least possible that these events were related to the occurrence of ARIA-E. All 6 seizure events occurred within the first 8 doses.

The overall incidence rate of hemorrhagic CNS vascular events was low and balanced between aducanumab (0.7 per 100 person-years) and placebo (0.6 per 100 person-years) in the placebo-controlled period of the Phase 3 studies. In Pool B, CNS hemorrhage events were reported in 10 participants concurrently with ARIA-E. In addition, some of these hemorrhagic events were anatomically proximal to ARIA-E. While treatment with aducanumab does not appear to increase the overall incidence of hemorrhagic CNS vascular events, it is possible that these events were related to the occurrence of ARIA-E. They typically occurred early during treatment, within the first 8 doses in 9 participants.

The FDA's Position:

FDA agrees that aducanumab results in dose-related ARIA-E and ARIA H and that the risk is increased in ApoE ϵ 4 carriers compared to non-carriers. FDA agrees that the risk appears to occur early during treatment. FDA notes that onset of the first event of ARIA-E in ApoE ϵ 4 carriers differs from noncarriers as shown in [Figure 28](#) and [Figure 29](#).

FDA agrees that the risk for intracranial hemorrhage (including cerebral hemorrhage, cerebral hematoma, subarachnoid hemorrhage, and intracranial hemorrhage) is low and balanced between aducanumab 10 mg/kg and placebo, with an incidence of 0.5% for 10 mg/kg and 0.5% for placebo in Pool A1.

4.4. Summary of Other Safety Areas of Interest

The Applicant's Position

No clinically significant laboratory, vital sign, urinalysis, ECG trends, or Columbia Suicide Severity Rating Scale (C-SSRS) results were identified with aducanumab treatment. Treatment emergent anti-aducanumab antibodies were rare and did not correlate with AEs. Two aducanumab-treated participants (Pool A2, 3 mg/kg group) that tested negative for anti-aducanumab antibodies had SAEs of hypersensitivity that were considered to be related to treatment.

Clinical Laboratory Evaluations

Overall, blood chemistry, hematology, and urinalysis results that met criteria as potentially clinically significant were reported infrequently and for similar proportions of participants in the placebo and aducanumab groups in Pool A1.

Vital Signs

No notable trends were observed for any of the vital sign measurements in the placebo or aducanumab groups in Pool A1. Increases or decreases in systolic or diastolic blood pressure or heart rate that met criteria as potentially clinically relevant were balanced with placebo.

ECG

The incidence of abnormal ECG findings was similar in the placebo and aducanumab groups in Pool A1.

Hypersensitivity

Two aducanumab-treated participants experienced related SAEs consistent with a hypersensitivity reaction. Both events occurred in the 3 mg/kg group. One participant experienced a serious event of allergic reaction to investigational product (preferred term: hypersensitivity) and the second participant experienced serious events of angioedema and urticarial rash (preferred terms: angioedema, urticaria). These events occurred during aducanumab infusions, and no alternative etiologic factors were identified to explain the events. For both participants, treatment with aducanumab was discontinued due to the events. The events resolved and anti-aducanumab antibody testing was negative in both participants.

Immunogenicity

The incidence of anti-aducanumab antibodies was very low and the incidence of treatment-emergent antibodies in Pool A1 was <1.0% in all dose groups. There was no apparent relationship between incidence of anti-aducanumab antibodies and dose group. There was no apparent impact of anti-aducanumab antibodies on the incidence of AEs.

Depression and Suicide

Psychiatric disorders are commonly associated with Alzheimer's disease, and individuals with psychiatric comorbidities (e.g., depression) may be at an increased risk for suicide [Dubois 2016]. Participants were evaluated with the C-SSRS, with suicidal ideation and suicidal behavior assessed at postbaseline visits throughout the placebo-controlled and long-term extension periods.

In Pool A1, a similar proportion of participants in the aducanumab and placebo groups reported suicidal ideation (placebo: 4.5%; 10 mg/kg: 3.2%) or exhibiting suicidal behavior (placebo: <0.1%; 10 mg/kg: <0.1%) at any postbaseline assessment.

SAEs of attempted suicide were reported for 1 participant in the placebo group and 2 participants in the aducanumab 3 mg/kg group, and 1 participant in the aducanumab 10 mg/kg group completed suicide. In the placebo group, there was also 1 participant who had an SAE of intentional self-injury, and 1 participant who had an SAE of suicidal ideation.

Based on these data, aducanumab is not expected to be associated with increases in suicidal ideation or behavior.

The FDA's Position:

The Applicant's analysis did not identify signals of concern for laboratory evaluations or vital signs analysis. FDA agrees with the applicant's conclusions regarding ECG analysis.

FDA agrees with the assessment of the hypersensitivity SAEs.

The low incidence of treatment emergent anti-aducanumab antibodies precludes an assessment of immunogenicity.

FDA agrees that there does not appear to be a signal for increased suicidal ideation or behavior in patients treated with aducanumab.

4.5. Safety Summary and Conclusions

The Applicant's Position

Aducanumab, when dosed at 10 mg/kg IV every 4 weeks after a titration period of 24 weeks, has an acceptable safety profile that would support use in individuals with Alzheimer's disease.

ARIA-E was the most common AE and it was mostly asymptomatic and transient. ARIA-E typically occurred during the first 8 doses of aducanumab. The incidence of ARIA-E was higher in ApoE ε4 carriers compared with noncarriers, whereas the incidence of symptomatic ARIA-E was comparable between carriers and noncarriers. Brain microhemorrhages and localized superficial siderosis were typically asymptomatic and were more common in patients with ARIA-E than in patients without ARIA-E.

Dose titration, routine MRI surveillance, and temporary dose suspension for certain ARIA events were used as risk minimization measures in clinical studies of aducanumab. Based on the profile of ARIA as established during the Phase 3 clinical trials (largely asymptomatic, with rare clinically significant symptoms), dose titration and temporary dose suspension for certain ARIA events are proposed for continued use as important components of ARIA risk minimization.

While severe symptoms associated with ARIA are uncommon, these symptoms are most likely to occur during the time period that ARIA is most likely to occur – namely during dose titration to 10 mg/kg and the first 2 doses of 10 mg/kg. Therefore, routine MRI surveillance is proposed for all patients during this time period. If ARIA is detected in a routine MRI, additional follow-up MRIs might be appropriate to further inform dosing decisions. Following this initial dosing period, it is suggested that that long-term routine MRI surveillance is not needed and that ad hoc MRI testing may be appropriate in the setting of potential symptoms of ARIA. Lastly, given that

patients and prescribers may not be familiar with ARIA, ARIA education is an important component of the proposed ARIA risk mitigation strategy.

The FDA's Position:

FDA agrees with the characterization of the safety profile for ARIA noted in this section, and notes that the incidence of onset of first ARIA primarily occurs between weeks 12 and 32 (and between weeks 12 and 44 in ApoE ϵ 4 carriers).

If efficacy is demonstrated, then appropriate labeling language will address adverse reactions of concern. We agree that prescriber and patient education regarding ARIA is important. The appropriate methods to communicate and educate healthcare providers and patients about the risk of ARIA are still under consideration.

5. POINTS FOR THE ADVISORY COMMITTEE TO CONSIDER

The FDA's Position:

The FDA seeks input from the committee on whether aducanumab is an effective treatment for Alzheimer's disease.

6. REFERENCES

Abu-Rumeileh S, Capellari S, Parchi P. Rapidly Progressive Alzheimer's Disease: Contributions to Clinical-Pathological Definition and Diagnosis. *J Alzheimers Dis.* 2018;63(3):887-897.

Albert MS, DeKosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 2011;7(3):270-9.

Alzheimer's Association. 2020 Alzheimer's disease facts and figures. *Alzheimers Dement.* 2020;16(3):391-485.

Amyvid (Florbetapir F 18 Injection) for intravenous use. Indianapolis, IN: Eli Lilly and Company.

Bekris LM, Yu CE, Bird TD, et al. Genetics of Alzheimer disease. *J Geriatr Psychiatry Neurol.* 2010;23(4):213-27.

Benilova I, Karran E, De Strooper B. The toxic A β oligomer and Alzheimer's disease: an emperor in need of clothes. *Nat Neurosci.* 2012;15(3):349-57. Epub 2012/01/29.

Birks J. Cholinesterase inhibitors for Alzheimer's disease. *Cochrane Database Syst Rev.* 2006(1):CD005593.

Brodaty H, Seeher K, Gibson L. Dementia time to death: a systematic literature review on survival time and years of life lost in people with dementia. *Int Psychogeriatr.* 2012;24(7):1034-45.

Bullock R, Touchon J, Bergman H, et al. Rivastigmine and donepezil treatment in moderate to moderately-severe Alzheimer's disease over a 2-year period. *Curr Med Res Opin.* 2005;21(8):1317-27.

Cedarbaum JM, Jaros M, Hernandez C, et al. Rationale for use of the Clinical Dementia Rating Sum of Boxes as a primary outcome measure for Alzheimer's disease clinical trials. *Alzheimers Dement.* 2013;9(1 Suppl):S45-55. Epub 2012 Jun 01.

Chapman KR, Bing-Canar H, Alosco ML, et al. Mini Mental State Examination and Logical Memory scores for entry into Alzheimer's disease trials. *Alzheimers Res Ther.* 2016;8:9. Epub 2016/02/22.

Coric V, van Dyck CH, Salloway S, et al. Safety and tolerability of the γ -secretase inhibitor avagacestat in a phase 2 study of mild to moderate Alzheimer disease. *Arch Neurol.* 2012;69(11):1430-40.

Cummings J. The Neuropsychiatric Inventory: Development and Applications. *J Geriatr Psychiatry Neurol.* 2020;33(2):73-84.

Delor I, Charoin JE, Gieschke R, et al. Modeling Alzheimer's Disease Progression Using Disease Onset Time and Disease Trajectory Concepts Applied to CDR-SOB Scores From ADNI. *CPT Pharmacometrics Syst Pharmacol*. 2013;2:e78. Epub 2013/10/02.

DeMattos RB, Bales KR, Cummins DJ, et al. Peripheral anti-A beta antibody alters CNS and plasma A beta clearance and decreases brain A beta burden in a mouse model of Alzheimer's disease. *Proc Natl Acad Sci U S A*. 2001;98(15):8850-5. Epub 2001/07/03.

Deng Q, Zhang YY, Roy D, et al. Superiority of combining two independent trials in interim futility analysis. *Stat Methods Med Res*. 2020;29(2):522-540. Epub 2019/04/08.

Doody RS, Thomas RG, Farlow M, et al. Phase 3 trials of solanezumab for mild-to-moderate Alzheimer's disease. *The New England journal of medicine*. 2014;370(4):311-21; correspondence *N Engl J Med*. 2014;370(15):1459-60.

Dubois B, Feldman HH, Jacova C, et al. Revising the definition of Alzheimer's disease: a new lexicon. *Lancet Neurol*. 2010;9(11):1118-27.

Dubois B, Padovani A, Scheltens P, et al. Timely Diagnosis for Alzheimer's Disease: A Literature Review on Benefits and Challenges. *J Alzheimers Dis*. 2016;49(3):617-31.

Dunbar SB, Khavjou OA, Bakas T, et al. Projected Costs of Informal Caregiving for Cardiovascular Disease: 2015 to 2035: A Policy Statement From the American Heart Association. *Circulation*. 2018;137(19):e558-e577. Epub 2018/04/09.

Egan MF, Kost J, Voss T, et al. Randomized Trial of Verubecestat for Prodromal Alzheimer's Disease. *The New England journal of medicine*. 2019;380(15):1408-1420.

Elias-Sonnenschein LS, Viechtbauer W, Ramakers IH, et al. Predictive value of APOE-ε4 allele for progression from MCI to AD-type dementia: a meta-analysis. *J Neurol Neurosurg Psychiatry*. 2011;82(10):1149-56. Epub 2011/04/14.

Evans S, McRae-McKee K, Wong MM, et al. The importance of endpoint selection: How effective does a drug need to be for success in a clinical trial of a possible Alzheimer's disease treatment? *Eur J Epidemiol*. 2018;33(7):635-644. Epub 2018/03/23.

Food and Drug Administration. Guidance for Industry - Early Alzheimer's Disease: Developing Drugs for Treatment Draft Guidance. 2018.

Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12(3):189-98.

Franco-Marina F, García-González JJ, Wagner-Echeagaray F, et al. The Mini-mental State Examination revisited: ceiling and floor effects after score adjustment for educational level in an aging Mexican population. *Int Psychogeriatr*. 2010;22(1):72-81. Epub 2009/09/07.

Galasko D, Abramson I, Corey-Bloom J, et al. Repeated exposure to the Mini-Mental State Examination and the Information-Memory-Concentration Test results in a practice effect in Alzheimer's disease. *Neurology*. 1993;43(8):1559-63.

Greenberg SM, Bacsikai BJ, Hernandez-Guillamon M, et al. Cerebral amyloid angiopathy and Alzheimer disease - one peptide, two pathways. *Nat Rev Neurol*. 2020;16(1):30-42. Epub 2019/12/11.

Guehne U, Luck T, Busse A, et al. Mortality in individuals with mild cognitive impairment. Results of the Leipzig Longitudinal Study of the Aged (LEILA75+). *Neuroepidemiology*. 2007;29(3-4):226-34.

Guehne U, Riedel-Heller S, Angermeyer MC. Mortality in dementia. *Neuroepidemiology*. 2005;25(3):153-62.

Hanseuw BJ, Betensky RA, Jacobs HIL, et al. Association of Amyloid and Tau With Cognition in Preclinical Alzheimer Disease: A Longitudinal Study. *JAMA Neurol*. 2019 Epub 2019/06/03.

Hardy J, Selkoe DJ. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science*. 2002;297(5580):353-6. Epub 2002 Jul 20.

Huang L, Su X, Federoff HJ. Single-chain fragment variable passive immunotherapies for neurodegenerative diseases. *Int J Mol Sci*. 2013;14(9):19109-27. Epub 2013/09/17.

Hyman BT, Phelps CH, Beach TG, et al. National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. *Alzheimers Dement*. 2012;8(1):1-13.

ICH. The extent of population exposure to assess clinical safety for drugs intended for long term treatment of non-life-threatening conditions. Presented at the International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use; 27 October 1994.

Ising C, Stanley M, Holtzman DM. Current thinking on the mechanistic basis of Alzheimer's and implications for drug development. *Clin Pharmacol Ther*. 2015;98(5):469-71.

Jack CR, Bennett DA, Blennow K, et al. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement*. 2018;14(4):535-562.

Jonsson T, Atwal JK, Steinberg S, et al. A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. *Nature*. 2012;488(7409):96-9.

Kang J, Lemaire HG, Unterbeck A, et al. The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface receptor. *Nature*. 1987;325(6106):733-6.

Karran E, Hardy J. A critique of the drug discovery and phase 3 clinical programs targeting the amyloid hypothesis for Alzheimer disease. *Ann Neurol*. 2014;76(2):185-205. Epub 2014/07/02.

Ketter N, Brashear HR, Bogert J, et al. Central Review of Amyloid-Related Imaging Abnormalities in Two Phase III Clinical Trials of Bapineuzumab in Mild-To-Moderate Alzheimer's Disease Patients. *J Alzheimers Dis.* 2017;57(2):557-573.

Klein G, Delmar P, Hofmann C, et al. Higher Dose Gantenerumab Leads to Significant Reduction in Amyloid Plaque Burden: Results for the Marguerite and Scarlet RoAD Open-Label Extension Studies. Presented at the AAN 2018; Los Angeles, CA.

Koffie RM, Meyer-Luehmann M, Hashimoto T, et al. Oligomeric amyloid beta associates with postsynaptic densities and correlates with excitatory synapse loss near senile plaques. *Proc Natl Acad Sci U S A.* 2009;106(10):4012-7.

Linse S, Scheidt T, Bernfur K, et al. Kinetic fingerprints differentiate the mechanisms of action of anti-A β antibodies. *Nature Structural & Molecular Biology.* 2020.

Llorens F, Schmitz M, Karch A, et al. Comparative analysis of cerebrospinal fluid biomarkers in the differential diagnosis of neurodegenerative dementia. *Alzheimers Dement.* 2016;12(5):577-89. Epub 2015/12/21.

Mattsson N, Groot C, Jansen WJ, et al. Prevalence of the apolipoprotein E ϵ 4 allele in amyloid β positive subjects across the spectrum of Alzheimer's disease. *Alzheimers Dement.* 2018;14(7):913-924. Epub 2018/03/28.

McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 2011;7(3):263-9.

McShane R, Areosa Sastre A, Minakaran N. Memantine for dementia. *Cochrane Database Syst Rev.* 2006(2):1-80.

Mohs RC, Knopman D, Petersen RC, et al. Development of cognitive instruments for use in clinical trials of antidementia drugs: additions to the Alzheimer's Disease Assessment Scale that broaden its scope. The Alzheimer's Disease Cooperative Study. *Alzheimer disease and associated disorders.* 1997;11 Suppl 2:S13-21.

Musiek ES, Holtzman DM. Three dimensions of the amyloid hypothesis: time, space and 'wingmen'. *Nat Neurosci.* 2015;18(6):800-6.

NEURACEQ (florbetaben F 18 injection), for intravenous use. Matran, Switzerland: Piramal Imaging.

Ostrowitzki S, Deptula D, Thurfjell L, et al. Mechanism of amyloid removal in patients with Alzheimer disease treated with gantenerumab. *Arch Neurol.* 2012;69(2):198-207. Epub 2011 Oct 10.

Ostrowitzki S, Lasser RA, Dorflinger E, et al. A phase III randomized trial of gantenerumab in prodromal Alzheimer's disease. *Alzheimers Res Ther.* 2017;9(1):95. Epub 2017/12/08.

- Penninkilampi R, Brothers HM, Eslick GD. Safety and Efficacy of Anti-Amyloid- β Immunotherapy in Alzheimer's Disease: A Systematic Review and Meta-Analysis. *J Neuroimmune Pharmacol*. 2017;12(1):194-203. Epub 2016/12/26.
- Rizzuto D, Bellocco R, Kivipelto M, et al. Dementia after age 75: survival in different severity stages and years of life lost. *Curr Alzheimer Res*. 2012;9(7):795-800.
- Roberts RO, Aakre JA, Kremers WK, et al. Prevalence and Outcomes of Amyloid Positivity Among Persons Without Dementia in a Longitudinal, Population-Based Setting. *JAMA Neurol*. 2018;75(8):970-979.
- Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. *Am J Psychiatry*. 1984;141(11):1356-64.
- Rozkalne A, Spires-Jones TL, Stern EA, et al. A single dose of passive immunotherapy has extended benefits on synapses and neurites in an Alzheimer's disease mouse model. *Brain Res*. 2009;1280:178-85.
- Salloway S, Sperling R, Fox NC, et al. Two phase 3 trials of bapineuzumab in mild-to-moderate Alzheimer's disease. *The New England journal of medicine*. 2014;370(4):322-33; accompanied by supplementary appendix.
- Salloway S, Sperling R, Gregg K, et al. Characterization of the Clinical Course of Placebo-Treated Amyloid-Negative Subjects With Mild-Moderate Alzheimer's Disease: Results From the Phase 3 PET Substudies of Bapineuzumab and Solanezumab. Presented at the Alzheimer's Association International Conference 2013; Boston, MA.
- Schmidt C, Haïk S, Satoh K, et al. Rapidly progressive Alzheimer's disease: a multicenter update. *J Alzheimers Dis*. 2012;30(4):751-6.
- Selkoe DJ. The molecular pathology of Alzheimer's disease. *Neuron*. 1991;6(4):487-98.
- Selkoe DJ. The therapeutics of Alzheimer's disease: where we stand and where we are heading. *Ann Neurol*. 2013;74(3):328-36.
- Selkoe DJ, Hardy J. The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Mol Med*. 2016;8(6):595-608. Epub 2016/06/01.
- Seubert P, Barbour R, Khan K, et al. Antibody capture of soluble Abeta does not reduce cortical Abeta amyloidosis in the PDAPP mouse. *Neurodegener Dis*. 2008;5(2):65-71. Epub 2008/01/04.
- Sevigny J, Chiao P, Bussière T, et al. The antibody aducanumab reduces A β plaques in Alzheimer's disease. *Nature*. 2016;537(7618):50-6.
- Spencer RJ, Wendell CR, Giggey PP, et al. Psychometric Limitations of the Mini-Mental State Examination Among Nondemented Older Adults: An Evaluation of Neurocognitive and Magnetic Resonance Imaging Correlates. *Experimental Aging Research*. 2013;39(4):382-397.

Sperling R, Salloway S, Brooks DJ, et al. Amyloid-related imaging abnormalities in patients with Alzheimer's disease treated with bapineuzumab: a retrospective analysis. *Lancet Neurol*. 2012;11(3):241-9.

Sperling RA, Aisen PS, Beckett LA, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011a;7(3):280-92.

Sperling RA, Jack Jr CR, Black SE, et al. Amyloid-related imaging abnormalities in amyloid-modifying therapeutic trials: Recommendations from the Alzheimer's Association Research Roundtable Workgroup. *Alzheimers Dement*. 2011b;7(4):367-385.

Spires-Jones TL, Mielke ML, Rozkalne A, et al. Passive immunotherapy rapidly increases structural plasticity in a mouse model of Alzheimer disease. *Neurobiol Dis*. 2009;33(2):213-20.

Suehs BT, Shah SN, Davis CD, et al. Household members of persons with Alzheimer's disease: health conditions, healthcare resource use, and healthcare costs. *J Am Geriatr Soc*. 2014;62(3):435-41. Epub 2014/02/27.

Suzuki K, Hirakawa A, Ihara R, et al. Effect of apolipoprotein E ϵ 4 allele on the progression of cognitive decline in the early stage of Alzheimer's disease. *Alzheimers Dement (N Y)*. 2020;6(1):e12007. Epub 2020/03/20.

Swanson CJ, Zhang Y, Dhadda S, et al. Treatment of Early AD Subjects with BAN2401, an Anti-AB Protofibril Monoclonal Antibody, Significantly Clears Amyloid Plaque and Significantly Reduces Clinical Decline. Presented at the AAIC 2018; Los Angeles, CA.

Tariot PN, Solomon PR, Morris JC, et al. A 5-month, randomized, placebo-controlled trial of galantamine in AD. The Galantamine USA-10 Study Group. *Neurology*. 2000;54(12):2269-76.

Tukey JW. Fences and outside values. *Exploratory data analysis*. Reading, MA: AddisonWesley; 1977. p 43.

U.S. Department of Health and Human Services. National Plan to address Alzheimer's Disease: 2019 update.

Vermunt L, Sikkes SAM, van den Hout A, et al. Duration of preclinical, prodromal, and dementia stages of Alzheimer's disease in relation to age, sex, and APOE genotype. *Alzheimers Dement*. 2019;15(7):888-898. Epub 2019/06/01.

Villemagne VL, Burnham S, Bourgeat P, et al. Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. *Lancet Neurol*. 2013;12(4):357-67.

VIZAMYL (flutemetamol F 18 injection) for intravenous use. Arlington Heights, IL: GE Healthcare.

Weiner MW, Veitch DP, Aisen PS, et al. Recent publications from the Alzheimer's Disease Neuroimaging Initiative: Reviewing progress toward improved AD clinical trials. *Alzheimers Dement*. 2017;13(4):e1-e85. Epub 2017/03/22.

Xie J, Brayne C, Matthews FE, et al. Survival times in people with dementia: analysis from population based cohort study with 14 year follow-up. *BMJ*. 2008;336(7638):258-62.

Zago W, Schroeter S, Guido T, et al. Vascular alterations in PDAPP mice after anti-A β immunotherapy: Implications for amyloid-related imaging abnormalities. *Alzheimers Dement*. 2013;9(5 Suppl):S105-15.

7. Appendix 1

FDA Clinical Review
Kevin Krudys, PhD
BLA 761178
Aduhelm (aducanumab)

FDA CLINICAL REVIEW

Application Type	BLA
Application Number(s)	761178
Priority or Standard	Priority
Submit Date(s)	07/07/2020
Received Date(s)	07/07/2020
PDUFA Goal Date	03/07/2021
Division/Office	Division of Neurology 1/Office of Neuroscience
Reviewer Name(s)	Kevin Krudys, PhD
Established/Proper Name	Aducanumab
(Proposed) Trade Name	Aduhelm
Applicant	Biogen Inc.
Dosage Form(s)	Solution for injection
Applicant Proposed Dosing Regimen(s)	10 mg/kg as an intravenous infusion every four weeks
Applicant Proposed Indication(s)/Population(s)	To delay clinical decline in patients with Alzheimer's disease

Table of Contents

1. Executive Summary	7
1.1. Product Introduction	7
1.2. Conclusions on Substantial Evidence of Effectiveness.....	7
2. Therapeutic Context	9
2.1. Analysis of Condition	9
2.2. Analysis of Current Treatment Options	11
3. Regulatory Background	11
3.1. U.S. Regulatory Actions and Marketing History	11
3.2. Summary of Presubmission/Submission Regulatory Activity	11
3.3. Foreign Regulatory Actions and Marketing History	13
4. Sources of Clinical Data and Review Strategy	13
4.1. Table of Clinical Studies.....	13
4.2. Review Strategy.....	16
5. Review of Relevant Individual Trials Used to Support Efficacy	17
5.1. Study 302 (221AD302): A Phase 3 Multicenter, Double-Blind, Placebo-Controlled, Parallel-Group Study to Evaluate the Efficacy and Safety of Aducanumab (BIB037) in Subjects With Early Alzheimer’s Disease.....	17
5.1.1. Study Design	17
5.1.2. Study Results	31
5.2. Study 301 (221AD301): A Phase 3 Multicenter, Double-Blind, Placebo-Controlled, Parallel-Group Study to Evaluate the Efficacy and Safety of Aducanumab (BIB037) in Subjects With Early Alzheimer’s Disease.....	44
5.2.1. Study Design	44
5.2.2. Study Results	44
5.3. Study 103 (221AD103): A Randomized, Double-Blinded, Placebo-Controlled Multiple Dose Study to Assess the Safety, Tolerability, Pharmacokinetics, and Pharmacodynamics of BIB037 in Subjects with Prodromal or Mild Alzheimer’s Disease.....	56
5.3.1. Study Design	56
5.3.2. Study Results.....	67

6. Integrated Review of Effectiveness	75
6.1. Assessment of Efficacy Across Trials	75
6.1.1. Primary Endpoints.....	75
6.1.2. Secondary and Other Endpoints	77
6.1.3. Subpopulations	80
6.1.4. Dose and Dose-Response.....	81
6.1.5. Onset, Duration, and Durability of Efficacy Effects	81
6.1.6. Early Termination and Interpretation of Study 301 and Study 302	83
6.1.7. Analysis of Discordant Results in Study 301 and Study 302	86
6.2. Integrated Assessment of Effectiveness	103

Table of Tables

Table 1: Clinical Studies Contributing Efficacy Data and Relevant to the Review of this BLA.....	14
Table 2: Dosing Scheme for Aducanumab by Treatment Group and ApoE ε4 Carrier Status.....	20
Table 3: Dose Modification/Discontinuation Rules for ARIA-E by Protocol Version (Study 302)	21
Table 4: Dose Modification/Discontinuation Rules for ARIA-H Microhemorrhage and Superficial Siderosis by Protocol Version (Study 302)	22
Table 5: Study 302 Schedule of Key Assessments	23
Table 6: Study 302 Patient Disposition	32
Table 7: Study 302 Baseline Demographics (ITT Population)	33
Table 8: Study 302 Disease Characteristics (ITT Population).....	35
Table 9: Study 302 Primary Endpoint Analysis	37
Table 10: Study 302 Primary Endpoint Analysis (Dataset Sensitivity)	37
Table 11: Study 302 Secondary Endpoint Analysis	40
Table 12: Study 301 Patient Disposition	44
Table 13: Study 301 Baseline Demographics (ITT Population)	47
Table 14: Study 301 Baseline Disease Characteristics (ITT Population)	48
Table 15: Study 301 Primary Endpoint Analysis	50
Table 16: Study 301 Primary Endpoint Analysis (Dataset Sensitivity)	51
Table 17: Study 301 Secondary Endpoint Analysis	51
Table 18: Dose Modification/Discontinuation Rules for ARIA-E by Protocol Version (Study 103)	60
Table 19: Dose Modification/Discontinuation Rules for ARIA-H (Microhemorrhage) by Protocol Version (Study 103).....	61
Table 20: Dose Modification/Discontinuation Rules for ARIA-H (Superficial Siderosis) by Protocol Version (Study 103).....	62
Table 21: Study 103 Schedule of Key Assessments	63
Table 22: Study 103 Patient Disposition	68
Table 23: Study 103 Baseline Demographics (ITT).....	70
Table 24: Study 103 Disease Characteristics (ITT Population).....	71
Table 25: Study 103 Clinical Endpoints (CDR-SB and MMSE) Analyses by ANCOVA	72
Table 26: Study 103 Clinical Endpoints (CDR-SB and MMSE) Analyses by MMRM.....	73
Table 27: Summary of Findings for CDR-SB in Studies 103, 301 and 302	77
Table 28: Summary of Findings for MMSE in Studies 103, 301 and 302	77
Table 29: Summary of Findings for ADAS-Cog 13 in Studies 301 and 302	78
Table 30: Summary of Findings for ADCS-ADL-MCI in Studies 301 and 302	78
Table 31: Summary of Findings for Amyloid PET Composite SUVR in Studies 103, 301 and 302	79
Table 32: Conditional Power for Change from Baseline in CDR-SB, MMSE and ADAS-Cog13 at Week 78 (Interim Futility Dataset)	84
Table 33: Conditional Power (Non-Pooled Analysis) for Change from Baseline in CDR-SB, MMSE and ADAS-Cog13 at Week 78 (Interim Futility Dataset)	85

Table 34: Number of Patients with Rapid Progression (CDR-SB Change from Baseline of >8 at Week 78)	89
Table 35: Study 301 Change from Baseline in CDR-SB, MMSE and ADAS-Cog13 at Week 78 Excluding Rapid Progressors (CDR-SB Change from Baseline > 8).....	89
Table 36: Study 302 Change from Baseline in CDR-SB, MMSE and ADAS-Cog13 at Week 78 Excluding Rapid Progressors (CDR-SB Change from Baseline > 8).....	89
Table 37: Change from Baseline in CDR-SB, MMSE, ADAS-Cog 13 and ADCS-ADL-MCI at Week 78: All Observations vs. Exclusion of Observations Post-ARIA	92
Table 38: Dosing in Subgroups Defined by Study, Randomized Dose Group, ApoE ε4 Status and Pre- and Post-Protocol Version 4.....	96
Table 39: Treatment Difference in CDR-SB at Week 78 by Cumulative Dose Threshold for High Dose in Study 301	98

Table of Figures

Figure 1: Study 302 High Dose Subgroup Analysis of the Primary Efficacy Endpoint (Demographics)	38
Figure 2: Study 302 High Dose Subgroup Analysis of the Primary Efficacy Endpoint (Baseline Characteristics)	39
Figure 3: Study 302 Change from Baseline in A β PET Composite SUVR.....	41
Figure 4: Study 302 Change from Baseline in CSF A β_{1-42} (left) and A β_{1-40} (right) at Week 78.....	42
Figure 5: Study 302 Change from Baseline in CSF p-Tau (left) and t-Tau (right) at Week 78.....	43
Figure 6: Study 301 Change from Baseline in A β PET Composite SUVR.....	53
Figure 7: Study 301 Change from Baseline in CSF A β_{1-42} (left) and A β_{1-40} (right) at Week 78.....	54
Figure 8: Study 301 Change from Baseline in CSF p-Tau (left) and t-Tau (right) at Week 78.....	54
Figure 9: Study 301 and Study 302 Pooled Analyses: Change from Baseline in Tau PET	55
Figure 10: Study 103 Schematic.....	57
Figure 11: Study 103 Change from Baseline in A β PET Composite SUVR.....	75
Figure 12: Group-Level Correlation Between Adjusted Mean Difference from Placebo in A β PET Composite SUVR and CDR-SB	80
Figure 13: Study 302 Clinical Endpoints Adjusted Mean Change from Baseline over Time	81
Figure 14: Typical Profile of SUVR for Planned Titration and No Dose Interruptions	83
Figure 15: Distribution of Change from Baseline in CDR-SB at Week 78 in Pooled Treatment Arms in Studies 301 and 302	88
Figure 16: Treatment Difference of Change from Baseline in CDR-SB, MMSE, ADAS-Cog 13 and ADCS-ADL-MCI at Week 78 Grouped by Study, Dose and ApoE Carrier Status and Including or Excluding Post-ARIA Observations. (x - Study 302, o – Study 301, C – carrier, NC – con-carrier; solid line is line of unity)	93
Figure 17: Treatment Difference in CDR-SB at Week 78 Vs. Mean Cumulative Dose by Study, Randomized Dose Group, ApoE ϵ 4 Carrier Status and Protocol Version	97
Figure 18: CDR-SB Adjusted Mean Change from Baseline % Difference from Propensity-Matched Placebo at Week 78 in Subgroups by Cumulative Dose in Studies 301 and 302	99
Figure 19: CDR-SB Adjusted Mean Change from Baseline % Difference from Propensity-Matched Placebo at Week 78 in Subgroups by Cumulative Dose in Studies 301 and 302 (Rapid Progressors Removed)	100
Figure 20: CDR-SB Adjusted Mean Change from Baseline % Difference from Propensity-Matched Placebo at Week 78 in Subgroups by Number of 10 mg/kg Doses in Studies 301 and 302.....	101
Figure 21: CDR-SB Adjusted Mean Change from Baseline % Difference from Propensity-Matched Placebo at Week 78 in Subgroups by Number of 10 mg/kg Doses in Studies 301 and 302 (Rapid Progressors Removed)	101

1. Executive Summary

This document has been prepared to provide the Advisory Committee with the clinical review of the information the applicant has provided supporting the effectiveness of aducanumab for the treatment of Alzheimer's disease.

1.1. Product Introduction

Aducanumab (previously BIIB037) is a human immunoglobulin gamma 1 (IgG1) anti-amyloid beta (A β) monoclonal antibody targeting aggregated forms of A β . Extracellular deposits of A β are one of the two pathological hallmarks of Alzheimer's disease, along with intracellular aggregates of hyperphosphorylated tau in the form of neurofibrillary tangles. Accumulation of A β in the brain has been proposed to be the primary driver of the disease process. Aducanumab reduces levels of brain A β plaque by targeting aggregated forms of A β including soluble oligomers and insoluble fibrils.

The applicant's proposed indication is to delay clinical decline in patients with Alzheimer's disease. The dosing regimen consists of an intravenous infusion over approximately one hour every four weeks. To initiate treatment, patients should receive two doses of 1 mg/kg, two doses of 3 mg/kg, then two doses of 6 mg/kg over a total of 24 weeks. Thereafter, the target maintenance dose is 10 mg/kg.

Aducanumab is a new molecular entity (NME) and is not marketed in any country. The proposed proprietary name is Aduhelm.

1.2. Conclusions on Substantial Evidence of Effectiveness

Evaluation of the evidence intended to establish the effectiveness of aducanumab was originally complicated by the fact that the two large, international pivotal trials (Studies 301 and 302) were near completion but were terminated prior to their planned conclusion with a public declaration of futility. Therefore, a critical and necessary first step in the consideration of these data was to assess the interpretability of the observed efficacy data considering the termination of the studies. Modeling and simulation methodologies were used to "virtually complete" the studies. The results of these analyses established that the results of Studies 301 and 302 were interpretable and suitable for additional consideration. Agreement on the interpretability of the data and the final dataset for further consideration was documented in the October 2019 Type C Meeting Minutes.

The applicant presents the results of Study 302 as the primary evidence of effectiveness. Study 302 is a large, multicenter trial that demonstrated aducanumab, as compared to placebo, reduced the change from baseline on the primary outcome measure, CDR-SB (-0.39, p=0.012).

The CDR-SB is an acceptable primary outcome measure and its use was accepted by FDA in a Special Protocol Assessment of Study 302 in September 2015. The effect of aducanumab in Study 302 was robust, including convincing effects on the primary endpoint, all three multiplicity-controlled secondary endpoints, and the tertiary clinical endpoint, and was exceptionally persuasive on several of the instruments used to evaluate efficacy. To some degree, each of these endpoints captures distinct information regarding cognitive decline. Statistically significant effects on ADAS-Cog 13 (-1.400, $p=0.0097$) and ADCS-ADL-MCI (1.7, 0.0006), endpoints which independently assess cognition and daily function, represent another acceptable co-primary endpoint approach in studies of Alzheimer's disease. Statistically significant treatment effects were maintained in sensitivity analyses, including one which included only patients who completed the study before its early termination. Further support for the results of the high dose is provided by the numerically favorable results for the low dose on the primary endpoint (-0.26, $p=0.0901$). Numerically favorable trends in the low dose were also observed for two of the secondary endpoints. These results are also suggestive of a dose-response relationship for aducanumab.

The treatment effect in Study 302 is supported by consistently favorable results for the primary and secondary endpoints across 79 of 80 prespecified subgroups of interest defined by demographic and baseline disease characteristics. Brain A β measured by PET was significantly reduced in a dose-dependent manner in all prespecified subgroups. Biomarkers reflecting target engagement (brain A β reduction and CSF A β_{1-42} levels), effects on downstream Alzheimer's tau pathophysiology (p-Tau and Tau PET), and neurodegeneration (t-Tau) supported the observations on the clinical outcomes. Dose- and concentration-dependent relationships for biomarkers offer further support to the apparent dose-response relationship for clinical endpoints.

Although designed primarily as a safety and tolerability study, Study 103 was an adequate and well-controlled study that explicitly included assessment of prespecified clinical (CDR-SB and MMSE) and biomarker endpoints and a prospectively identified statistical analysis plan. The results of Study 103 are appropriately viewed as supportive evidence of effectiveness. Despite the limitations of a trial designed primarily to assess safety and tolerability, the 10 mg/kg dose arm was able to show a reduction, as compared to placebo, in the change from baseline in the clinical endpoints, CDR-SB (-1.08, $p=0.0464$) and MMSE (1.9, $p=0.0356$). The dose-response relationship for A β reduction provides support for the positive finding and is consistent with the dose-response relationship observed in the study for CDR-SB and MMSE.

Study 301 is a negative study and does not contribute to the evidence of effectiveness. One key observation, however, is that primary and secondary endpoints for the low dose had responses that were numerically favorable and similar in magnitude to those in Study 302. The response of the high dose was the notable difference between the studies. At the June 14, 2019, Type C Meeting, the Division clearly stated that, "available data do not suggest the future use of Study

301 as an efficacy study providing independent evidence of effectiveness supporting the approval of aducanumab.” Rather, the Division noted the possibility that analyses “may be understood well enough... to not represent evidence that the drug is ineffective.” Given the possibility that aducanumab is an effective drug for a disease with an enormous unmet medical need, extensive resources were brought to bear on achieving a maximum understanding of the existing data. The analyses were exploratory by design but limited in scope and focused on pre-defined areas of interest. These explorations were not intended to provide supportive evidence of effectiveness nor were they intended to fully account for the partially discordant results in Studies 301 and 302. The detailed considerations given to those analyses in this review, however, attest to the relevance of these investigations. The rapid progressor analysis indicated that a small imbalance in the number of rapid progressing patients in the high-dose arm in Study 301 had a disproportionate impact on the estimate of the treatment effect using the primary analysis method. An examination of dosing in Study 301 indicates that patients with higher sustained exposure to the 10 mg/kg dose in Study 301 had similar responses to patients in Study 302. These two factors contribute to the overall understanding of Study 301.

Based on the considerations above, the applicant has provided substantial evidence of effectiveness to support approval. Study 302 provides the primary evidence of effectiveness as a robust and exceptionally persuasive study demonstrating a treatment effect on a clinically meaningful endpoint and reinforced by effects on secondary endpoints, biomarkers, and in relevant subgroups. Study 103 was an adequate and well-controlled study which included design components consistent with Study 302 and demonstrated a persuasive treatment effect on both clinical endpoints. The dose-response relationship for A β reduction provides support for the positive finding in the 10 mg/kg treatment arm and to the apparently dose-related effects observed on clinical outcomes in Studies 103 and 302. Study 301 does not contribute to the evidence of effectiveness. The results of exploratory analyses, however, contribute to the overall understanding of Study 301 and together do not meaningfully detract from the persuasiveness of Study 302.

2. Therapeutic Context

2.1. Analysis of Condition

Alzheimer’s disease is a progressive, degenerative brain disorder that affects memory, thinking and behavior and is the most common cause of dementia. An estimated 5.8 million Americans age 65 and older are currently living with Alzheimer’s disease. In the absence of interventions to prevent or slow the disease the number is projected to reach 13.8 million by 2050.

Alzheimer’s disease is the sixth-leading cause of death in the United States and the fifth-leading cause of death for those age 65 and older. Almost two-thirds of Americans with Alzheimer’s disease are women. Older African Americans and Latinos are disproportionately more likely to

have Alzheimer's disease than White Americans.

Alzheimer's disease exists on a continuum from pathophysiological changes in the brain which are undetectable to the person affected, to subtle problems with memory and thinking, and ultimately difficulties with memory, language, problem-solving and other skills that affect an individual's ability to perform everyday activities. The disease process may begin 20 years or more before symptoms arise. Life expectancy varies depending on many factors, but after a diagnosis of Alzheimer's dementia the average survival is 4 to 8 years. The long duration of the disease contributes to the burden not only of the individuals with the disease, but also their families and caregivers who provide most of the patient care and are at an increased risk for emotional distress and negative mental and physical outcomes.

The two pathological hallmarks of Alzheimer's disease are extracellular deposits of A β , or plaques, and intracellular aggregates of hyperphosphorylated tau in the form of neurofibrillary tangles. Accumulation of A β in the brain has been proposed to be the primary driver of the disease process and precedes the accumulation of tau pathology and neural degeneration. Consequently, therapies to inhibit A β production or enhance A β clearance have been investigated in an attempt to slow or halt the disease process. Anti-amyloid therapies have thus far had a disappointing record with many high-profile, late-stage failures. Critical evaluation of these programs has revealed many factors that may, in part, contribute to their lack of success, including insufficient dosing, unknown target engagement, off-target effects, and inclusion of individuals in trials without evidence of brain A β pathology or at later stages of Alzheimer's disease. Importantly, "anti-amyloid" therapies are not a distinct class of drugs, but rather reflect many different modes of action. Even among anti-A β monoclonal antibodies there are differences due to effector function, binding at different epitopes, and selectivity for A β variants (e.g. monomers, soluble oligomers, aggregated forms). Therefore, previous late-stage failures of anti-A β therapies do not constitute a demonstrated "class failure" and are not particularly informative for the assessment of the effectiveness of aducanumab.

The aducanumab development program in many ways stands apart from these previous failures. The pivotal trials of aducanumab included patients with evidence of brain A β pathology who were early in the disease process. Moreover, an early trial (Study 103) demonstrated target engagement and confirmed reduction of A β plaque burden. Accordingly, aducanumab may more appropriately be grouped with agents which have also demonstrated plaque reduction at appropriate dosages with some early evidence suggesting favorable effects on clinical endpoints.

Some anti-A β monoclonal antibodies, including aducanumab, have been associated with the occurrence of amyloid related imaging abnormalities (ARIA) that require special attention with respect to dosing and monitoring. ARIA covers a spectrum of imaging findings detected on brain magnetic resonance imaging (MRI) which include ARIA-edema (ARIA-E) and ARIA-hemorrhage

(ARIA-H).

2.2. Analysis of Current Treatment Options

Current treatment goals for patients with Alzheimer's disease are often directed to maintain quality of life, treat cognitive symptoms, and manage behavioral and psychological symptoms of dementia. Currently approved Alzheimer's disease treatments include the cholinesterase inhibitors donepezil, rivastigmine, and galantamine, and the N-methyl-D-aspartate antagonist memantine. The most recent approval of a novel medication was for memantine in 2004.

There clearly remains an enormous unmet clinical need for effective treatments for Alzheimer's disease. Current treatments do not target the underlying pathology of Alzheimer's disease and are only able to modestly affect the manifestations of the disease. Furthermore, there are no treatments explicitly approved for earlier stages of the disease.

3. Regulatory Background

3.1. U.S. Regulatory Actions and Marketing History

Aducanumab is a new molecular entity (NME) and is not currently marketed in the United States for any indication.

3.2. Summary of Presubmission/Submission Regulatory Activity

IND 106230 for aducanumab (previously BIIB037) was opened in the United States on April 6, 2011, for the treatment of Alzheimer's disease following a pre-IND meeting between the applicant and FDA on October 13, 2009. Relevant regulatory interactions between FDA and the applicant include the following:

- December 16, 2014 – A Type B End of Phase 2 Meeting was held. The meeting included preliminary discussion regarding study population, endpoints and dosing for the applicant's two proposed Phase 3 studies, and the Division suggested a Special Protocol Assessment (SPA) be requested for an in-depth review of the protocols. The Division and applicant agreed to the safety monitoring plan and the safety database.
- September 28, 2015 – SPA agreements for Study 301 and Study 302 were reached with the Division.
- March 21, 2019 – Biogen announces the termination of the Phase 3 program (Studies

301 and 302) based on results of a pre-specified interim futility analysis. The first formal request from Biogen to the FDA regarding a discussion of the decision to terminate the studies came on May 15, 2019, in the form of a Type C Meeting Request.

- June 14, 2019 – A Type C Meeting was held to discuss the applicant’s analysis of the intent to treat (ITT) populations of Study 301 and Study 302 including all data prior to the March 21, 2019, announcement of the termination of the studies. The Division advised the applicant that the development of aducanumab should not be abandoned as the available clinical data suggest the drug may be clinically active and do not provide convincing evidence that the drug is ineffective. The Division recommended that further analyses of the available data should be conducted to understand the effect of early termination of the studies on the interpretability of the data and to address the partially conflicting results for Study 301 as compared with those for Study 302.
- October 21, 2019 – A Type C Meeting was held to discuss the additional analyses proposed at the June 14, 2019, meeting. Based on these analyses, the Division agreed that the results of Study 301 and Study 302 are interpretable and suitable for additional consideration. The Division further advised that planning for submission of a marketing application was a reasonable option.
- February 20, 2020 – The applicant opened BLA 761178 and submitted nonclinical information.
- February 27, 2020 – Another Type C Meeting was held to discuss scientific questions raised at the October 21, 2019, meeting. The Division agreed to review a protocol for an open-label, uncontrolled re-dosing study in subjects who were enrolled in previous aducanumab studies when they were halted.
- June 17, 2020 – The goal of this Type C meeting held via teleconference was to discuss the efficacy and safety data that were available for aducanumab in the context of the planned BLA submission. The Division informed the applicant that the extent to which the results of Study 301, Study 103, and the additional data analyses support or undermine the results of Study 302 would be a matter of review. The Division and applicant also agreed on the safety data to support filing. On June 8, 2020, the applicant submitted a meeting request for a Type B pre-BLA meeting. The Division determined that the questions for that meeting request were limited in scope and therefore provided answers as part of the June 17, 2020, meeting.
- July 7, 2020 – BLA submission complete.

3.3. Foreign Regulatory Actions and Marketing History

Aducanumab is not approved or marketed in any foreign country.

4. Sources of Clinical Data and Review Strategy

4.1. Table of Clinical Studies

A summary of clinical studies pertinent to the evaluation of efficacy is presented in Table 1.

Table 1: Clinical Studies Contributing Efficacy Data and Relevant to the Review of this BLA

Trial Identity/NCT no.	Trial Design	Regimen/ schedule/ route	Study Endpoints	Treatment Duration/ Follow Up	No. of patients enrolled	Study Population (per categorization at the time of enrollment)	No. of Centers and Countries
<i>Controlled Studies to Support Efficacy and Safety</i>							
221AD302 (Study 302)/ NCT02484547	Randomized, double-blind, placebo-controlled, parallel group	IV infusion once every 4 weeks <u>Placebo</u> saline infusion <u>Low Dose</u> Titration to 3 mg/kg for ApoE ε4 carriers or 6 mg/kg for ApoE ε4 non-carriers <u>High Dose</u> Titration to 10 mg/kg	<u>Primary</u> Change from baseline in CDR-SB at Week 78 <u>Secondary</u> Change from baseline in MMSE, ADAS-Cog 13 and ADCS-ADL-MCI at Week 78	76-week treatment period 18-week follow-up period (or enrollment in LTE up to 5 years)	1643	MCI due to Alzheimer's disease or mild Alzheimer's disease dementia CDR global score of 0.5, MMSE between 24 to 30 and positive amyloid PET scan 50 to 85 years of age	181 centers in 13 countries
221AD301 (Study 301)/ NCT02477800	Randomized, double-blind, placebo-controlled, parallel group	IV infusion once every 4 weeks <u>Placebo</u> saline infusion	<u>Primary</u> Change from baseline in CDR-SB at Week 78	76-week treatment period 18-week	1653	MCI due to Alzheimer's disease or mild Alzheimer's disease	169 centers in 14 countries

		<u>Low Dose</u> Titration to 3 mg/kg for ApoE ε4 carriers or 6 mg/kg for ApoE ε4 non-carriers <u>High Dose</u> Titration to 10 mg/kg	<u>Secondary</u> Change from baseline in MMSE, ADAS-Cog 13 and ADCS-ADL-MCI at Week 78	follow-up period (or enrollment in LTE up to 5 years)		dementia (with CDR global score of 0.5, MMSE of 24 to 30 and positive amyloid PET scan) 50 to 85 years of age	
221AD103 (Study 103)/ NCT01677572	Randomized, double-blind, placebo-controlled, 4-cohort, multiple-dose	IV infusion once every 4 weeks <u>Fixed Dose Cohorts</u> Placebo 1 mg/kg 3 mg/kg 6 mg/kg 10 mg/kg <u>Titration Cohort</u> Placebo 10 mg/kg (44-week titration)	<u>Primary</u> Safety and tolerability <u>Secondary</u> Change from baseline in Aβ PET composite SUVR at Week 54 <u>Exploratory</u> Change from baseline in CDR-SB at Week 54 and MMSE at Week 52	52-week treatment period 18-week follow-up period (or enrollment in LTE)	197	Prodromal Alzheimer's disease (IWG criteria) or mild Alzheimer's dementia CDR global score of 0.5 or 1.0 and MMSE between 20 and 30 and positive amyloid PET scan 50 to 90 years of age	27 centers in the United States

4.2. Review Strategy

Before evaluating the evidence submitted in support of effectiveness, the efficacy review addresses the applicant's declaration of futility and subsequent termination of Study 301 and Study 302. During the June 2019 Type C Meeting, the applicant and the Division agreed that it would have been more appropriate if futility had not been declared. Therefore, a critical first step was to assess the interpretability of the observed efficacy data considering the termination of the studies. Virtual completion of the studies using modeling and simulation was used to explore the range of plausible outcomes had the studies been run to completion. This analysis established the results of Studies 301 and 302 as interpretable and suitable for additional consideration. The applicant and Division agreed that the appropriate dataset should include all data from all randomized and dosed participants with censoring of data collected after the futility announcement (March 20, 2019). A new subheading has been added to the review template (Section 6.1.6) to fully describe the futility analysis and its consequences. Although an unusual situation, the early termination of the studies does not render the findings uninterpretable.

After concluding the studies were interpretable, the applicant had one study which, on face, met its primary and secondary endpoints (Study 302) and appeared exceptionally persuasive and one which did not meet the primary endpoint (Study 301). Both studies are reviewed in Section 5. As the positive study, Study 302 is presented first and in greater detail, with sensitivity and subgroup analyses. Study 301 results are also presented with a high-level focus on potential similarities or differences with Study 302. A more detailed consideration of the partially discordant results between Studies 301 and 302 is presented in Section 6.1.7. The analyses presented in this section are exploratory but were hypothesis-driven, designed to address specific questions, and predetermined to the furthest extent possible. Importantly, the analyses were not intended to obtain independent support from Study 301, nor were they intended to completely explain the discordant results. Study 301 is clearly a negative study. The intention was to provide maximum understanding for why Study 301 had a negative outcome and to determine the degree to which it supports or undermines the persuasive results from Study 302.

The partially discordant results in Studies 301 and 302 elevate the importance of considering the results of Study 103. Although designed primarily as a safety and tolerability study, Study 103 was a randomized, double-blind, multicenter study which incorporated many of the characteristics of Study 301 and Study 302, including blinded efficacy and pharmacodynamic assessments in a similar patient population. This study is therefore reviewed in detail in Section 5.3.

Given the unique nature of the aducanumab data, an integrated assessment of effectiveness is

particularly relevant for this review. Section 1.2 provides an integrated summary of the evidence in support of the effectiveness of aducanumab for the treatment of Alzheimer's disease.

This review focuses solely on clinical efficacy. This application is being reviewed separately for safety by Drs. Branagan and Trummer.

5. Review of Relevant Individual Trials Used to Support Efficacy

5.1. Study 302 (221AD302): A Phase 3 Multicenter, Double-Blind, Placebo-Controlled, Parallel-Group Study to Evaluate the Efficacy and Safety of Aducanumab (BIIB037) in Subjects With Early Alzheimer's Disease

5.1.1. Study Design

Overview and Objective

Study 302 was one of two identically designed studies to evaluate the efficacy and safety of aducanumab in patients with Alzheimer's disease. The primary efficacy objective was to evaluate the efficacy of monthly doses of aducanumab in slowing cognitive and functional impairment as measured by changes in the Clinical Dementia Rating Scale-Sum of Boxes (CDR-SB) score as compared with placebo in participants with mild cognitive impairment (MCI) due to Alzheimer's disease or mild Alzheimer's disease dementia.

Trial Design

Study Design

Study 302 was a multicenter, randomized, double-blind, placebo-controlled, parallel-group study in patients with MCI due to Alzheimer's disease or mild Alzheimer's disease dementia. The study was conducted in 181 centers globally. Randomization was stratified by site and by ApoE ε4 carrier status (carrier or non-carrier) and enrollment was monitored such that 80% of the population included patients with a baseline clinical stage of MCI due to Alzheimer's disease. The study included an 8-week screening period, a 78-week placebo-controlled treatment period and a safety follow-up period of 18 weeks after the final dose. For the placebo-controlled period, patients were randomized to aducanumab low dose, aducanumab high dose, or placebo treatment arms in a 1:1:1 ratio. Patients who completed the placebo-controlled period had the option to enter the 5-year, dose-blind, long-term extension period.

Diagnostic Criteria

At the time of enrollment, patients fulfilled clinical criteria for either MCI due to Alzheimer's disease or mild Alzheimer's disease dementia as defined by the 2011 National Institute on Aging-Alzheimer's Association (NIA-AA) framework. Patients were also required to have evidence of brain A β pathology by visual read of a PET scan.

Key Inclusion Criteria

1. Male or female patients age 50 to ≤ 85 years
2. At least 6 years of education or work experience
3. Positive amyloid positron emission tomography (PET) scan
4. CDR global score of 0.5
5. Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) score ≤ 85
6. Mini-Mental State Examination (MMSE) score ≥ 24
7. Must consent to ApoE genotyping
8. Has one informant/care partner who is able to provide accurate information about the subject's cognitive and functional abilities and should be available for the duration of the study

Key Exclusion Criteria

1. Any uncontrolled medical or neurological condition (other than Alzheimer's disease) that may be a contributing cause of the subject's cognitive impairment
2. Clinically significant unstable psychiatric illness within 6 months prior to screening
3. Transient ischemic attack or stroke or any unexplained loss of consciousness within 1 year prior to screening
4. Brain MRI performed at screening that shows evidence of any of the following: acute or sub-acute hemorrhage, prior microhemorrhage or prior subarachnoid hemorrhage (unless finding is not due to an underlying structural or vascular hemorrhage), greater than 4 microhemorrhages, cortical infarct, >1 lacunar infarct, superficial siderosis or history of diffuse white matter disease
5. Contraindications to having a brain MRI or PET scan
6. History of bleeding disorder
7. Use of medications with platelet anti-aggregant or anti-coagulant properties (unless aspirin at ≤ 325 mg daily)
8. Uncontrolled hypertension or history of unstable angina, myocardial infarction, chronic heart failure or clinically significant conduction abnormalities
9. Participation in any active immunotherapy study targeting A β , any passive immunotherapy study targeting A β within 12 months of screening or any study with purported disease-modifying effect in AD within 12 months of screening unless documentation of receipt of placebo

Reviewer Comment: The patient population is consistent with Stage 3 and Stage 4 patients as described in the FDA 2018 Guidance for Industry Early Alzheimer's Disease: Developing Drugs for Treatment.

Dose Selection

Two dose levels (hereafter referred to as low dose and high dose) were chosen to balance dose-dependent A β reduction and occurrence of ARIA. The high dose was chosen to maximize A β reduction and the low dose was chosen to achieve considerable, but lower, reduction of A β while lowering the incidence of ARIA. Dose levels were initially based on the results of the fixed-dose cohorts from Study 103 (1, 3, 6 and 10 mg/kg) which demonstrated a time- and dose-dependent reduction of brain amyloid and a dose-dependent slowing of decline for the exploratory endpoints CDR-SB and MMSE. The incidence of ARIA was also dose- and ApoE ϵ 4 carrier-dependent in the fixed-dose cohorts of Study 103. It was hypothesized that dose titration would minimize the incidence of ARIA in Study 302. Therefore, dosing was initially dependent on ApoE ϵ 4 carrier status with ApoE ϵ 4 carriers titrated to lower doses (3 mg/kg low dose and 6 mg/kg high dose) than ApoE ϵ 4 non-carriers (6 mg/kg low dose and 10 mg/kg high dose). In parallel with the initiation of Study 302, the applicant added a cohort to Study 103 to assess the impact of titration to 10 mg/kg in ApoE ϵ 4 carriers on the incidence and severity of ARIA. The incidence of ARIA and discontinuations from study treatment due to ARIA in this cohort appeared to be reduced compared to ApoE ϵ 4 carriers who received a fixed dose of 10 mg/kg throughout the study. Following this analysis, the high dose (after titration) for ApoE ϵ 4 carriers in Study 302 was increased from 6 mg/kg to 10 mg/kg. The target 14 doses of 10 mg/kg administered over 78 weeks in Study 302 is equivalent to the number of 10 mg/kg doses administered over 54 weeks in the fixed-dose cohort in Study 103 (see Section 5.3.1).

Study Treatments

IV infusions of aducanumab or placebo were administered every 4 weeks over 76 weeks for a total of 20 doses. The titration period lasted 8 or 24 weeks (2 to 6 doses) depending on the target dose. Target doses for both the low- and high-dose treatment arms initially depended on ApoE ϵ 4 carrier status. After implementation of protocol version 4, the target dose in the high-dose arm was 10 mg/kg for both ApoE ϵ 4 carriers and non-carriers. After implementation of protocol version 4, subjects enrolled under protocol versions prior to version 4 and assigned to the ApoE ϵ 4 carrier high-dose arm were titrated to 10 mg/kg following receipt of at least 2 doses of 6 mg/kg. The dosing scheme is illustrated in Table 2.

Table 2: Dosing Scheme for Aducanumab by Treatment Group and ApoE ε4 Carrier Status

Dose (Week)		0	4	8	12	16	20	24 to 76
Treatment Group		Dose (mg/kg)						
ApoE ε4 carrier	Low Dose	1	1	3	3	3	3	3
	High Dose (Protocol Version 1-3)	1	1	3	3	3	3	6
	High Dose (Protocol Version ≥4)	1	1	3	3	6	6	10
	Placebo		Saline					
ApoE ε4 non-carrier	Low Dose	1	1	3	3	3	3	6
	High Dose	1	1	3	3	6	6	10
	Placebo		Saline					

Adapted from Table 12 in the Study 302 protocol.

Assignment to Treatment

An automated interactive voice/web response system (IxRS) was used to manage randomization and treatment assignment. Patients were randomized in a 1:1:1 ratio to placebo, low dose or high dose. Randomization was stratified by site and by ApoE ε4 carrier status (carrier or non-carrier) and enrollment was monitored such that 80% of the population included patients with a baseline clinical stage of MCI due to Alzheimer’s disease.

Blinding

A placebo match was not provided for the study, so an unblinded pharmacist managed all aspects of study treatment receipt, dispensing and preparation. All other study site staff and patients were blinded to treatment assignment during the placebo-controlled period. Efficacy assessments were conducted by two independent health care professionals (HCPs) not involved in patient care or management who were to remain blinded to treatment assignment and any other information that had the potential to reveal treatment assignment, including status of dosing with aducanumab, concomitant therapy, laboratory data, imaging data or adverse events (AEs). The treating HCP was not allowed to discuss AEs, including occurrence of ARIA, with the independent rating HCPs. The occurrence of certain ARIA events during the titration period required additional monitoring to continue through completion of the titration period. The titration period varied depending on the target dose, so to maintain the blind it was just assumed that all patients who required additional monitoring were titrating to the 10 mg/kg dose. For the long-term extension (LTE) period, study site staff and patients were blinded to dose level.

Reviewer Comment: The occurrence of ARIA has the potential to cause functional unblinding of investigators, patients and caregivers because ARIA is associated with aducanumab treatment and it prompts differential management of patients, including additional MRIs and dose modification. The potential for this functional unblinding was unavoidable and the applicant has

taken steps to address it in the study. This was recognized and reflected in the issuance of FDA agreement to the applicant's request for a special protocol assessment of the protocol. One important safeguard against this potential unblinding is the use of rating HCPs who are independent of patient care and blinded to dosing and adverse events. Patients and caregivers may become functionally unblinded though, which can influence rating scales that require more of their input, such as the ADCS-ADL-MCI. Ultimately, the potential of functional unblinding to influence the efficacy results is an important issue and is addressed in greater detail in Section 6.1.7.

Dose modification/Dose Discontinuation

Dose modification criteria were established to account for the expected occurrence of ARIA-E and ARIA-H. Dose reduction, suspension, or termination were dependent on the radiographic severity of ARIA as detected by MRI, the presence or absence of clinical symptoms, and the severity of symptoms, if present. ARIA dose management was initially informed by results from fixed-dose cohorts of Study 103 and expert recommendations but evolved over time as the scientific field gained experience managing ARIA and as additional data became available from Study 103. Specifically, protocol versions 3 and above enabled more patients to continue treatment with aducanumab and to reach their assigned target dose. Dosing guidelines for different protocol versions are summarized for ARIA-E in Table 3 and ARIA-H in Table 4. Patients who developed ARIA-H coincident with ARIA-E were to follow the more restrictive guideline. If dosing was suspended prior to reaching the target dose the patient must receive at least two doses at the restart dose before titrating to the next dose level and must complete the required number of doses at that dose level per their assigned group.

Table 3: Dose Modification/Discontinuation Rules for ARIA-E by Protocol Version (Study 302)

Clinical Symptom Severity		MRI Severity	Protocol Version ¹		
			Version 1	Version 3	Version 4-6
Asymptomatic		Mild	Continue dosing at same dose and schedule		
		Moderate or Severe	Suspend dosing and restart at next lower dose	Suspend dosing and restart at the same dose	
Symptomatic	Mild or Moderate	Any	Permanently discontinue		
	Severe				
	Serious “other medically				

	important event”²			
	Serious, except “other than medically important event”		Permanently discontinue	

Created by reviewer, modified from Table 11 in Study 302 CSR

¹ No patients were consented under Protocol Version 2

² “Other medically important events” include SAEs that were not life-threatening, did not require inpatient hospitalization or prolongation of existing hospitalization and did not result in significant or permanent disability

Note: Dosing was restarted when ARIA-E and/or clinical symptoms resolved

Note: From protocol version 3-5, if a second ARIA event (either ARIA-E or ARIA-H) required dose suspension, patients were to restart at the next lower dose. If a third ARIA event required dose suspension, drug was to be permanently discontinued. For protocol version 6, patients were to restart at the same dose after the third ARIA event

Table 4: Dose Modification/Discontinuation Rules for ARIA-H Microhemorrhage and Superficial Siderosis by Protocol Version (Study 302)

Clinical Symptom Severity		MRI Severity	Protocol Version ¹	
			Version 1	Version 3-6
Asymptomatic		Mild	Continue dosing at same dose and schedule	
		Moderate	Suspend dosing and restart at next lower dose	Suspend dosing and restart at same dose
		Severe	Permanently discontinue	
Symptomatic	Mild or Moderate	Mild or Moderate	Suspend dosing and restart at next lower dose	Suspend dosing and restart at same dose
		Severe	Permanently discontinue	Permanently discontinue
	Severe	Mild or Moderate		Suspend dosing and restart at same dose
		Severe		Permanently discontinue
	Serious “other medically important event” ²	Mild or Moderate		Suspend dosing and restart at same dose
		Severe		Permanently discontinue
	Serious, except “other than medically important event”	Any		

Created by reviewer, modified from Table 12 in Study 302 CSR

¹ No patients were consented under Protocol Version 2

² “Other medically important events” include SAEs that were not life-threatening, did not require inpatient hospitalization or prolongation of existing hospitalization and did not result in significant or permanent disability

Note: Dosing was restarted when ARIA-H stabilized, and clinical symptoms resolved

Note: From protocol version 3-5, if a second ARIA event (either ARIA-E or ARIA-H) required dose suspension, patients were to restart at the next lower dose. If a third ARIA event required dose suspension, drug was to be permanently discontinued. For protocol version 6, patients were to restart at the same dose after the third ARIA event

Administrative Structure

A steering committee blinded to treatment assignment was established to provide scientific and medical direction and to oversee the administrative progress of the study. Members of the committee included study medical directors from Biogen and external experts in Alzheimer's disease. An independent Data Monitoring Committee (IDMC) was formed to monitor the overall conduct of the study and review safety on an ongoing basis and could recommend protocol modifications, dose suspension, dose termination, or study termination. The IDMC was also responsible for reviewing the futility analysis. At each meeting the IDMC made a recommendation to continue, stop, or modify the study.

A centralized imaging laboratory was selected to read and interpret PET and MRI scans. The central laboratory was to notify the PI and the applicant of ARIA-E and ARIA-H findings. For the purpose of study conduct, readings from the central reader prevailed over those from the local radiologist.

A central electronic clinical outcomes vendor was selected to ensure standardization of clinical outcome assessments. Selected assessments were reviewed by central raters for consistency. If the central review resulted in feedback to the site rater, the site rater either agreed with the feedback and updated the assessment or provided rationale for not making the update.

Procedures and Schedule

The schedule for key assessments is presented in Table 5. The screening period consisted of three visits within a 60-day period before administration of the first dose. Eligible subjects reported to the study site to receive study treatment every 4 weeks for 76 weeks. The end-of-treatment visit occurred on Week 78 and the follow-up visit occurred on Week 94.

Table 5: Study 302 Schedule of Key Assessments

Assessment	Schedule
Eligibility Criteria	Screening V1, V2 and V3
ApoE Genotyping	Screening V1
Neurological and Physical Examinations	Screening V1, Week 12, 24, 48, 72, 78, 94
Brain MRI	Screening V2, Weeks 14, 22, 30, 42, 54, 66 and 78
Study Drug Infusion	Weeks 1, 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, 56, 60, 64, 68, 72, 76
Anti-Aducanumab Ab	Weeks 1, 24, 32, 60, 78, 94
Aducanumab Concentration	Weeks 1, 4, 12, 16, 20, 26, 28, 52, 56, 78, 94

RNA, Serum and Plasma for Biomarkers	Screening V1, Weeks 8, 12, 24, 28, 56, 78, 94
CSF Collection (optional)	Screening V1, Week 78
Amyloid PET	Screening V3, Weeks 26 and 78
Tau PET	Screening V3, Weeks 26 and 78
RBANS	Screening V1
CDR	Screening V1, Weeks 26, 50, 78 and 94
MMSE	Screening V1, Weeks 26, 50, 78 and 94
ADCS-ADL-MCI, ADAS-Cog 13	Screening V2, Weeks 26, 50, 78 and 94
NPI-10 and EQ-5D	Screening V2, Weeks 26, 50, 78

Created by reviewer, modified from Table 1 and Table 2 in Study 302 protocol

Note: There were three screening visits (V1, V2 and V3) within 60 days of randomization

Note: Amyloid PET and tau PET were assessed in different subgroups of patients

Note: Additional MRIs were collected in response to ARIA

Concurrent Medications

Medications for treatment of Alzheimer’s disease, including but not limited to donepezil, rivastigmine, galantamine, and memantine were allowed if patients were receiving a stable dose for at least 8 weeks prior to Screening Visit 1 and were to stay on a stable dose during the study. Patients were instructed to continue the medications they were receiving at enrollment without changes and to avoid starting new medications. An unscheduled visit for collection of all clinical assessments for the primary and secondary endpoints should have occurred before any change in Alzheimer’s disease medications.

The following medications were not allowed:

- Medications with platelet anti-aggregant or anti-coagulant properties, except aspirin at doses ≤ 325 mg per day
- Non-prescription narcotic medication
- Immunosuppressive drugs including systemic corticosteroids.
- Parenteral immunoglobulin, blood products, plasma derivatives, plasma exchange and plasmapheresis
- Any investigational drug

Subject Completion, Discontinuation or Withdrawal

Patients who completed the 78-week treatment period and follow-up visit were considered to have completed the study.

Patients who discontinued treatment were to remain in the study, attend a follow-up visit 18 weeks after the final visit and continue protocol-required tests and assessments at a subset of the clinical visits listed in Table 5 until the end of the study or until the patient withdrew consent. The most notable treatment-related reason for study discontinuation was the occurrence of ARIA as outlined in Table 3 and Table 4. Other reasons for discontinuation include

withdrawal of consent, medical emergency, AEs that do not resolve, severe infusion reaction, and discretion of the investigator for medical reasons or noncompliance. The reason for discontinuation of study treatment was to be recorded in the patient's case report form (CRF).

Patients who were withdrawn from the study after receiving at least one dose of study treatment were expected to complete the end-of-treatment visit and return to the site for a follow-up visit 18 weeks after receiving their last dose of study treatment. Efficacy assessments were not required if the patients discontinued treatment within 3 months of the previous efficacy assessment and significant changes in cognitive status were not suspected by the investigator. Reasons for withdrawal of patients include withdrawal of consent, unwillingness or inability to comply with the protocol and discretion of the investigator or applicant. The reason for withdrawal was to be recorded in the patient's CRF.

Study Endpoints

Primary Efficacy Endpoint

The primary efficacy endpoint was the change from baseline in CDR-SB at Week 78. The CDR-SB assesses three domains of cognition (memory, orientation, judgment/problem solving) and three domains of function (community affairs, home/hobbies, personal care) using semi-structured interviews with the patient and a reliable companion or informant. A qualified rater uses interview data and clinical judgment to assign scores for each domain ranging from none=0, questionable = 0.5, mild = 1, moderate = 2 to severe = 3. The personal care domain does not include the 0.5 score. Scores from each domain are summed to provide the CDR-SB value ranging from 0 to 18, with higher scores indicating greater disease severity. CDR-SB has been described in the literature as a suitable primary endpoint for clinical trials in patients with early Alzheimer's disease due to its psychometric properties and its ability to assess both cognitive and functional disability and is routinely accepted by FDA as an acceptable primary outcome assessment for studies of Alzheimer's disease intended to demonstrate substantial evidence of effectiveness. It has been widely used as the primary efficacy endpoint in clinical trials for other investigational drugs in this population.

CDR-SB assessments were conducted by an independent HCP not involved in patient care or management who remained blinded to treatment assignment. Given the importance of the reliable informant or companion, the inclusion criteria included a stipulation that the informant should be available for the duration of the study and the use of the same informant for the duration of the study was encouraged. Raters were required to complete qualification and training prior to being eligible to administer the assessment. All sites were asked to maintain the same rater throughout the study to ensure consistency. If a rater administered the CDR-SB to a patient, that rater was not allowed to administer the other neurocognitive assessments to that patient at any point during the study. A contract research organization (CRO), Medavante,

was selected to manage rater training and eligibility and to ensure standardization of clinical outcome assessments with special focus on CDR-SB. Selected assessments were reviewed by central raters for consistency. If the central review resulted in feedback to the site rater, the site rater either agreed with the feedback and updated the assessment or provided rationale for not making the update.

Reviewer Comment: CDR-SB is an integrated scale that adequately and meaningfully assesses both daily function and cognitive effects in early Alzheimer's disease and is consistent with FDA guidance on clinical endpoints in Stage 3 patients. The distinction between cognitive and functional domains for the CDR-SB is somewhat artificial because the effects on cognition are measured in a way that reflect impact on function and are clinically meaningful. The SPA agreement for this protocol indicated FDA's concurrence with the choice of primary endpoint.

Secondary Endpoints

MMSE

The MMSE is a widely used performance-based assessment of cognitive ability consisting of 11 tasks evaluating orientation, word recall, attention and calculation, language, and visuospatial functions. The scores from the 11 tests are summed to obtain a total score, which ranges from 0 to 30 with lower scores indicating greater cognitive impairment. It is often used in clinical practice or as a staging instrument for trial inclusion, and is also used as an efficacy assessment in clinical trials.

ADAS-Cog 13

The ADAS-Cog is a cognitive assessment consisting of clinical ratings and cognitive tasks that was originally developed for use in clinical trials of patients with later stages of Alzheimer's disease dementia. ADAS-Cog 11 includes 11 tasks measuring disturbances of memory, language, and praxis. Many of the items of the ADAS-Cog 11 are at the measurement floor in patients with mild disease and may not show decline over the length of a typical clinical trial. Therefore, two additional tasks, delayed word recall and number cancellation, were added to create the ADAS-Cog 13 for use in this earlier disease population. The ADAS-Cog 13 includes 9 items that test performance (up to 65 points) and 4 clinician-rated items that test language and memory (up to 20 points) with a total score ranging from 0 to 85 with higher scores indicating greater cognitive impairment. The 13 items assess word recall, ability to follow commands, constructional praxis, naming, ideational praxis, orientation, word recognition, comprehension of spoken language, memory, word-finding, language ability, delayed word recall, and concentration.

The ADAS-Cog was a key clinical endpoint in the trials that led to the approval of the acetylcholinesterase inhibitors. For many of the initial trials of anti-A β antibodies, ADAS-Cog 11

was a co-primary endpoint with a functional outcome such as ADCS-ADL. More recently, another variation of the ADAS-Cog, ADAS-Cog 14, was used as the primary endpoint in the trial for solanezumab in patients with mild dementia due to Alzheimer's disease.

ADCS-ADL-MCI

The ADCS-ADL-MCI is a questionnaire for informants that consists of 17 instrumental items and 1 basic item (getting dressed) intended to reflect activities of daily living. Informants are asked whether the patient attempted each item during the prior 4 weeks and their level of performance. Responses are "Yes," "No," or "Don't Know" with additional sub-ratings depending on the item. The total score ranges from 0 to 53 with lower scores indicating greater impairment. The ADCS-ADL-MCI was adapted from the ADCS-ADL, which was developed for a population with more advanced disease and served as a key endpoint in many of the acetylcholinesterase trials.

The secondary endpoints were assessed by a second HCP who was independent from the rater responsible for administering the CDR-SB. As with CDR-SB, all sites were asked to maintain the same rater throughout the study to ensure consistency. If raters administered the CDR-SB to a patient, they were not allowed to administer the secondary endpoint assessments to that patient at any point during the study.

The applicant performed a principal components analysis of the 48 individual items from CDR-SB, MMSE, ADAS-Cog 13 and ADCS-ADL-MCI using baseline and change from baseline data from Study 301 and Study 302 to assess the degree to which the endpoints capture similar or distinct information. The applicant concluded that the overlap between the items was 5-25% and therefore each scale carries independent information on the clinical state of the patient.

Reviewer Comment: FDA routinely encounters the use of these measures and they are appropriate selections for use in supporting an effect on an acceptable primary measure. The principal components analysis indicated that while there may be overlap among the 4 clinical endpoints, each also captures distinct information regarding cognitive decline. Effects on each of these endpoints can independently contribute to the persuasiveness of a specific study.

Tertiary Endpoint

NPI-10 was included as a tertiary endpoint and was completed by interview with the informant. The NPI-10 assesses the presence, frequency, and severity of 10 neuropsychiatric domains: delusions, hallucinations, dysphoria, apathy, euphoria, disinhibition, agitation/aggression, irritability/liability, anxiety, and aberrant motor behavior. The total score ranges from 0 to 120 with higher scores indicating worse symptoms.

Key Pharmacodynamic Endpoints

Key pharmacodynamic endpoints include the following:

- Change from baseline in amyloid signal as measured by ^{18}F -florbetapir PET and quantified by a composite SUVR was assessed in a subset of sites and patients (approximately 400) at Week 26 and Week 78. The standard uptake value ratio (SUVR) was calculated for a composite of brain regions consisting of frontal, parietal, lateral temporal, sensorimotor, and anterior and posterior cingulate, and occipital cortices with whole cerebellum as a reference region.
- Change from baseline in CSF levels of $\text{A}\beta_{1-42}$, $\text{A}\beta_{1-40}$, phosphorylated tau at residue 181 (p-Tau), and total tau (t-Tau) at Week 78 in a subset of patients
- Change from baseline in tau PET as measured by ^{18}F -MK-6240 PET and quantified by a composite SUVR at the Week 78 or end-of-treatment visit in a subset of patients. The SUVR was calculated for 6 composite regions of interest (frontal, temporal, medial temporal, parietal, cingulate, and occipital) with cerebellar cortex as the reference region.
- Change in brain volume (whole brain, whole cortex, hippocampus, and lateral ventricle) as measured by MRI at Week 30 and Week 78.

Statistical Analysis Plan

The Statistical Analysis Plan (SAP) was finalized on September 11, 2018, before the termination of the study and an addendum was added on November 4, 2019 prior to database lock, in response to the futility declaration of March 21, 2019. The addendum to the SAP did not alter the prespecified primary analysis methods as documented in the September 2018 SAP, except for specifying that the primary analysis would exclude efficacy data collected after March 20, 2019, per agreement with the Division at the October 21, 2019, Type C Meeting. Additionally, no censoring of data was applied to biomarker, PK, and safety data.

Interim Analysis – Blinded Sample Size Reassessment

The sample size was planned to have approximately 90% power to detect a true mean difference of 0.5 in change from baseline CDR-SB at Week 78. The mean difference of 0.5 represents an approximately 25% reduction assuming the placebo mean change in CDR-SB is 2 at Week 78. Sample size was reassessed approximately 3 months before enrollment completion in November of 2017. At the time of this reassessment approximately 10% of the data was available on the primary endpoint from the pooled blinded data from Studies 301 and 302. As a result of the analysis, the sample size was increased from 1350 to 1605 (from 450 to 535 per treatment arm).

Interim Analysis – Futility Assessment

An interim futility analysis was planned to occur after approximately 50% of the patients enrolled in Studies 301 and 302 had the opportunity to complete the Week 78 visit. Futility criteria were primarily based on conditional power, which was defined as the chance that the primary efficacy endpoint analysis would be statistically significant in favor of aducanumab at the planned final analysis, given the data at the futility analysis. Conditional power assumed that the future unobserved treatment effect would be equal to an estimate based on pooled observed data from Studies 301 and 302. The studies were to be considered futile if both studies had a conditional power for the primary efficacy analysis that was <20% in both the high-dose and low-dose treatment groups. The SAP specified that other data in addition to the pre-specified criteria could also be considered.

The data cutoff date for the futility analysis was December 26, 2018, at which point 49% of patients from Study 302 had the opportunity to complete the Week 78 visit. Data from patients who were enrolled in the study but had not had the opportunity to complete the Week 78 visit were not included in the analysis. An independent, unblinded statistician from a CRO (IQVIA) conducted the futility analysis. The IDMC received the futility analysis package on March 8, 2019, and met with the unblinded statistician 4 days later. The IDMC communicated its recommendation to the Biogen Senior Decision Team, which was independent of the study team, on March 19, 2019. The Biogen Senior Decision Team requested additional analyses which were conducted by unblinded statisticians who were employed by the applicant but not part of the study team. The Biogen Senior Decision Team recommended termination of the studies to the Biogen Executive Decision Team on March 20, 2019, and the Biogen Executive Team made the final decision to terminate all ongoing studies of aducanumab. On March 21, 2019, the applicant publicly announced the discontinuation of the aducanumab program. See Section 6.1.6 for a detailed discussion of the futility analysis and its implications.

Definitions of Statistical Analysis Populations

The following analysis populations were defined:

- Intent-to-Treat (ITT) Population – all randomized subjects who received at least one dose of study treatment and excluding data collected after March 20, 2019
- Per-Protocol Population – all subjects in the ITT population who had no violations for the following inclusion criteria: 6 years of education or work experience, positive PET scan, CDR global score of 0.5, RBANS score of 85 or lower, and MMSE score between 24 and 30. Also, subjects had ≥70% expected infusions and did not make changes to concomitant Alzheimer’s disease medications during the study

- Opportunity to Complete (OTC) Population – subjects in the ITT population who had the opportunity to complete the Week 78 visit by March 20, 2019
- ITT Population during the Double-Blind Period – subjects in the ITT population including all data collected until April 17, 2019. The applicant began releasing treatment assignments to the CRO on April 18, 2019 to be distributed to the sites upon request.
- Uncensored ITT Population – subjects in the ITT population with all data collected during the study
- ¹⁸F-florbetapir Amyloid PET Analysis Population – all randomized subjects who received at least one dose of study treatment, used ¹⁸F-florbetapir ligand for their amyloid PET scan and had an evaluable amyloid PET SUVR value for the composite region-of-interest using cerebellum as the reference region

Analysis Method for Primary Endpoint

A mixed model repeated measures (MMRM) model was used to analyze change from baseline CDR-SB using fixed effects of treatment, time, treatment-by-time interaction, baseline CDR-SB, baseline CDR-SB by time, baseline MMSE, Alzheimer's disease medication use at baseline, region and laboratory ApoE ε4 status.

Missing Data

For the MMRM analysis, missing data were assumed to be missing at random. Different assumptions for missing data were explored as part of sensitivity analyses.

Adjustments for Multiplicity

A sequential (closed) testing procedure was used to control the overall Type I error rate according to the following order: high-dose aducanumab versus placebo and low-dose aducanumab versus placebo. All comparisons after the initial comparison with $p > 0.05$ will not be considered statistically significant. Secondary endpoints are rank prioritized according to the following order: MMSE, ADAS-Cog 13, ADCS-ADL-MCI. To control for Type I error for each of the secondary endpoints, a sequential (closed) procedure, including both the order of the secondary endpoints and treatment comparisons were used.

Subgroup Analyses

Subgroup analyses for CDR-SB, MMSE, ADAS-Cog 13, and ADCS-ADL-MCI were conducted for

the following pre-defined groups:

- Laboratory ApoE status (carrier or non-carrier)
- Baseline clinical stage (MCI due to Alzheimer's disease or mild Alzheimer's disease)
- Use of Alzheimer's disease concomitant medication at baseline (yes or no)
- Baseline MMSE (MMSE ≤ 26 or MMSE ≥ 27)
- Region (United States, Europe/Canada/Australia, Asia)
- Age (≤ 64 , 65-74, ≥ 75)

Reviewer Comment: Baseline clinical stage and baseline MMSE are both ways to evaluate disease severity. The distinction between MCI due to Alzheimer's disease and mild Alzheimer's disease is somewhat subjective. Also, only 20% of enrolled patients had mild Alzheimer's disease, further complicating interpretation of potential subgroup effects. The baseline MMSE criterion is more objective and roughly divides the population in half.

Protocol Amendments

The original protocol (Version 1) was issued on April 9, 2015. No patients were consented under Version 2. The main consequence of subsequent protocol versions was to increase the dose level or number of doses received by patients. Notably, Protocol Version 4, dated March 24, 2017, allowed ApoE $\epsilon 4$ carriers randomized to aducanumab high dose to receive the same aducanumab dose received by ApoE $\epsilon 4$ non-carriers (titration to 10 mg/kg). Table 3 and Table 4 outline the changes to the protocol, including Protocol Version 3 dated July 21, 2016, which modified dosing in relation to ARIA management. Protocol Version 6, dated June 28, 2018, updated the sample size from 450 to 535 per treatment group following the blinded sample size re-estimation.

5.1.2. Study Results

Compliance with Good Clinical Practices

The applicant asserts that the study was performed in accordance with 21 CFR parts 50, 54, 56 and 312 Subpart D, ICH Guideline on GCP (E6) and the ethical principles outlined in the Declaration of Helsinki.

Financial Disclosure

The applicant has adequately disclosed financial interests or agreements with clinical investigators as outlined in the guidance for industry Financial Disclosures by Clinical Investigators.

Patient Disposition

A total of 6757 patients were screened for entry into the study and 1643 patients were randomized. The most common reasons reported for screening failure were having a CDR global score, MMSE score, or RBANS score outside the allowed range (62%) or not having a positive amyloid PET scan (16%). There were 5 patients who were randomized but did not receive study treatment, with the most common reason being that the patient should have been classified as a screening failure. Patient disposition is summarized in Table 6. The most common reason for treatment discontinuation is “other.” This category includes patients who terminated the study due to the administrative decision to terminate aducanumab studies based on the interim futility analysis (placebo, 221 (40.3%); low dose, 198 (36.5%); high dose, 187 (34.2%)).

Table 6: Study 302 Patient Disposition

Disposition	Study 303		
No. of patients screened	6757		
No. of patients not randomized	5114		
	Aducanumab Low Dose N=543 n (%)	Aducanumab High Dose N=547 n (%)	Placebo N=548 n (%)
Patients randomized	547	547	549
ITT population	543 (100%)	547 (100%)	548 (100%)
Per-protocol population	387 (71.3%)	365 (66.7%)	427 (77.9%)
OTC population	329 (60.6%)	340 (62.2%)	313 (57.1%)
¹⁸ F-florbetapir amyloid PET	159 (29.3%)	170 (31.1%)	159 (29.0%)
Discontinued treatment	264 (48.6%)	265 (48.4%)	254 (46.4%)
Adverse event	42 (7.7%)	49 (9.0%)	17 (3.1%)
Consent withdrawn	22 (4.1%)	20 (3.7%)	6 (1.1%)
Other reasons	200 (36.8%)	196 (35.7%)	231 (42.2%)
Discontinued study	252 (46.4%)	252 (46.1%)	260 (47.4%)
Adverse event	13 (2.4%)	20 (3.7%)	10 (1.8%)
Consent withdrawn	32 (5.9%)	26 (4.8%)	15 (2.7%)
Other reasons	207 (38.1%)	206 (37.6%)	235 (42.9%)

Created by the reviewer using ie.xpt and Tables 13, 20 and 50 in Study 302 CSR

Protocol Violations/Deviations

Two-thirds of patients had at least one major protocol deviation during the placebo-controlled period (placebo, 349 (63.7%); low dose, 354 (65.2%); high dose, 388 (70.9%)). The only major protocol deviation with ≥5% difference between the high-dose treatment arm and placebo was in the “study procedures” category. The “study procedures” category included a variety of protocol deviations and none seemed to predominate. Overall, the most common categories of

major protocol deviations were regarding informed consent (27.8%), study procedures (26.9%), and investigational product compliance (23.7%). These deviations are not expected to affect interpretation of the overall results. During the combined screening and placebo-controlled periods, 101 patients (36 placebo, 29 low dose, 36 high dose) had a major protocol deviation for eligibility or entry criteria. There were 4 patients who did not meet the CDR global score criterion, 2 who did not meet the MMSE criterion and 10 who did not meet the RBANS criterion. These numbers are small relative to the overall population enrolled and balanced across the treatment arms and are therefore not expected to complicate efficacy analyses. A total of 10.1% of patients received disallowed concomitant medications during the placebo-controlled period, but many of these were anti-coagulants or corticosteroids, disallowed because of a possible interaction with the safety of aducanumab and were not anticipated to have an impact on the effectiveness of aducanumab. Also, 13 patients were randomized to a carrier status that was discordant with laboratory assay results. There was one patient randomized to placebo who received aducanumab treatment in the visit before the start of the LTE. Single dose administration of aducanumab near the end of the study would not have an impact on the efficacy results.

Table of Demographic Characteristics

Table 7 contains information regarding demographic characteristics for each treatment arm in the ITT population. Demographic characteristics were balanced across the treatment arms and generally representative of the patient population except for an under-representation of African American patients. Overall, 39.8% of patients were enrolled in the United States.

Table 7: Study 302 Baseline Demographics (ITT Population)

Demographic Parameters	Placebo (N=548) n (%)	Treatment Group		
		Aducanumab Low Dose (N=543) n (%)	Aducanumab High Dose (N=547) n (%)	Total (N=1638) n (%)
Sex				
Male	258 (47.1%)	274 (50.5%)	263 (48.1%)	795 (48.5%)
Female	290 (52.9%)	269 (49.5%)	284 (51.9%)	843 (51.5%)
Age				
Mean years (SD)	70.8 (7.4)	70.6 (7.5)	70.6 (7.5)	70.7 (7.4)
Median (years)	71.0	72.0	72.0	72.0
Min, max (years)	50, 85	50, 85	50, 85	50, 85
Age Group				
≤ 64 years	104 (19.0%)	106 (19.5%)	113 (20.7%)	323 (19.7%)
> 64 – <75 years	255 (46.5%)	262 (48.3%)	257 (47.0%)	774 (47.3%)
≥ 75 years	189 (34.5%)	175 (32.2%)	177 (32.4%)	541 (33.0%)
Race				

FDA Clinical Review
Kevin Krudys, PhD
BLA 761178
Aduhelm (aducanumab)

White	431 (78.6%)	432 (79.6%)	422 (77.1%)	1285 (78.4%)
Black or African American	1 (0.2%)	6 (1.1%)	4 (0.7%)	11 (0.7%)
Asian	47 (8.6%)	39 (7.2%)	42 (7.7%)	128 (7.8%)
American Indian or Alaska Native	1 (0.2%)	0	0	1 (<0.1%)
Native Hawaiian or Other Pacific Islander	0	0	0	0
Other	1 (0.2%)	1 (0.2%)	3 (0.5%)	5 (0.3%)
Ethnicity				
Hispanic or Latino	22 (4.0%)	22 (4.1%)	23 (4.2%)	67 (4.1%)
Not Hispanic or Latino	470 (85.8%)	470 (86.6%)	461 (84.3%)	1401 (85.5%)
Not Reported ¹	56 (10.2%)	51 (9.4%)	62 (11.3%)	169 (10.3%)
Region				
United States	218 (39.8%)	218 (40.1%)	216 (39.5%)	652 (39.8%)
Rest of the World				
Europe/Canada	287 (52.4%)	287 (52.9%)	291 (53.2%)	865 (52.8%)
Asia	43 (7.8%)	38 (7.0%)	40 (7.3%)	121 (7.4%)

Source: Table 14 and Table 49 in Study 302 CSR

¹ Data on race and/or ethnicity were not collected because of local regulations.

Other Baseline Characteristics (disease characteristics, important concomitant drugs)

Table 8 contains a summary of baseline disease characteristics and baseline use of concomitant Alzheimer's disease medications. The disease characteristics are balanced across treatment arms and reflect a population of patients who are early in the course of Alzheimer's disease. By design, the overall population consisted of 80% of patients with MCI due to Alzheimer's disease and 20% with mild Alzheimer's disease dementia. The percentage of the population who were ApoE ε4 carriers is consistent with previous reports. Most patients were receiving concomitant medications for Alzheimer's disease (51.8%). Additionally, 15.3% of patients received any Alzheimer's disease medication and stopped prior to entering the study.

Table 8: Study 302 Disease Characteristics (ITT Population)

Disease Characteristics	Placebo (N=548) n (%)	Treatment Group		
		Aducanumab Low Dose (N=543) n (%)	Aducanumab High Dose (N=547) n (%)	Total (N=1638) n (%)
Baseline Clinical Stage				
MCI due to AD	446 (81.4%)	452 (83.2%)	438 (80.1%)	1336 (81.6%)
Mild AD	102 (18.6%)	91 (16.8%)	109 (19.9%)	302 (18.4%)
Laboratory ApoE ε4 Status				
Carrier	368 (67.2%)	362 (66.7%)	365 (66.7%)	1095 (66.8%)
Homozygote	92 (16.8%)	97 (17.9%)	77 (14.1%)	266 (16.2%)
Heterozygote	276 (50.4%)	265 (48.8%)	288 (52.7%)	829 (50.6%)
Non-carrier	178 (32.5%)	178 (32.8%)	181 (33.1%)	537 (32.8%)
Undetermined	2 (0.4%)	3 (0.6%)	1 (0.2%)	6 (0.4%)
Number of Years of Formal Education				
Mean years (SD)	14.5 (3.7)	14.5 (3.6)	14.5 (3.6)	14.5 (3.6)
Median (years)	15.0	15.0	15.0	15.0
Min, Max (years)	5, 29	3, 27	3, 29	3, 29
Number of Years Since Diagnosis of AD				
Mean years (SD)	1.3 (1.4)	1.3 (1.5)	1.3 (1.6)	1.3 (1.5)
Median (years)	0.8	0.8	0.8	0.8
Min, Max (years)	0, 10.0	0, 14.4	0, 16.1	0, 16.1
Concomitant AD medication				
Any AD medication at baseline	282 (51.5%)	281 (51.7%)	285 (52.1%)	848 (51.8%)
Only cholinesterase	235 (42.9%)	230 (42.4%)	228 (41.7%)	693 (42.3%)

inhibitors				
Only memantine	8 (1.5%)	15 (2.8%)	21 (3.8%)	44 (2.7%)
Both cholinesterase inhibitors and memantine	39 (7.1%)	36 (6.6%)	36 (6.6%)	111 (6.8%)
Baseline CDR-SB				
Mean (SD)	2.47 (1.00)	2.46 (1.01)	2.51 (1.05)	2.48 (1.02)
Median	2.50	2.50	2.50	2.50
Min, Max	0.5, 6.0	0.5, 5.5	0.5, 5.5	0.5, 6.0
Baseline MMSE				
<24	0	1 (0.2%)	1 (0.2%)	2 (0.1%)
≥24 - <27	296 (54.0%)	314 (57.8%)	296 (54.1%)	906 (55.3%)
≥27 - ≤30	252 (46.0%)	228 (42.0%)	250 (45.7%)	730 (44.6%)
Baseline ADAS-Cog 13	545	542	546	1633
Mean (SD)	21.9 (6.7)	22.5 (6.8)	22.2 (7.1)	22.2 (6.9)
Median	21.7	22.3	21.8	21.7
Min, Max	4.7, 45.7	2.0, 46.0	5.3, 57.7	2.0, 57.7
Baseline ADCS-ADL-MCI	545	540	545	1630
Mean (SD)	42.6 (5.7)	42.8 (5.5)	42.5 (5.8)	42.6 (5.7)
Median	43.0	44.0	43.0	43.0
Min, Max	11, 52	19, 53	5, 53	5, 53

Source: Tables 16, 17 and Table 59 in Study 302 CSR

Treatment Compliance, Concomitant Medications, and Rescue Medication Use

During study treatment, compliance with study treatment was 96.2%. Although compliance with study treatment was high, protocol-defined dose modification for ARIA, changes to the protocol and the termination of the study meant that some patients may have received fewer target doses. For example, only 9% of ApoE ε4 carriers received all 14 doses of aducanumab 10 mg/kg, compared to 37.2% of ApoE ε4 non-carriers. Similarly, 80.1% of ApoE ε4 carriers received any 10 mg/kg doses, compared to 91.1% of ApoE ε4 non-carriers. More patients in the lower dose treatment arm received the maximum dose level (89.7% for ApoE ε4 non-carriers receiving 6 mg/kg and 98.6% for ApoE ε4 carriers receiving 3 mg/kg).

Efficacy Results – Primary Endpoint

The primary efficacy endpoint analysis, change from baseline in CDR-SB at Week 78, demonstrated a statistically significant treatment effect in the aducanumab high-dose treatment arm compared to placebo (-0.39 [-22%], p=0.0120) (Table 9). The low-dose treatment arm demonstrated a numerical advantage compared to placebo (-0.26 [-15%]) but failed to reach statistical significance (p=0.0901). The results are consistent with a dose-response relationship.

Table 9: Study 302 Primary Endpoint Analysis

	Placebo (N=548)	Aducanumab Low Dose (N=543)	Aducanumab High Dose (N=547)
Baseline CDR-SB			
n	548	543	547
Mean	2.47	2.46	2.51
Change from Baseline in CDR-SB at Week 78			
n	288	290	299
Adjusted mean	1.74	1.47	1.35
Standard error	0.115	0.116	0.115
95% CI	(1.513, 1.963)	(1.247, 1.701)	(1.124, 1.573)
Difference from placebo		-0.26	-0.39
95% CI for difference		(-0.569, 0.041)	(-0.694, -0.086)
% difference vs. placebo		-15%	-22%
p-value (compared with placebo)		0.0901	0.0120

Source: Table 22 from Study 302 CSR

Reviewer Comment: The treatment effect is not much different than the estimate used to power the study (-25%). Power calculations assumed a 2-point decline in CDR-SB over 78 weeks, whereas the observed decline in Study 302 was lower (1.74).

Several SAP-defined and post hoc sensitivity analyses were performed for the primary endpoint. The copy increment from reference method and the jump to reference method were used to test the assumption that missing data were missing at random. Results from these analyses demonstrated that the statistically significant results for the primary endpoint were not sensitive to departures from the missing at random assumption. Notably, a sensitivity analysis for normality revealed that the distribution of the data was skewed, but the primary analysis was not sensitive to departures from normality.

Table 10 displays the results of the supplementary analyses of the primary endpoint in some of the key populations as defined in Section 5.1.1. Results in the uncensored population suggest that the magnitude of the treatment effect improves with additional data, including data after patients have stopped dosing for a period of time. Overall, the results support the robustness of the analysis of the primary endpoint in the ITT population.

Table 10: Study 302 Primary Endpoint Analysis (Dataset Sensitivity)

ITT Population during Double-Blind Period	Uncensored ITT Population	OTC Population
---	---------------------------	----------------

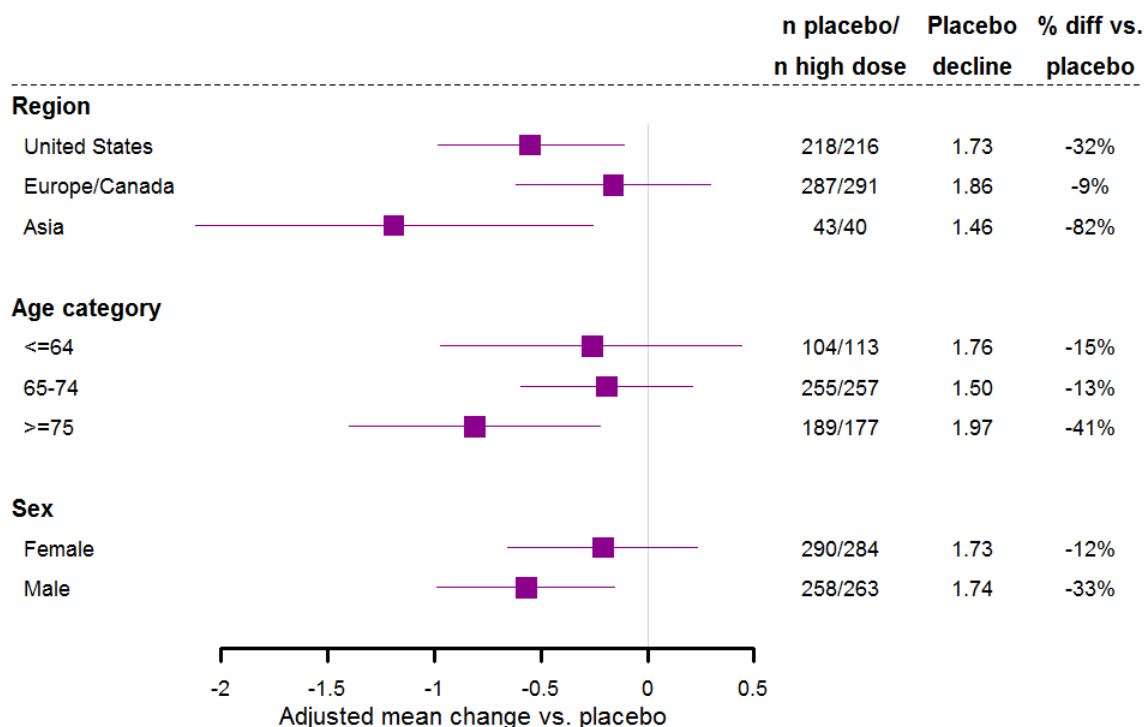
Difference vs. Placebo at Week 78 (%) p-value				Difference vs. Placebo at Week 78 (%) p-value			Difference vs. Placebo at Week 78 (%) p-value		
	PBO Decline (N=548)	Low Dose (N=543)	High Dose (N=547)	PBO Decline (N=548)	Low Dose (N=543)	High Dose (N=547)	PBO Decline (N=313)	Low Dose (N=329)	High Dose (N=340)
CDR-SB	n=340	n=329	n=349	n=408	n=399	n=403	n=288	n=290	n=298
	1.76	-0.23	-0.41	1.79	-0.22	-0.44	1.61	-0.27	-0.36
		-13%	-23%		-12%	-25%		-17%	-22%
		0.1339	0.0061		0.1273	0.0029		0.1188	0.0368

Source: Tables 73, 77 and 79 in Study 302 CSR
p-values are nominal

Subgroup Analysis of the Primary Efficacy Endpoint

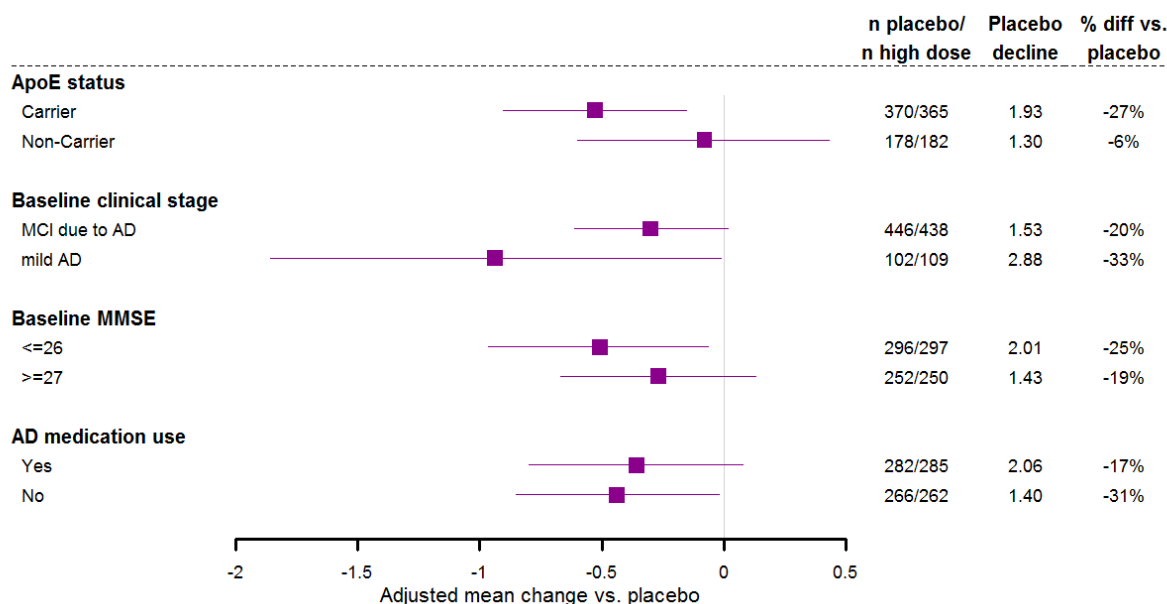
Subgroup analyses were performed on the primary efficacy endpoint for the groups described in Section 5.1.1 and are reported for the high-dose treatment arm in Figure 1 and Figure 2. All results trended in favor of high-dose aducanumab compared to placebo.

Figure 1: Study 302 High Dose Subgroup Analysis of the Primary Efficacy Endpoint (Demographics)



Source: Tables 135, 136 and 137 in Study 302 CSR.

Figure 2: Study 302 High Dose Subgroup Analysis of the Primary Efficacy Endpoint (Baseline Characteristics)



Source: Tables 131 to 134 in Study 302 CSR.

Data Quality and Integrity

There were no major data quality issues identified during the review of Study 301.

Efficacy Results – Secondary endpoints

A summary of the analysis results for the secondary endpoints is provided in Table 11. Statistically significant differences from placebo were observed for the high-dose treatment arm at Week 78 for all secondary endpoints. The low-dose treatment arm demonstrated favorable numerical trends for ADAS-Cog 13 and ADCS-ADL-MCI, but not for MMSE. Sensitivity analyses were also performed for secondary endpoints and were consistent with the primary analysis except for MMSE after censoring for intercurrent events. Although the p-value for this sensitivity analysis increased from 0.0493 to 0.1008, the magnitude of the treatment effect only dropped from -18% to -16%. Like CDR-SB, the secondary endpoints were also not normally distributed. Supplementary analysis using different datasets demonstrated similar results to the ITT population. Subgroup analyses for the secondary endpoints showed that all groups favored high-dose aducanumab except for one (MMSE in ApoE ε4 non-carriers). Subgroup findings for ApoE ε4 populations are discussed further in Section 6.1.3.

Table 11: Study 302 Secondary Endpoint Analysis

	Placebo (N=548)	Aducanumab Low Dose (N=543)	Aducanumab High Dose (N=547)
Baseline MMSE			
n	548	543	547
Mean	26.4	26.3	26.3
Change from Baseline in MMSE at Week 78			
n	288	293	299
Adjusted mean	-3.3	-3.3	-2.7
Standard error	0.22	0.22	0.21
95% CI	(-3.68, -2.83)	(-3.77, -2.92)	(-3.11, -2.27)
Difference from placebo		-0.1	0.6
95% CI for difference		(-0.65, 0.48)	(0.00, 1.13)
% difference vs. placebo		3%	-18%
p-value (compared with placebo)		0.7578	0.0493
Baseline ADAS-Cog 13			
n	545		
Mean	21.87		
Change from Baseline in ADAS-Cog 13 at Week 78			
n	287	289	293
Adjusted mean	5.16	4.46	3.76
Standard error	0.40	0.41	0.40
95% CI	(4.37, 5.96)	(3.66, 5.26)	(2.97, 4.55)
Difference from placebo		-0.70	-1.40
95% CI for difference		(-1.76, 0.36)	(-2.46, -0.34)
% difference vs. placebo		-14%	-27%
p-value (compared with placebo)		0.1962	0.0097
Baseline ADCS-ADL-MCI			
n	545	540	545
Mean	42.6	42.8	42.5
Change from Baseline in ADCS-ADL-MCI at Week 78			
n	283	286	295
Adjusted mean	-4.3	-3.5	-2.5
Standard error	0.38	0.38	0.38
95% CI	(-5.02, -3.53)	(-4.29, -2.79)	(-3.27, -1.79)
Difference from placebo		0.7	1.7
95% CI for difference		(-0.27, 1.73)	(0.75, 2.74)
% difference vs. placebo		-16%	-40%
p-value (compared with placebo)		0.1515	0.0006

Source: Tables 92, 104 and 116 in Study 302 CSR

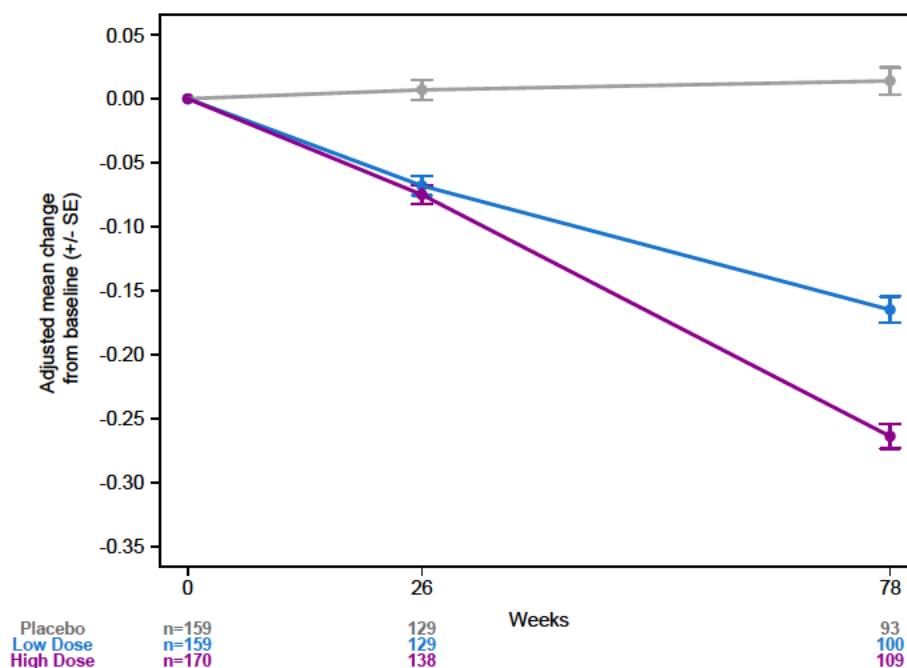
Efficacy Results – Tertiary endpoint

The adjusted mean change from baseline in NPI-10 at Week 78 compared to placebo was -1.3 (-87%, nominal $p = 0.0215$) for aducanumab high dose and -0.5 (-33%, $p=0.3921$) for aducanumab low dose.

Efficacy Results – Pharmacodynamic endpoints

Change from baseline in brain amyloid signal as measured by SUVR was analyzed with an MMRM model with fixed effects of treatment group, visit, treatment group-by-visit interaction, baseline SUVR, baseline SUVR-by-visit interaction, baseline MMSE, ApoE $\epsilon 4$ status, and baseline age. The adjusted mean change from baseline relative to placebo was similar in the low-dose and high-dose groups at Week 26, -0.075 and -0.082, respectively due to similar dosing during titration. At Week 78, the adjusted mean change from placebo was -0.179 and -0.278 in the low-dose and high-dose groups, respectively, indicating time- and dose-dependent relationships (Figure 3). Consistent relationships were observed for other brain regions and when using additional reference regions.

Figure 3: Study 302 Change from Baseline in A β PET Composite SUVR

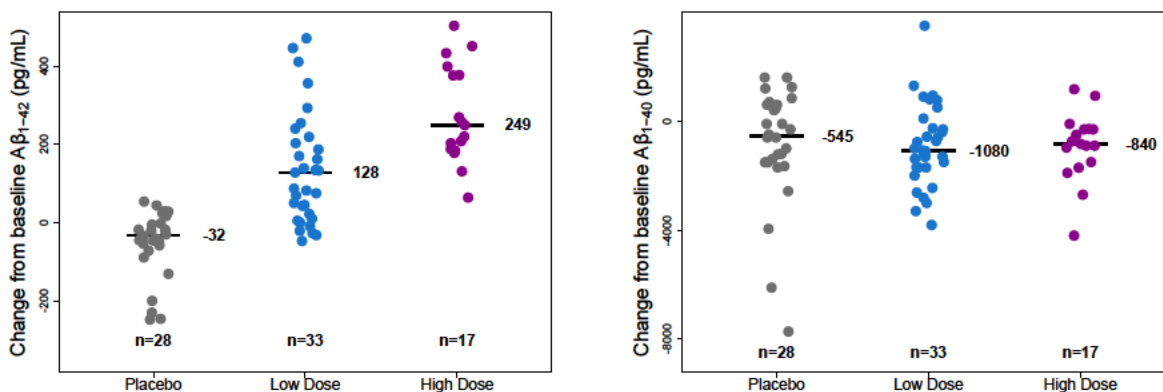


Created by reviewer from Table 23 in Study 302 CSR

There was a dose-related increase in CSF A β_{1-42} levels with overlap between groups, but no apparent relationship between dose and CSF A β_{1-40} levels (Figure 4). The change in CSF A β_{1-42}

levels may reflect an increased input of $A\beta_{1-42}$ into the CSF, decreased clearance from the CSF due to the prolonged half-life of the aducanumab: $A\beta_{1-42}$ complex, or dispersion out of aggregated forms.

Figure 4: Study 302 Change from Baseline in CSF $A\beta_{1-42}$ (left) and $A\beta_{1-40}$ (right) at Week 78

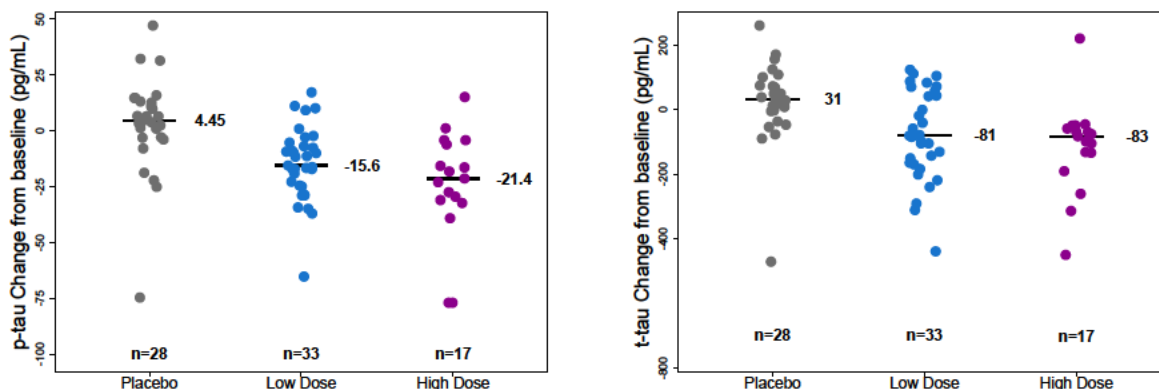


Created by reviewer using adbk.xpt

Change from baseline in markers of downstream Alzheimer's Tau pathophysiology and neurodegeneration are illustrated in Figure 5 and suggest a dose-dependent decrease in both p-Tau and t-Tau with significant variability between patients following aducanumab treatment. There were no differences between either treatment group and placebo for change from baseline at Week 78 in MRI volume measures for whole brain, whole cortex and hippocampus. There was an increase in change from baseline to Week 78 lateral ventricle volume in both dose groups.

A pooled analysis of Studies 301 and 302 was performed for tau PET and is further described in Section 5.2.2.

Figure 5: Study 302 Change from Baseline in CSF p-Tau (left) and t-Tau (right) at Week 78



Created by reviewer using adbk.xpt

Additional Analyses Conducted on the Individual Trial

Correlation analyses of CDR-SB and brain A β plaque levels were performed for the ^{18}F -florbetapir amyloid PET analysis population. A total of 329 patients were included in the analyses. A positive, but relatively weak (Spearman correlation of 0.19 [95% CI: 0.048, 0.327]) relationship was observed between change from baseline in PET composite SUVR at Week 78 and change from baseline in CDR-SB Week 78.

Reviewer Comment: It has been recognized that brain A β burden is not well correlated with disease progression and there is not an a priori expectation of precisely how reduction in A β burden would relate to changes on cognitive scales. It is not even clear if the key metric is a relative reduction in SUVR or an absolute reduction below a certain level. The correlation presented by the applicant is a snapshot of Week 78 data and does not take into account the dynamic nature of A β reduction or the potential for downstream effects, or delays between reduction in A β and changes in CDR-SB. The correlation analyses were performed in a subgroup of the overall population (~30-35%) and do not account for other baseline factors which may affect the relationship. Furthermore, the 25th to 75th percentiles of the change from baseline of CDR-SB at Week 78 in the overall population are 0 and 2.5, respectively, making it difficult to resolve relationships over this narrow range, especially in small subgroups with limited range of change in SUVR. For these reasons, it is not surprising that despite a treatment effect, there is not a tight correlation between reduction in brain A β plaque levels and CDR-SB. Notably, an exposure-response model which considers individual dosing and the time course of SUVR was able to detect a relationship between SUVR and CDR-SB.

5.2. Study 301 (221AD301): A Phase 3 Multicenter, Double-Blind, Placebo-Controlled, Parallel-Group Study to Evaluate the Efficacy and Safety of Aducanumab (BIIB037) in Subjects With Early Alzheimer’s Disease

5.2.1. Study Design

Study 301 was identically designed to Study 302. At the time of the futility analysis on December 26, 2018, 57% of patients from Study 301 had the opportunity to complete the Week 78 visit. Please refer to Section 5.1.1 for study design elements as these are applicable to Study 301.

5.2.2. Study Results

Compliance with Good Clinical Practices

The applicant asserts that the study was performed in accordance with 21 CFR parts 50, 54, 56 and 312 Subpart D, ICH Guideline on GCP (E6) and the ethical principles outlined in the Declaration of Helsinki.

Financial Disclosure

The applicant has adequately disclosed financial interests or agreements with clinical investigators as outlined in the guidance for industry Financial Disclosures by Clinical Investigators.

Patient Disposition

A total of 6173 patients were screened for the study and 1653 patients were randomized. The most common reasons reported for screening failure were having a CDR global score, MMSE score, or RBANS score outside the allowed range (67%) or not having a positive amyloid PET scan (12%). There were 6 patients who were randomized but did not receive study treatment, with the most common reason being that the patient should have been classified as a screening failure. There were also 5 patients randomized to placebo who received aducanumab treatment in the visit before the start of the LTE. Patient disposition in the study is summarized in Table 12. The most common reason noted for treatment or study discontinuation is “other.” This category includes patients who terminated the study due to the administrative decision to terminate aducanumab studies based on the interim futility analysis (placebo, 166 (30.5%); low dose, 160 (29.3%); high dose, 190 (34.2%)).

Table 12: Study 301 Patient Disposition

Disposition	Study 301
No. of patients screened	6173

No. of patients not randomized	4520		
	Aducanumab Low Dose N=547 n (%)	Aducanumab High Dose N=555 n (%)	Placebo N=545 n (%)
Patients randomized	549	556	548
ITT population	547 (100%)	555 (100%)	545 (100%)
Per-protocol population	398 (72.8%)	378 (68.1%)	435 (79.8%)
OTC population	370 (67.6%)	345 (62.2%)	369 (67.7%)
¹⁸ F-florbetapir amyloid PET	198 (36.2%)	183 (33.0%)	204 (37.4%)
Discontinued treatment	221 (40.4%)	280 (50.5%)	219 (40.2%)
Adverse event	45 (8.2%)	63 (11.4%)	28 (5.1%)
Consent withdrawn	9 (1.6%)	16 (2.9%)	14 (2.6%)
Other reasons	167 (30.6%)	201 (36.2%)	177 (32.5%)
Discontinued study	222 (40.6%)	267 (48.1%)	220 (40.4%)
Adverse event	25 (4.6%)	28 (5.0%)	16 (2.9%)
Consent withdrawn	16 (2.9%)	27 (4.9%)	26 (4.8%)
Other reasons	181 (33.1%)	212 (39.2%)	178 (32.7%)

Created by the reviewer using ie.xpt and Tables 13, 20 and 50 in Study 301 CSR.

Reviewer Comment: Patient disposition in Study 301 was largely similar to that in Study 302. Relatively fewer patients in Study 301 discontinued the study because it began before Study 302 and therefore more patients had the opportunity to complete the study at the time of study termination.

Protocol Violations/Deviations

Most patients (60.8%) had at least one major protocol deviation during the placebo-controlled period, but the incidence was similar across treatment groups. The most common categories of major protocol deviations were regarding informed consent (25.9%), study procedures (24.4%), and investigational product compliance (18.0%). These deviations are not expected to affect interpretation of the overall results. During the combined screening and placebo-controlled periods, 83 patients (35 placebo, 21 low dose, and 27 high dose) had a major protocol deviation for eligibility or entry criteria. There were 3 patients who did not meet the CDR global score criterion, 5 who did not meet the MMSE criterion and 13 who did not meet the RBANS criterion. These numbers are small compared to the overall population and balanced across the treatment arms and are therefore not expected to affect the efficacy analyses. A total of 10.9% of patients received disallowed concomitant medications during the placebo-controlled period, but many of these were anti-coagulants or corticosteroids, disallowed because of a possible role in the event of ARIA and were not anticipated to have an impact on the effectiveness of aducanumab. Also, seven patients were randomized to a carrier status that was discordant with laboratory assay results. There were also 5 patients randomized to placebo who received aducanumab treatment in the visit before the start of the LTE. Single dose administration of aducanumab at the end of the study would not have an impact on the efficacy results.

Tables of Demographic and Disease Characteristics

Table 13 contains information regarding demographic characteristics for each treatment arm in the ITT population. Demographic characteristics were well-balanced across the treatment arms and generally representative of the patient population except for an under-representation of African American patients. Overall, 46.3% of patients were enrolled in the United States.

Table 13: Study 301 Baseline Demographics (ITT Population)

Demographic Parameters	Placebo (N=545) n (%)	Treatment Group		
		Aducanumab Low Dose (N=547) n (%)	Aducanumab High Dose (N=555) n (%)	Total (N=1647) n (%)
Sex				
Male	258 (47.3%)	263 (48.1%)	263 (47.4%)	784 (47.6%)
Female	287 (52.7%)	284 (51.9%)	292 (52.6%)	863 (52.4%)
Age				
Mean years (SD)	69.8 (7.7)	70.4 (7.0)	70.0 (7.7)	70.1 (7.5)
Median (years)	70.0	71.0	71.0	71.0
Min, max (years)	50, 85	51, 85	50, 85	50, 85
Age Group				
≤ 64 years	130 (23.9%)	116 (21.2 %)	124 (22.3%)	370 (22.5%)
> 64 – <75 years	246 (45.1%)	264 (48.3%)	255 (45.9%)	765 (46.4%)
≥ 75 years	169 (31.0%)	167 (30.5%)	176 (31.7%)	512 (31.1%)
Race				
White	413 (75.8%)	412 (75.3%)	413 (74.4%)	1238 (75.2%)
Black or African American	5 (0.9%)	1 (0.2%)	2 (0.4%)	8 (0.5%)
Asian	55 (10.1%)	55 (10.1%)	65 (11.7%)	175 (10.6%)
American Indian or Alaska Native	0	0	0	0
Native Hawaiian or Other Pacific Islander	0	1 (0.2%)	0	1 (<0.1%)
Other	3 (0.6%)	4 (0.7%)	3 (0.5%)	10 (0.6%)
Ethnicity				
Hispanic or Latino	13 (2.4%)	11 (2.0%)	13 (2.3%)	37 (2.2%)
Not Hispanic or Latino	489 (89.7%)	492 (89.9%)	499 (89.9%)	1480 (89.9%)
Not Reported ¹	43 (7.9%)	44 (8.0%)	43 (7.7%)	130 (7.9%)
Region				
United States	251 (46.1%)	260 (47.5%)	252 (45.4%)	763 (46.3%)
Rest of the World				
Europe/Canada/Australia	242 (44.4%)	236 (43.1%)	243 (43.8%)	721 (43.8%)
Asia	52 (9.5%)	51 (9.3%)	60 (10.8%)	163 (9.9%)

Source: Table 14 and Table 49 in Study 301 CSR

¹ Data on race and/or ethnicity were not collected because of local regulations.

Reviewer Comment: Any differences in the demographics between Studies 301 and 302 were minor and not expected to explain differences in study outcomes. Study 301 enrolled more patients in the United States (46.3%) than Study 302 (39.8%). The most evident difference in geographic enrollment was that only Study 302 included patients from Poland (12% of the total

study population at 14 sites). The estimated treatment effect for patients in Poland favored placebo over aducanumab and thus does not contribute to the discordant results in the two studies.

Other Baseline Characteristics (disease characteristics and important concomitant drugs)

Table 14 contains a summary of baseline disease characteristics and baseline use of concomitant Alzheimer's disease medications. The disease characteristics are well-balanced across treatment groups and reflect a population of patients who are early in the course of Alzheimer's disease. By design, the overall population consisted of 80% of patients with MCI due to Alzheimer's disease and 20% with mild Alzheimer's disease dementia. The percentage of the population who were ApoE ε4 carriers is consistent with previous reports. Most patients were receiving concomitant medications for Alzheimer's disease (56.4%). Additionally, 13% of patients received any Alzheimer's disease medication and stopped prior to entering the study.

Table 14: Study 301 Baseline Disease Characteristics (ITT Population)

Disease Characteristics	Placebo (N=545) n (%)	Treatment Group		
		Aducanumab Low Dose (N=547) n (%)	Aducanumab High Dose (N=555) n (%)	Total (N=1647) n (%)
Baseline Clinical Stage				
MCI due to AD	443 (81.3%)	440 (80.4%)	442 (79.6%)	1325 (80.4%)
Mild AD	102 (18.7%)	107 (19.6%)	113 (20.4%)	322 (19.6%)
Laboratory ApoE ε4 Status				
Carrier	376 (69.0)	391 (71.5%)	378 (68.1%)	1145 (69.5%)
Homozygote	104 (19.1%)	101 (18.5%)	104 (18.7%)	309 (18.8%)
Heterozygote	272 (49.9%)	290 (53.0%)	274 (49.4%)	836 (50.8%)
Non-carrier	167 (30.6%)	156 (28.5%)	176 (31.7%)	499 (30.3%)
Undetermined	2 (0.4%)	0	1 (0.2%)	3 (0.2%)
Number of Years of Formal Education				
Mean years (SD)	14.7 (3.7)	14.6 (3.8)	14.6 (3.7)	14.6 (3.7)
Median (years)	16.0	15.0	15.0	15.0
Min, Max (years)	4, 26	3, 27	4, 27	3, 27
Number of Years Since Diagnosis of AD				
Mean years (SD)	543	546	555	1644
Median (years)	1.1 (1.2)	1.3 (1.6)	1.2 (1.4)	1.2 (1.4)
Min, Max (years)	0.7	0.7	0.8	0.7
Min, Max (years)	0.0, 11.9	0.0, 21.6	0.0, 17.6	0.0, 21.6
Concomitant AD medication				
Any AD medication at	299 (54.9%)	317 (58.0%)	313 (56.4%)	929 (56.4%)

baseline				
Only cholinesterase inhibitors	242 (44.4%)	257 (47.0%)	264 (47.6%)	763 (46.3%)
Only memantine	16 (2.9%)	15 (2.7%)	13 (2.3%)	44 (2.7%)
Both cholinesterase inhibitors and memantine	41 (7.5%)	45 (8.2%)	36 (6.5%)	122 (7.4%)
Baseline CDR-SB	545	547	554	1646
Mean (SD)	2.40 (1.01)	2.43 (1.01)	2.40 (1.01)	2.41 (1.01)
Median	2.50	2.50	2.50	2.50
Min, Max	0.5, 7.0	0.5, 8.0	0.5, 5.5	0.5, 8.0
Baseline MMSE	545	547	555	1647
<24	3 (0.6%)	1 (0.2%)	1 (0.2%)	5 (0.3%)
≥24 - <27	284 (52.1 %)	282 (51.6%)	302 (54.4%)	868 (52.7%)
≥27 - ≤30	258 (47.3%)	264 (48.3%)	252 (45.4%)	774 (47.0%)
Baseline ADAS-Cog 13	542	547	553	1642
Mean (SD)	22.5 (6.6)	22.5 (6.3)	22.4 (6.5)	22.5 (6.5)
Median	22.2	22.3	22.0	22.0
Min, Max	4.7, 48.0	8.0, 41.7	6.0, 43.3	4.7, 48.0
Baseline ADCS-ADL-MCI	541	546	553	1640
Mean (SD)	43.0 (5.6)	42.9 (5.7)	42.9 (5.7)	42.9 (5.7)
Median	44.0	44.0	44.0	44.0
Min, Max	22, 53	21, 53	12, 53	12, 53

Source: Table 16, 17 and Table 59 in Study 301 CSR

Reviewer Comment: Any differences in the baseline disease characteristics between Studies 301 and 302 were minor and not expected to explain differences in study outcomes.

Treatment Compliance, Concomitant Medications, and Rescue Medication Use

The mean percentage of study treatment received was 96.4%. Although compliance with study treatment was high, protocol-defined dose modification for ARIA, changes to the protocol, and the termination of the study meant that some patients may have received fewer target doses. For example, only 6.1% of ApoE ε4 carriers received all 14 doses of aducanumab 10 mg/kg compared to 33.5% of ApoE ε4 non-carriers. Similarly, 80.2% of ApoE ε4 carriers received any 10 mg/kg doses, compared to 88.6% of ApoE ε4 non-carriers. More patients in the lower dose treatment arm received the maximum dose level (91.7% for ApoE ε4 carriers receiving 6 mg/kg and 99.5% for ApoE ε4 non-carriers receiving 3 mg/kg).

Overall, 11.7% of patients made changes to their concomitant Alzheimer's disease medications at some point during the study. The number of patients who made changes was similar across treatment arms (placebo, 68 (12.5%); low dose, 66 (12.1%); high dose, 58 (10.5%)).

Reviewer Comment: Because of the timing of the studies, the implementation of Protocol Version 3 and 4 occurred earlier in Study 302 than Study 301. Therefore, patients in the high-dose treatment arm in Study 302 were more likely to receive a full treatment course of 10 mg/kg.

Efficacy Results – Primary Endpoint

The primary efficacy endpoint analysis, change from baseline in CDR-SB at Week 78, did not show a difference in the aducanumab high-dose treatment arm compared to placebo (0.03, 2%, $p = 0.8330$). A numeric difference of -12% in favor of the aducanumab low-dose treatment arm compared to placebo was observed (Table 15). Study 301 was a negative study, so sensitivity analyses were not performed, and subgroup analyses are not reported here.

Table 15: Study 301 Primary Endpoint Analysis

	Placebo (N=545)	Aducanumab Low Dose (N=547)	Aducanumab High Dose (N=555)
Baseline CDR-SB			
n	545	547	554
Mean	2.40	2.43	2.40
Change from Baseline in CDR-SB at Week 78			
n	333	331	295
Adjusted mean	1.56	1.38	1.59
Standard error	0.108	0.108	0.111
95% CI	(1.344, 1.768)	(1.164, 1.590)	(1.370, 1.805)
Difference from placebo		-0.18	0.03
95% CI for difference		(-0.469, 0.110)	(-0.262, 0.326)
% difference vs. placebo		-12%	2%
p-value (compared with placebo)		0.2250	0.8330

Source: Table 22 from Study 301 CSR

Reviewer Comment: The placebo decline in Study 301 (1.56) was smaller than in Study 302 (1.74). The treatment differences for the low dose, although not statistically significant, were remarkably consistent in the two studies and suggest the outcomes between the two studies are only partially discordant, namely for the high-dose treatment arms.

Table 16 displays the results of the supplementary analyses of the primary endpoint in different populations as described in Section 5.1.1. Results in the uncensored population with the most data provide the most favorable results for aducanumab but are still negative.

Table 16: Study 301 Primary Endpoint Analysis (Dataset Sensitivity)

ITT Population during Double-Blind Period				Uncensored ITT Population			OTC Population		
Difference vs. Placebo at Week 78 (%) p-value				Difference vs. Placebo at Week 78 (%) p-value			Difference vs. Placebo at Week 78 (%) p-value		
	PBO Decline (N=545)	Low Dose (N=547)	High Dose (N=555)	PBO Decline (N=545)	Low Dose (N=547)	High Dose (N=555)	PBO Decline (N=369)	Low Dose (N=370)	High Dose (N=345)
CDR-SB	n=369	n=362	n=335	n=414	n=421	n=398	n=332	n=331	n=293
	1.57	-0.17	-0.02	1.60	-0.20	-0.08	1.46	-0.12	0.08
		-11%	-1%		-13%	-5%		-8%	5%
		0.2351	0.8719		0.1530	0.5851		0.4527	0.6326

Source: Tables 74, 76 and 78 in Study 301 CSR

Data Quality and Integrity

There were no major data quality issues identified during the review of Study 301.

Efficacy Results – Secondary endpoints

Although Study 301 was a negative study, a summary of the analysis results for the secondary endpoints is provided in Table 17 for comparison to results from Study 302.

Table 17: Study 301 Secondary Endpoint Analysis

	Placebo (N=545)	Aducanumab Low Dose (N=547)	Aducanumab High Dose (N=555)
Baseline MMSE			
n	545	547	554
Mean	26.4	26.4	26.4
Change from Baseline in MMSE at Week 78			
n	322	334	297
Adjusted mean	-3.5	-3.3	-3.6
Standard error	0.21	0.21	0.21
95% CI	(-3.94, -3.13)	(-3.75, -2.93)	(-4.02, -3.19)
Difference from placebo		0.2	-0.1
95% CI for difference		(-0.35, 0.74)	(-0.62, 0.49)
% difference vs. placebo		-6%	3%
p-value (compared with placebo)		0.4795	0.8106
Baseline ADAS-Cog 13			
n	542	547	553
Mean	22.48	22.52	22.40

Change from Baseline in ADAS-Cog 13 at Week 78			
n	331	332	294
Adjusted mean	5.14	4.56	4.55
Standard error	0.38	0.38	0.39
95% CI	(4.40, 5.88)	(3.82, 5.30)	(3.79, 5.31)
Difference from placebo		-0.58	-0.59
95% CI for difference		(-1.58, 0.42)	(-1.61, 0.43)
% difference vs. placebo		-11%	-11%
p-value (compared with placebo)		0.2536	0.2578
Baseline ADCS-ADL-MCI			
n	541	546	553
Mean	43.0	42.9	42.9
Change from Baseline in ADCS-ADL-MCI at Week 78			
n	331	330	298
Adjusted mean	-3.8	-3.1	-3.1
Standard error	0.35	0.35	0.35
95% CI	(-4.48, -3.12)	(-3.76, -2.39)	(-3.81, -2.42)
Difference from placebo		0.7	0.7
95% CI for difference		(-0.19, 1.64)	(-0.25, 1.61)
% difference vs. placebo		-18%	-18%
p-value (compared with placebo)		0.1225	0.1506

Sources: Tables 92, 104 and 116 in Study 301 CSR

Reviewer Comment: Results for the secondary endpoints for the low-dose treatment arms in Studies 301 and 302 are remarkably similar and suggest a numerical difference in favor of aducanumab, further suggesting that the outcomes between the two studies are only partially discordant, namely for the high-dose treatment arms.

Efficacy Results – Tertiary endpoint

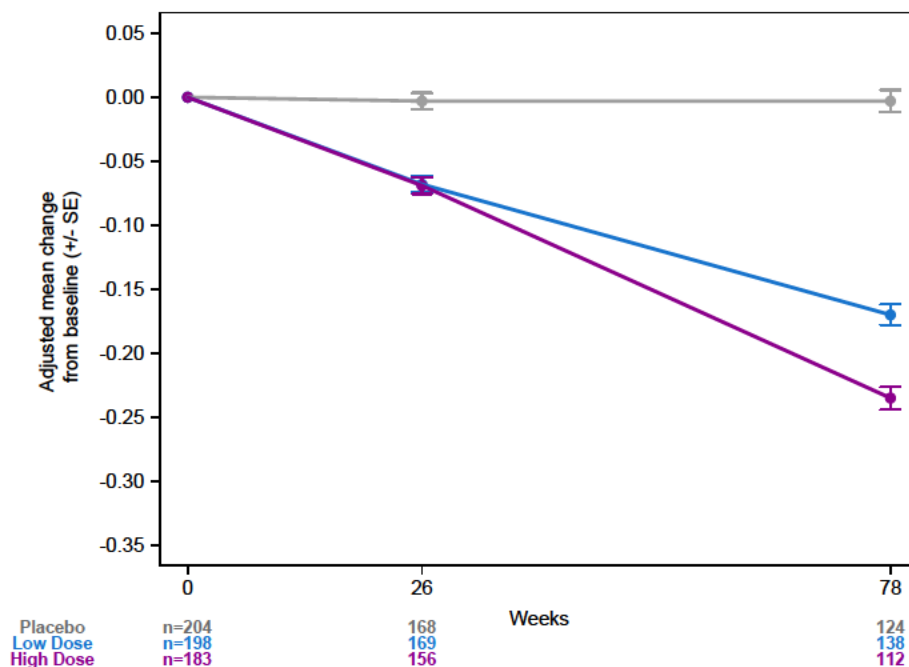
The adjusted mean change from baseline in NPI-10 at Week 78 compared to placebo was 0.1 (8%, $p = 0.9071$) for aducanumab high dose and -1.0 (-83%, nominal $p=0.0460$) for aducanumab low dose.

Efficacy Results – Pharmacodynamic endpoints

Change from baseline in brain amyloid signal as measured by SUVR was analyzed with an MMRM model with fixed effects of treatment group, visit, treatment group-by-visit interaction, baseline SUVR, baseline SUVR-by-visit interaction, baseline MMSE, ApoE $\epsilon 4$ status, and baseline age. The adjusted mean change from baseline relative to placebo was similar in the low-dose and high-dose groups at Week 26, -0.065 and -0.066, respectively due to similar dosing during titration. At Week 78, the adjusted mean change from placebo was -0.167 for the low dose and -0.232 for the high dose, indicating time- and dose-dependent relationships (Figure 6).

Consistent relationships were found for other brain regions and when using additional reference regions. No correlation was observed between change from baseline CDR-SB at Week 78 and change from baseline SUVR at Week 78.

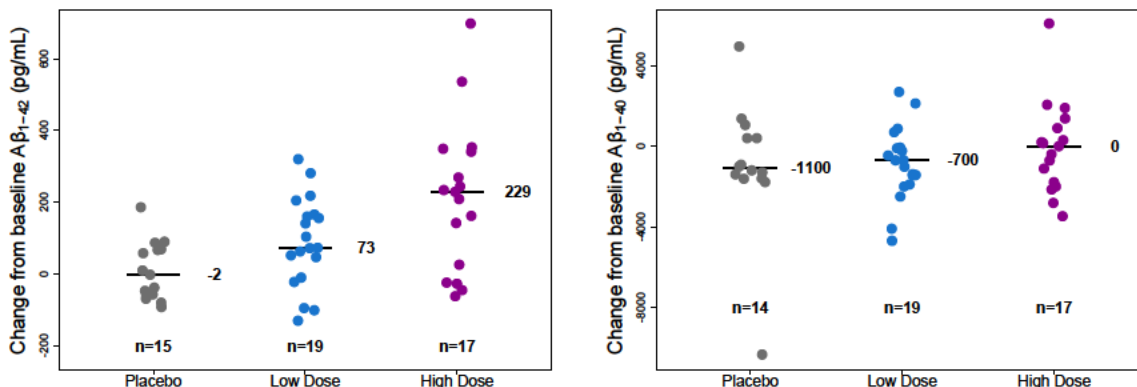
Figure 6: Study 301 Change from Baseline in A β PET Composite SUVR



Created by reviewer from Table 23 in Study 301 CSR

There was an apparent dose-related increase in CSF A β_{1-42} levels, but some overlap between groups (Figure 7). The high-dose treatment arm exhibited no median change in A β_{1-40} compared to decreases in the placebo and low-dose arms, but there was considerable overlap between groups.

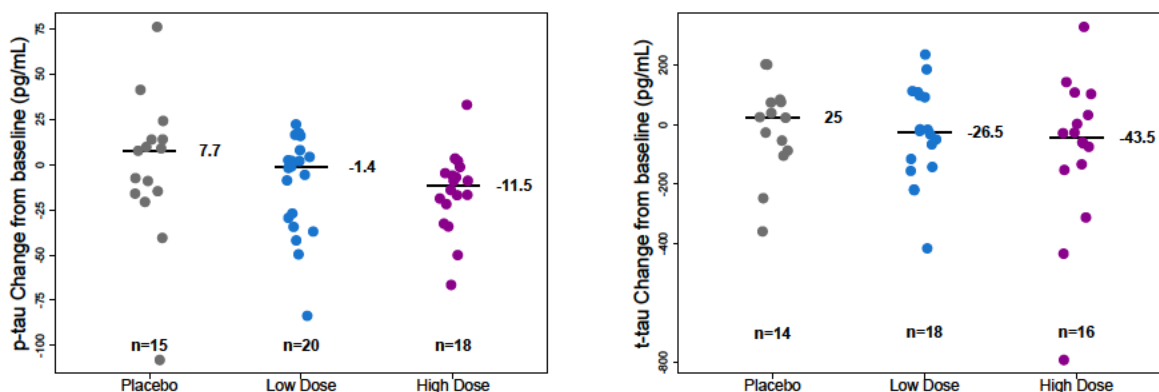
Figure 7: Study 301 Change from Baseline in CSF $A\beta_{1-42}$ (left) and $A\beta_{1-40}$ (right) at Week 78



Created by reviewer using adbk.xpt

Change from baseline in markers of downstream Alzheimer's Tau pathophysiology and neurodegeneration are illustrated in Figure 8. Dose-dependent trends were consistent with those observed in Study 302, but with smaller decreases following aducanumab treatment. There were no differences between either treatment group and placebo for change from baseline at Week 78 in MRI volume measures for whole brain, whole cortex and hippocampus. There was an increase in change from baseline to Week 78 lateral ventricle volume in both dose groups.

Figure 8: Study 301 Change from Baseline in CSF p-Tau (left) and t-Tau (right) at Week 78

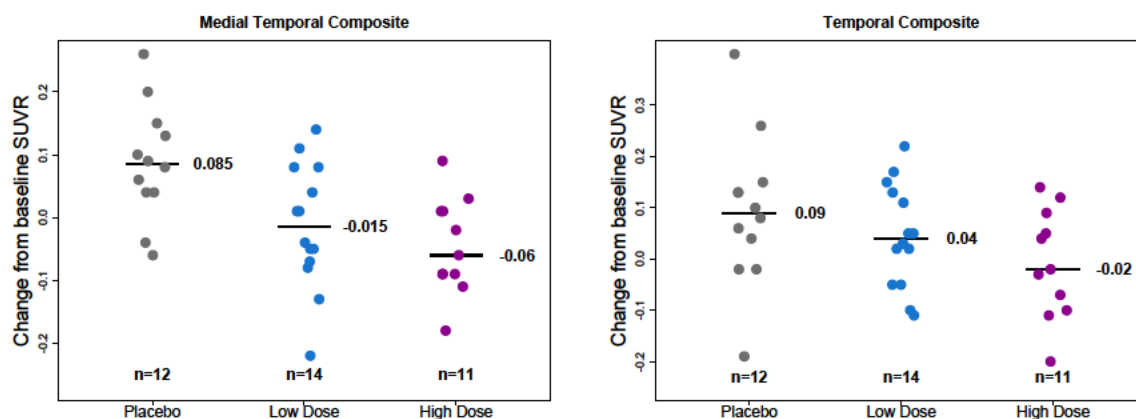


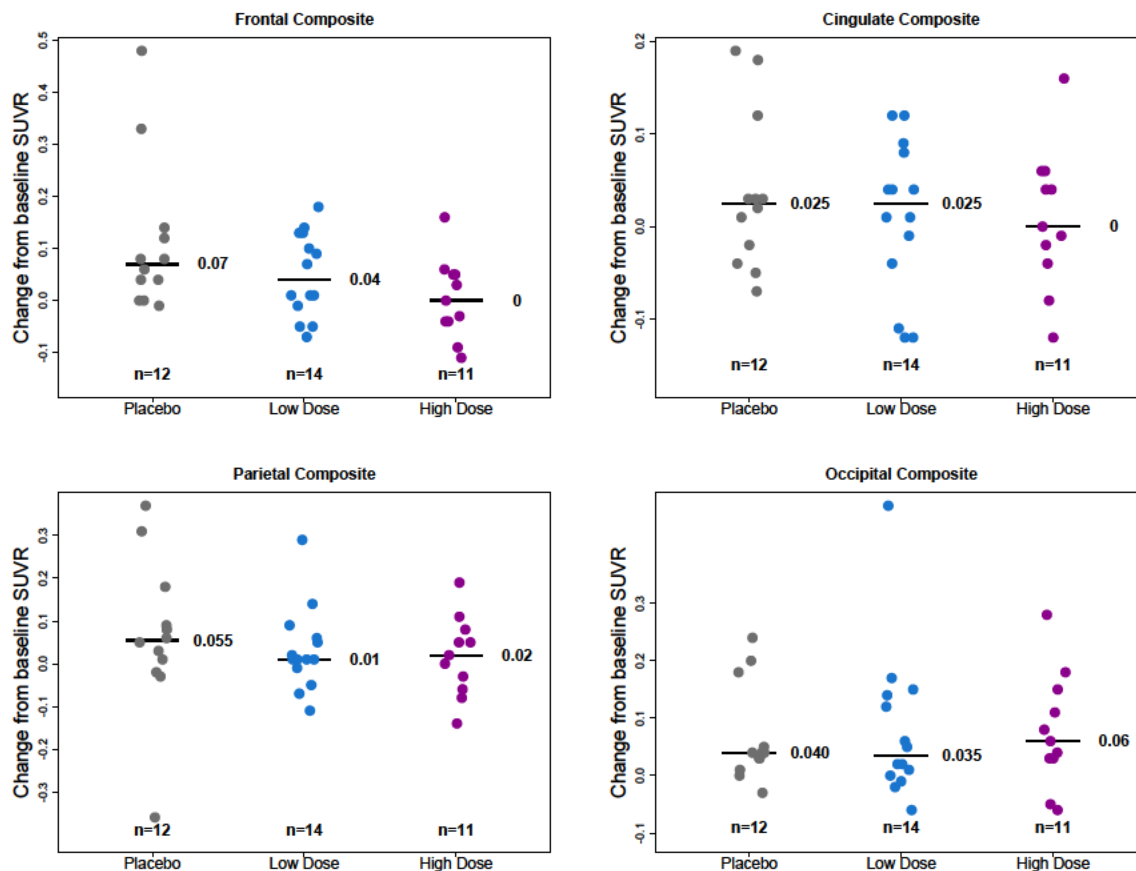
Created by reviewer using adbk.xpt

Reviewer Comment: The change from baseline in SUVR at Week 78 was similar for the low-dose arms in Studies 301 and 302, but lower for the high dose in Study 301 (-0.232) than in Study 302 (-0.278). This is consistent with the expectation that patients in Study 301, especially those enrolled early in the studies as in the PET subpopulation, achieved lower aducanumab exposure than patients in Study 302. Similar trends are observed for CSF biomarkers, but it is more difficult to draw conclusions because of the limited sample size and considerable variability.

A pooled analysis of Studies 301 and 302 was performed for tau PET measurements because the number of patients with tau PET imaging was small (31 in Study 301 and 6 in Study 302). Measurement of tau PET was introduced relatively later in the studies and therefore this subpopulation included more patients in the high-dose arms who had the opportunity to receive the 10 mg/kg dose. The observations for different brain regions are illustrated in Figure 9. Post-baseline samples were collected earlier in some patients (mean visit time of 13.6 months) before some received the full course of aducanumab treatment. For example, the mean cumulative dose for the pooled high-dose group for tau PET data was 86 mg/kg compared to 122 mg/kg for the high dose group of patients in Study 301 who contributed p-Tau data. Dose-dependent reductions in brain tau SUVR were observed in the medial temporal, temporal, and frontal composite regions with trends favoring active treatment for the cingulate and parietal composite brain regions.

Figure 9: Study 301 and Study 302 Pooled Analyses: Change from Baseline in Tau PET





Created by reviewer using adtau.xpt

5.3. Study 103 (221AD103): A Randomized, Double-Blinded, Placebo-Controlled Multiple Dose Study to Assess the Safety, Tolerability, Pharmacokinetics, and Pharmacodynamics of BIIB037 in Subjects with Prodromal or Mild Alzheimer's Disease

5.3.1. Study Design

Overview and Objective

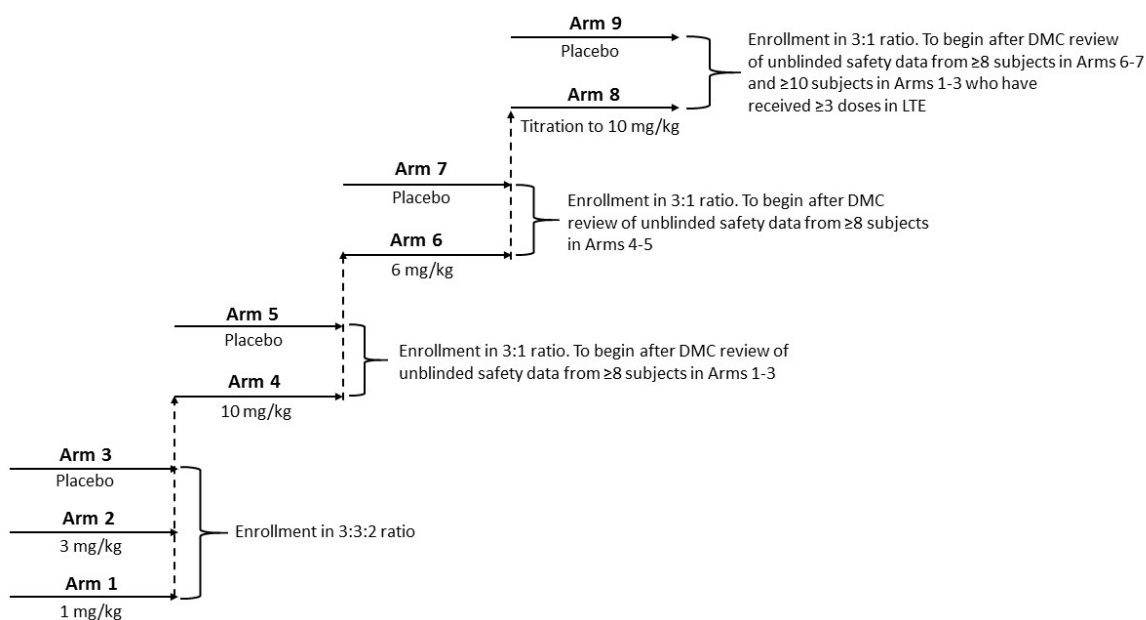
Study 103 was the first multiple-dose study of aducanumab in patients with Alzheimer's disease. The primary objective of the study was to evaluate the safety and tolerability of multiple doses of aducanumab in subjects with prodromal or mild Alzheimer's disease. Although described as a safety and tolerability study, Study 103 was a randomized, double-blind, multicenter study which incorporated many of the elements of Study 301 and Study 302, including efficacy and pharmacodynamic assessments in a similar patient population.

Trial Design

Study Design

Study 103 was a multicenter, randomized, double-blind, placebo-controlled, staggered, parallel-group study of aducanumab in patients with prodromal or mild Alzheimer's disease (i.e., mild Alzheimer's disease dementia). The study was conducted in 27 sites in the United States. A schematic of the staggered, parallel-group design of the study is illustrated in Figure 10. Initially, Arms 1-3 were enrolled in parallel. Upon review of ongoing unblinded safety data in a subset of patients enrolled in Arms 1-3, Arms 4-5 were enrolled in parallel, then Arms 6-7, and then Arms 8-9. Patients were randomized into each treatment group within Arms 1-3, Arms 4-5, Arms 6-7 and Arms 8-9, with each cohort including a placebo arm. Randomization was stratified by ApoE ϵ 4 status (carrier or non-carrier) except for Arms 8-9 which only included ApoE ϵ 4 carriers to test the hypothesis that titration would lower the incidence of ARIA in this population.

Figure 10: Study 103 Schematic



Created by reviewer, modified from Figure 2 in Study 103 protocol

Note: Safety data from the single ascending dose study (Study 101) was also be considered for dose escalation.

The study included an 8-week screening period, a 52-week placebo-controlled treatment period and a safety follow-up period of 18 weeks after the final dose. Patients who completed the placebo-controlled period had the option to enter a dose-blind long-term extension period.

Diagnostic Criteria

Patients with mild Alzheimer's disease fulfilled the clinical criteria as defined by the 2011 NIA-AA framework. The clinical criteria for prodromal patients were consistent with those described for MCI due to Alzheimer's disease in the 2011 NIA-AA framework. Patients were also required to have evidence of brain A β pathology by visual read of a PET scan.

Key Inclusion Criteria

1. Male or female patients age 50 to ≤ 90 years
2. Positive amyloid PET scan
3. Must consent to ApoE genotyping
4. Must meet the following core clinical criteria:
 - Prodromal Alzheimer's disease
 - MMSE ≥ 24
 - A spontaneous memory complaint
 - Free recall score ≤ 27 on the Free and Cued Selective Reminding Test (FCSRT)
 - CDR global score of 0.5
 - Absence of significant levels of impairment in other cognitive domains
 - Essentially preserved activities of daily living and an absence of dementia
 - Mild Alzheimer's disease
 - MMSE between 20-26, inclusive
 - CDR global score of 0.5 or 1.0
 - Meet NIA-AA core clinical criteria for probable Alzheimer's disease
5. Has one informant/care partner able to provide accurate information about the subject's cognitive and functional abilities and should be available for the duration of the study

Key Exclusion Criteria

1. Any medical or neurological condition (other than Alzheimer's disease) that may be a contributing cause of the subject's cognitive impairment
2. Any medication that may contribute to cognitive impairment, put the subject at higher risk for AEs, or impair the subject's ability to perform cognitive testing
3. Clinically significant unstable psychiatric illness within 6 months prior to screening
4. Brain MRI performed at screening that shows evidence of any of the following: acute or sub-acute hemorrhage, prior microhemorrhage or prior subarachnoid hemorrhage (unless finding is not due to an underlying structural or vascular hemorrhage), greater than 4 microhemorrhages, cortical infarct, >1 lacunar infarct, superficial siderosis, or

history of diffuse white matter disease

5. Contraindications to having a brain MRI or PET scan
6. History of bleeding disorder
7. Use of medications with platelet anti-aggregant or anti-coagulant properties (unless aspirin at ≤ 325 mg daily)
8. Uncontrolled hypertension or history of unstable angina, myocardial infarction, chronic heart failure, or clinically significant conduction abnormalities
9. Participation in any active immunotherapy study targeting A β , any passive immunotherapy study targeting A β within 48 weeks of screening or any study with purported disease-modifying effect in Alzheimer's disease within 24 weeks of screening

Reviewer Comment: The population enrolled in Studies 301 and 302 is essentially represented within Study 103. Study 103 also includes patients at a later stage of the disease continuum. Specifically, the entry criteria for Study 103 allows for enrollment of patients with an MMSE as low as 20 (as compared to 24 in Studies 301 and 302) and a CDR global score of 0.5 or 1 (as compared to only 0.5 in Studies 301 and 302).

Dose Selection

Dosing was originally based on results from nonclinical studies, the PK/PD relationship of mouse chimeric aducanumab on brain amyloid in mice, safety and PK data from the single ascending dose study in subjects with mild to moderate Alzheimer's disease, and the projected PK/PD relationship in humans. The intended doses were 1 mg/kg, 3 mg/kg, up to 10 mg/kg and up to 30 mg/kg. Based on review of ongoing safety data and IDMC recommendations, the actual doses studied were fixed doses of 1 mg/kg, 3 mg/kg, 6 mg/kg and 10 mg/kg. An additional arm investigating titration to 10 mg/kg was added to test the hypothesis that titration could mitigate the incidence and severity of ARIA.

Study Treatments

IV infusions of aducanumab or placebo were administered approximately every 4 weeks over 52 weeks for a total of 14 doses. Fixed doses were administered in Arms 1-7 (Figure 10). Titration in Arm 8 consisted of 1 mg/kg for the first 2 doses, 3 mg/kg for the next 4 doses, 6 mg/kg for the next 5 doses and 10 mg/kg thereafter (3 doses) for a 44-week titration period.

Reviewer Comment: The 10 mg/kg titration regimen in Study 103 is much slower (44 weeks) than the titration regimen in Studies 301 and 302 (24 weeks). The 10 mg/kg fixed-dose arm (14 doses of 10 mg/kg over 54 weeks) is the relevant arm for comparisons to the high dose regimen in Studies 301 and 302 (14 doses of 10 mg/kg over 78 weeks).

Assignment to Treatment

IxRS was used to manage randomization and treatment assignment. The randomization ratios for different arms are included in Figure 10. Randomization was stratified by ApoE ε4 status (carrier or non-carrier) except for Arms 8-9 which included only ApoE ε4 carriers. The ratio of ApoE ε4 carriers was to be no more than 2:1 and no less than 1:2.

Blinding

An unblinded pharmacist managed all aspects of study treatment receipt, dispensing, and preparation. All other study site staff and patients were blinded to treatment assignment during the placebo-controlled period. The rater who conducted the CDR for a patient was not allowed to complete any other rating scales for the same patient. For the LTE period, patients were to remain blinded to their treatment assignment. Investigators were blinded to whether the patient received a titrated dose of study treatment.

The Biogen study team was blinded until the interim analysis of the Week 26 data. Information from Arms 1-5 was unblinded after all patients in these arms completed evaluation at the Week 26 visit, information from Arms 6-7 was unblinded after all patients in these arms completed evaluation at the Week 26 visit, and information from Arms 8-9 was unblinded after all patients in these arms completed evaluation at the Week 26 visit.

Reviewer Comment: There was no central review of ratings. On the other hand, the rater who conducted CDR was not able to complete other rating assessments.

Dose modification/Dose Discontinuation

Dose modification criteria were established to account for the expected occurrence of ARIA-E or ARIA-H. Dose reduction, suspension or termination were dependent on the radiographic severity of ARIA as detected by MRI, the presence or absence of clinical symptoms, and the severity of symptoms, if present. Dosing guidelines evolved during the study and are summarized in Table 18, Table 19, and Table 20 for ARIA-E, ARIA-H microhemorrhage and ARIA-H superficial siderosis, respectively. The investigator was allowed to choose a more conservative course of action than specified. Patients who developed ARIA-H coincident with ARIA-E were to follow the more restrictive guideline.

Table 18: Dose Modification/Discontinuation Rules for ARIA-E by Protocol Version (Study 103)

Clinical Symptom Severity	MRI Severity	Protocol Version ¹
---------------------------	--------------	-------------------------------

			Version 2 3	Version 4-8	Version 9-11
Asymptomatic		Mild	Suspend dosing and restart at next lower dose	Continue dosing at same dose and schedule	
		Moderate or Severe		Suspend dosing and restart at next lower dose	Suspend dosing and restart at same dose and titrate to target
Symptomatic	Mild	Any	Permanently discontinue	Suspend dosing and restart at next lower dose	
	Moderate			Permanently discontinue	
	Severe/Serious “other medically important event” ²				
	Serious, except “other than medically important event”				Permanently discontinue

Created by reviewer, modified from Table 11 in Study 103 CSR

¹ Participant enrollment began under Protocol Version 2

² "Other medically important events" include SAEs that were not life-threatening, did not require inpatient hospitalization or prolongation of existing hospitalization and did not result in significant or permanent disability

Note: Dosing was restarted when ARIA-E and/or clinical symptoms resolved

Table 19: Dose Modification/Discontinuation Rules for ARIA-H (Microhemorrhage) by Protocol Version (Study 103)

Clinical Symptom Severity		MRI Severity	Protocol Version ¹		
			Version 2 and 3	Version 4-8	Version 9-11
Asymptomatic		Mild	Continue dosing at same dose and schedule		
		Moderate		Permanently Discontinue	Suspend dosing and restart at the same dose
		Severe			Permanently discontinue
Symptomatic	Mild	Mild	Permanently Discontinue	Suspend dosing and restart at next lower dose	Suspend dosing and restart at the same dose
		Moderate		Permanently Discontinue	Permanently discontinue
		Severe			Permanently discontinue
	Moderate/Severe	Mild or Moderate			Suspend dosing and restart at the same dose

		Severe			Permanently discontinue
	Serious “other medically important event” ²	Mild or Moderate			Suspend dosing and restart at the same dose
		Severe			Permanently discontinue
	Serious, except other than medically important event”	Any			

Created by reviewer, modified from Table 12 in Study 103 CSR

¹ Participant enrollment began under Protocol Version 2

² “Other medically important events” include SAEs that were not life-threatening, did not require inpatient hospitalization or prolongation of existing hospitalization and did not result in significant or permanent disability

Note: Dosing was restarted when ARIA-E and/or clinical symptoms resolved

Note: As of Protocol Version 9, mild on MRI = 1 to 4 cumulative microhemorrhages; moderate on MRI = 5 to 9 cumulative microhemorrhages; severe on MRI ≥10 cumulative microhemorrhages

Table 20: Dose Modification/Discontinuation Rules for ARIA-H (Superficial Siderosis) by Protocol Version (Study 103)

Clinical Symptom Severity		MRI Severity	Protocol Version ¹		
			Version 2-4	Version 5-8	Version 9-11
Asymptomatic		Mild	Permanently discontinue	Suspend dosing and restart at the next lower dose	Continue dosing at same dose and schedule
		Moderate		Permanently discontinue	Suspend dosing and restart at same dose
		Severe			Permanently discontinue
Symptomatic	Mild	Mild		Suspend dosing and restart at the next lower dose	Suspend dosing and restart at same dose
		Moderate		Permanently discontinue	Permanently discontinue
		Severe			Permanently discontinue
	Moderate/Severe	Mild or Moderate			Suspend dosing and restart at same dose
		Severe			Permanently discontinue
	Serious “other medically important event” ²	Mild or Moderate			Suspend dosing and restart at same dose
		Severe			

	Serious, except “other than medically important event”	Any			Permanently discontinue
--	---	------------	--	--	----------------------------

Created by reviewer, modified from Table 13 in Study 103 CSR

¹ Participant enrollment began under Protocol Version 2

² “Other medically important events” include SAEs that were not life-threatening, did not require inpatient hospitalization or prolongation of existing hospitalization and did not result in significant or permanent disability

Note: Dosing was restarted when ARIA-E and/or clinical symptoms resolved

Note: As of Protocol Version 9, mild on MRI = 1 cumulative area of superficial siderosis; moderate on MRI = 2 cumulative areas of superficial siderosis; severe on MRI >2 superficial siderosis

Reviewer Comment: ARIA management was more conservative in Study 103 than Studies 301 and 302. Protocol Versions 4 through 8 of Study 103 were generally similar to Version 1 of Studies 301 and 302. ARIA management rules starting in Version 9 of the Study 103 protocol were essentially the same as Protocol Version 3 in Studies 301 and 302.

Administrative Structure

An IDMC was formed to monitor the overall conduct of the study and was the same IDMC used for Studies 301 and 302. At each closed session of the IDMC meeting, unblinded data was reviewed and the IDMC made recommendations to continue, modify, or stop the study. The IDMC played a key role reviewing safety, tolerability, and PK data and recommending dose levels to be studied in subsequent arms.

A centralized imaging laboratory was selected to read and interpret PET and MRI scans. The central laboratory was to notify the PI and the applicant of ARIA-E and ARIA-H findings. For the purposes of study conduct, readings from the central reader prevailed over those from the local radiologist.

The applicant used a CRO (ePharmaSolutions) to provide monitoring of neurocognitive assessments and training and oversight to ensure rating consistency.

Procedures and Schedule

The schedule for key assessments is presented in Table 21. The screening period consisted of four visits within a 60-day period before administration of the first dose of study treatment. Eligible subjects reported to the study site to receive study treatment every 4 weeks for 52 weeks. The final study visit was 18 weeks after the final dose.

Table 21: Study 103 Schedule of Key Assessments

Assessment	Schedule
------------	----------

Eligibility Criteria	Screening V1
ApoE Genotyping	Screening V1
Neurological and Physical Examinations	Screening V1, Day 1, Weeks 4, 8, 12, 22, 40, 52, 70
Study Drug Infusion	Day 1, Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52
Anti-Aducanumab Ab	Day 1, Weeks 4, 8, 12, 20, 28, 36, 44, 52, 70
Aducanumab Concentration (Arms 1-7)	Day 1, Weeks 4, 6, 8, 12, 16, 20, 24, 26, 28, 32, 36, 40, 42, 44, 48, 52, 70
Aducanumab Concentration (Arms 8-9)	Day 1, Weeks 4, 6, 8, 12, 14, 16, 20, 22, 24, 28, 30, 32, 36, 40, 42, 44, 48, 50, 52, 70
CDR	Screening V1, Weeks 26, 54, 70
MMSE	Screening V1, Weeks 24, 52, 70
CSF Collection (optional)	Screening, Week 3, Week 51
Amyloid PET	Screening V3, Weeks 26, 54, 70

Created by reviewer, modified from Table 1 and Table 2 in Study 103 protocol

Note: There were four screening visits (V1, V2, V3 and V4) within 60 days of randomization

Concurrent Medications

Medications for treatment of Alzheimer’s disease were allowed if patients were on a stable dose for 4 weeks prior to and during screening and were to stay on a stable dose while in the study. Medications for chronic diseases were also allowed if patients had been stable on the medication for 4 weeks prior to and during screening. Patients were instructed to continue the medications they were receiving at enrollment and avoid starting any new medications during the study. If a patient’s medication was to be changed during the study, the applicant could have been consulted to determine whether the patient’s study treatment should be suspended.

The following medications were not allowed:

- Medications with platelet anti-aggregant or anti-coagulant properties, except aspirin at doses ≤ 325 mg per day
- Immunosuppressive drugs including systemic corticosteroids.
- Parenteral immunoglobulin, blood products, plasma derivatives, plasma exchange, and plasmapheresis
- Any investigational drug

Subject Completion, Discontinuation or Withdrawal

Patients who completed the 54-week treatment period were considered completers.

Patients who discontinued treatment could remain in the study and continue protocol-defined assessments. The main treatment-related reason for study discontinuation was the occurrence of ARIA as outlined in Table 18, Table 19, and Table 20. Other reasons for discontinuation include withdrawal of consent, medical emergency, and discretion of the investigator for medical reasons or noncompliance. The reason for discontinuation of study treatment was to be recorded in the patient's CRF.

Patients who withdrew from the study after receiving at least one dose of study treatment were encouraged to complete all scheduled safety evaluations or the end-of-treatment visit. Reasons for withdrawal of patients include withdrawal of consent, unwillingness or inability to comply with the protocol and discretion of the investigator or applicant. The reason for withdrawal was to be recorded in the patient's CRF.

Patients who discontinued study treatment or withdrew from the placebo-controlled part of the study may have been replaced. Replacement subjects were to be randomized into a treatment group within the same group (i.e., Arms 1-3, Arms 4-5, Arms 6-7, or Arms 8-9) as the subject who withdrew.

Study Endpoints

The primary endpoint was safety and tolerability as measured by incidence of adverse events and monitored via several clinical assessments. The secondary endpoint, for which the study was powered, was the change from baseline in amyloid signal as measured by ¹⁸F-florbetapir PET at Week 26. The change from baseline in amyloid signal at Week 54 was considered an exploratory endpoint. Notably, change from baseline in CDR-SB and MMSE were also included as exploratory endpoints. For a description of CDR-SB and MMSE, please refer to Section 5.1.1. The rater who conducted the CDR for a given patient was not allowed to complete any other rating scales for that patient. Raters had to be certified before they could perform assessments and sites were encouraged to use the same rater to administer a given test across all visits. Central review of ratings was not performed.

Statistical Analysis Plan

The SAP was finalized on February 7, 2014, and modified on August 5, 2016, to specify the analysis plan for Arms 8-9 and the integrated plan for Arms 1-9 and on May 30, 2018 to specify details for the analyses of the LTE.

Interim Analyses

The SAP specified three interim analyses for the purpose of planning future studies of aducanumab. The three analyses were to be conducted: (1) after all subjects in Arms 1-5 had completed the Week 26 visit, (2) after all subjects in Arms 1-7 had completed the Week 26 visit

and (3) after all subjects in Arms 1-5 had completed the Week 54 visit. The Biogen study team were unblinded at these interim analyses. Data evaluated during these analyses included subject disposition, dosing information, safety data, change from baseline in amyloid signal and change from baseline in CDR and MMSE.

Definitions of Statistical Analysis Populations

The following relevant populations include:

- Pharmacodynamic (PD) Analysis Population – all subjects who were randomized, received at least one dose of study treatment and have both baseline and at least one post-baseline PET SUVR assessment
- Efficacy Analysis Population – all subjects who were randomized, received at least one dose of study treatment and have both baseline and at least one post-baseline CDR or MMSE assessment for at least one scheduled timepoint
- Per-Protocol Population – all subjects who were randomized, received at least one dose of study treatment, had at least 80% study treatment compliance and did not take any disallowed medications

Sample Size Considerations

The study was planned to have approximately 90% power to detect a treatment difference of 1 SD for change from baseline to Week 26 in PET SUVR at a 2-sided significance level of 0.05 and assuming a dropout rate of 20%.

Analysis Method for Clinicals Endpoints (CDR-SB and MMSE)

CDR-SB and MMSE were analyzed by ANCOVA adjusting for their baseline values and ApoE ε4 carrier status at Week 24 and Week 52 separately. A MMRM model was used as a sensitivity analysis. Placebo data were pooled across Arms 3, 5, 7, and 9.

Missing Data

Values for missing data were not imputed for the ANCOVA analysis.

Adjustments for Multiplicity

This study was designed as a safety and tolerability study, so there were no adjustments for multiplicity.

Subgroup Analyses

Subgroup analyses for PET SUVR were performed for ApoE ϵ 4 status (carrier or non-carrier), baseline clinical stage (prodromal or mild AD), and baseline composite SUVR (\leq median or $>$ median). Subgroup analysis for clinical endpoints was not conducted.

Protocol Amendments

Enrollment began under Protocol Version 2, dated April 20, 2012. Many of the protocol amendments specified changes to dose levels to be studied or extended the study duration. Protocol Version 8, dated July 13, 2015, modified the titration regimen in Arm 8 from 1 mg/kg to 6 mg/kg to 1 mg/kg to 10 mg/kg. Table 18, Table 19, and Table 20 outline the changes to the protocol which modified dosing in relation to ARIA management.

5.3.2. Study Results

Compliance with Good Clinical Practices

The applicant asserts that the study was performed in accordance with 21 CFR parts 50, 54, 56, and 312 Subpart D, ICH Guideline on GCP (E6) and the ethical principles outlined in the Declaration of Helsinki.

Financial Disclosure

The applicant has adequately disclosed financial interests or agreements with clinical investigators as outlined in the guidance for industry Financial Disclosures by Clinical Investigators.

Patient Disposition

A total of 197 patients were randomized into the study and 196 received at least one dose of study treatment. One patient was randomized to the aducanumab 3 mg/kg treatment arm but was withdrawn from the study before receiving study treatment because he met the exclusion criterion of uncontrolled hypertension. Patient disposition is summarized in Table 22. Treatment and study discontinuation was highest in the fixed-dose 10 mg/kg treatment arm due to discontinuation rules outlined in Table 18, Table 19, and Table 20. Of the 7 patients who discontinued the study from the 10 mg/kg fixed-dose treatment arm for “other reasons,” 6 were simply because they did not require a follow-up visit due to their treatment being discontinued for longer than 18 weeks.

Table 22: Study 103 Patient Disposition

Disposition	Study 103					
	Aducanumab 1 mg/kg N=31 n (%)	Aducanumab 3 mg/kg N=32 n (%)	Aducanumab 6 mg/kg N=30 n (%)	Aducanumab 10 mg/kg N=32 n (%)	Aducanumab Titration N=23 n (%)	Placebo N=48 n (%)
Patients randomized	31	33	30	32	23	48
PD analysis population	26 (84%)	29 (91%)	24 (80%)	28 (88%)	18 (78%)	42 (88%)
Efficacy analysis population	30 (97%)	32 (100%)	29 (97%)	30 (94%)	22 (96%)	46 (96%)
Discontinued treatment	7 (23%)	6 (19%)	5 (17%)	12 (38%)	4 (17%)	10 (21%)
Adverse event	3 (10%)	2 (6%)	3 (10%)	10 (31%)	2 (9%)	3 (6%)
Consent withdrawn	2 (6%)	1 (3%)	1 (3%)	1 (3%)	0	0
Other reasons	2 (6%)	3 (9%)	1 (3%)	1 (3%)	2 (9%)	7 (14%)
Discontinued study	8 (26%)	6 (19%)	4 (13%)	11 (34%)	3 (13%)	10 (21%)
Adverse event	3 (10%)	1 (3%)	0	3 (9%)	0	3 (6%)
Consent withdrawn	2 (6%)	1 (3%)	2 (7%)	1 (3%)	0	0
Other reasons	3 (10%)	4 (12%)	2 (7%)	7 (22%)	3 (13%)	7 (14%)

Source: Table 36 in Study 103 CSR

Protocol Violations/Deviations

Overall, 30% of patients had at least one major protocol deviation during the placebo-controlled period, but the incidence was similar across treatment arms. The most common categories of major protocol deviations were regarding study procedures criteria (15%) and informed consent (14%). These deviations are not expected to affect interpretation of the overall results. During the screening period there were 6 patients who did not meet the inclusion criteria for MMSE but were randomized. Two of these patients were in the placebo arm and none were in the fixed-dose 10 mg/kg arm. One patient randomized to placebo mistakenly received 2 doses of aducanumab 1 mg/kg. Two doses of aducanumab 1 mg/kg would not have an impact on the efficacy results for the placebo arm.

Table of Demographic Characteristics

Table 23 contains information regarding demographic characteristics for each treatment arm in the ITT population. Demographic characteristics are representative of the patient population

FDA Clinical Review
Kevin Krudys, PhD
BLA 761178
Aduhelm (aducanumab)

and are reasonably balanced across the treatment arms considering the number of patients in each arm. Patients in Study 103 were enrolled in 27 sites in the United States.

Table 23: Study 103 Baseline Demographics (ITT)

Demographic Parameters	Placebo (N=48) n (%)	Treatment Group					
		1 mg/kg (N=31) n (%)	3 mg/kg (N=32) n (%)	6 mg/kg (N=30) n (%)	10 mg/kg (N=32) n (%)	Titration (N=23) n (%)	Total (N=196) n (%)
Sex							
Male	20 (42%)	18 (58%)	15 (47%)	15 (50%)	17 (53%)	13 (57%)	98 (50%)
Female	28 (58%)	13 (42%)	17 (53%)	15 (50%)	15 (47%)	10 (43%)	98 (50%)
Age							
Mean years (SD)	73.3 (6.8)	72.6 (7.8)	70.5 (8.2)	73.3 (9.3)	73.7 (8.3)	73.1 (7.8)	72.8 (7.9)
Median (years)	73.5	74.0	71.0	72.0	74.5	73.0	73.0
Min, max (years)	54, 85	55, 88	54, 86	57, 91	51, 87	52, 88	51, 91
Age Group							
≥ 51 - ≤60 years	3 (6%)	2 (6%)	4 (13%)	2 (7%)	2 (6%)	1 (4%)	14 (7%)
≥ 61 - ≤70 years	13 (27%)	9 (29%)	10 (31%)	11 (37%)	6 (19%)	6 (26%)	55 (28%)
≥ 71 - ≤80 years	26 (54%)	16 (52%)	14 (44%)	8 (27%)	19 (59%)	12 (52%)	95 (48%)
≥ 81 - ≤90 years	6 (13%)	4 (13%)	4 (13%)	8 (27%)	5 (16%)	4 (17%)	31 (16%)
>90 years	0	0	0	1 (3%)	0	0	1 (<1%)
Race							
White	48 (100%)	31 (100%)	31 (97%)	28 (93%)	30 (94%)	23 (100%)	191 (97%)
Black or African American	0	0	0	2 (7%)	0	0	2 (1%)
Asian	0	0	0	0	1 (3%)	0	1 (<1%)
American Indian or Alaska Native	0	0	0	0	1 (3%)	0	1 (<1%)
Native Hawaiian or Other Pacific Islander	0	0	0	0	0	0	0
Other ¹	0	0	1 (3%)	0	0	0	1 (<1%)
Ethnicity							
Hispanic or Latino	1 (2%)	2 (6%)	1 (3%)	0	1 (3%)	1 (4%)	6 (3%)
Not Hispanic or Latino	47 (98%)	29 (94%)	31 (97%)	30 (100%)	30 (94%)	22 (96%)	189 (96%)

Source: Table 14 in Study 103 CSR

¹ Data on race and/or ethnicity were not collected because of local regulations.

Other Baseline Characteristics (disease characteristics, important concomitant drugs)

Table 24 contains a summary of baseline disease characteristics and baseline use of concomitant Alzheimer's disease medications. The disease characteristics represent a population similar to Studies 301 and 302 but with more mild patients as evidenced by 23% of patients with a CDR global score of 1 and 39% of patients with baseline MMSE less than 24. Baseline CDR-SB was lower in the placebo arm (2.69) than any of the aducanumab treatment arms (3.14 to 3.50). The rate of concomitant Alzheimer's disease medication was lowest in the fixed 10 mg/kg and titration arms. By design, the aducanumab titration arm included only ApoE ε4 carriers. Otherwise, the distribution of ApoE ε4 carrier status was similar to that observed in Studies 301 and 302.

Table 24: Study 103 Disease Characteristics (ITT Population)

Demographic Parameters	Placebo (N=48) n (%)	Treatment Group					
		1 mg/kg (N=31) n (%)	3 mg/kg (N=32) n (%)	6 mg/kg (N=30) n (%)	10 mg/kg (N=32) n (%)	Titration (N=23) n (%)	Total (N=196) n (%)
Baseline Clinical Stage							
Prodromal AD	22 (46%)	10 (32%)	14 (44%)	12 (40%)	13 (41%)	13 (57%)	84 (43%)
Mild AD	26 (54%)	21 (68%)	18 (56%)	18 (60%)	19 (59%)	10 (43%)	112 (57%)
Laboratory ApoE ε4 Status							
Carrier	34 (71%)	19 (61%)	21 (66%)	21 (70%)	20 (63%)	23 (100%)	138 (70%)
Homozygote	7 (15%)	1 (3%)	5 (16%)	4 (13%)	6 (19%)	3 (13%)	26 (13%)
Heterozygote	27 (56%)	18 (58%)	16 (50%)	17 (57%)	14 (44%)	20 (87%)	112 (57%)
Non-carrier	14 (29%)	12 (39%)	11 (34%)	9 (30%)	12 (38%)	0	58 (30%)
Number of Years of Formal Education							
Mean years (SD)	15.5 (2.98)	15.5 (3.15)	15.5 (2.42)	16.1 (2.76)	15.2 (2.35)	14.5 (3.33)	15.4 (2.84)
Median (years)	16.0	16.0	16.0	16.0	16.0	16.0	16.0
Min, Max (years)	12, 24	12, 27	12, 21	12, 23	10, 20	5, 21	5, 27
Concomitant AD medication							
Any AD medication at baseline	32 (67%)	21 (68%)	28 (88%)	20 (67%)	17 (53%)	12 (52%)	130 (66%)
Cholinesterase inhibitors	30 (63%)	20 (65%)	27 (84%)	19 (63%)	17 (53%)	11 (48%)	124 (63%)
Memantine	12 (25%)	6 (19%)	5 (16%)	8 (27%)	5 (16%)	3 (13%)	39 (20%)
Baseline CDR-SB							
Mean (SD)	2.69 (1.54)	3.40 (1.76)	3.50 (2.06)	3.32 (1.54)	3.14 (1.71)	3.24 (1.84)	3.17 (1.74)

Median	2.50	3.00	3.00	3.25	2.75	3.00	3.00
Min, Max	0.5, 7.0	0.5, 7.0	0.5, 8.0	1.0, 8.0	0.5, 7.0	0.5, 10.0	0.5, 10.0
Baseline CDR global score							
0.5	40 (83%)	22 (71%)	22 (69%)	25 (83%)	24 (75%)	18 (78%)	151 (77%)
1.0	8 (17%)	9 (29%)	10 (31%)	5 (17%)	8 (25%)	5 (22%)	45 (23%)
Baseline MMSE							
≤20	4 (8%)	3 (10%)	7 (22%)	0	0	0	14 (7%)
≥21 - <24	15 (31%)	11 (35%)	6 (19%)	11 (37%)	10 (31%)	9 (39%)	62 (32%)
≥24	29 (61%)	17 (55%)	19 (59%)	19 (63%)	22 (69%)	14 (61%)	120 (61%)

Source: adsl.xpt and Table 15 and Table 16 in Study 103 CSR

Treatment Compliance, Concomitant Medications, and Rescue Medication Use

During study treatment, compliance with study treatment was ≥94% in all treatment arms. Because of dose modifications due to ARIA, only 14 patients (44%) in the aducanumab fixed-dose 10 mg/kg arm received all 14 infusions compared to 26 patients (54%) in the placebo arms. Similarly, 63% of patients receiving aducanumab 10 mg/kg fixed-dose were on treatment for ≥48 weeks, compared to 86% of placebo patients.

Overall, 21 patients (11%) either changed the dose or initiated treatment with cholinesterase inhibitors and/or memantine during the placebo-controlled period but the number was reasonably balanced across the treatment arms.

Efficacy Results – Clinical Endpoints

The results for CDR-SB and MMSE are presented in Table 25. The reduction in clinical decline of CDR-SB and MMSE reached nominal significance for the aducanumab fixed-dose 10 mg/kg treatment arm. Based on analyses (linear contrast in ANCOVA) of the aducanumab fixed-dose treatment arms only, a dose-dependent relationship was found for both clinical endpoints. An MMRM model with the same covariates as the ANCOVA model specified in the SAP was used as a sensitivity analysis and provided similar results for the 10 mg/kg treatment arm for both endpoints (Table 26). Subgroup analysis was not performed due to the small sample size in the treatment arms.

Table 25: Study 103 Clinical Endpoints (CDR-SB and MMSE) Analyses by ANCOVA

	Placebo (N=46)	1 mg/kg (N=30)	3 mg/kg (N=32)	6 mg/kg (N=29)	10 mg/kg (N=30)	Titration (N=22)
Baseline CDR-SB						
n	44	28	32	29	30	22
Mean	2.65	3.27	3.47	3.22	3.14	3.23
Change from Baseline in CDR-SB at Week 54						

n	39	23	27	26	23	21
Adjusted mean	1.89	1.69	1.33	1.09	0.63	0.70
Standard error	0.350	0.441	0.413	0.423	0.446	0.499
95% CI	(1.198, 2.581)	(0.819, 2.564)	(0.517, 2.149)	(0.258, 1.930)	(-0.251, 1.511)	(-0.287, 1.686)
Difference from placebo		-0.20	-0.56	-0.80	-1.26	-1.19
95% CI for difference		(-1.308, 0.912)	(-1.612, 0.499)	(-1.855, 0.264)	(-2.356, -0.163)	(-2.343, -0.037)
% difference vs. placebo		-11%	-30%	-42%	-67%	-63%
p-value (compared with placebo)		0.7249	0.2995	0.1398	0.0246	0.0432
Baseline MMSE						
n	45	26	29	28	30	21
Mean	24.82	23.65	22.97	24.32	24.90	24.67
Change from Baseline in MMSE at Week 52						
n	40	25	26	26	25	21
Adjusted mean	-2.45	-2.21	-0.75	-1.99	-0.55	-1.00
Standard error	0.594	0.738	0.732	0.734	0.741	0.864
95% CI	(-3.627, -1.279)	(-3.664, -0.748)	(-2.198, 0.694)	(-3.435, -0.536)	(-2.010, 0.918)	(-2.702, 0.710)
Difference from placebo		0.25	1.70	0.47	1.91	1.46
95% CI for difference		(-1.612, 2.106)	(-0.141, 3.543)	(-1.356, 2.291)	(0.061, 3.754)	(-0.531, 3.445)
% difference vs. placebo		-8%	-68%	-20%	-76%	-60%
p-value (compared with placebo)		0.7932	0.0700	0.6133	0.0430	0.1496

Source: Table 23 and Table 24 in Study 103 CSR
p-values are nominal

Table 26: Study 103 Clinical Endpoints (CDR-SB and MMSE) Analyses by MMRM

	Placebo (N=46)	1 mg/kg (N=30)	3 mg/kg (N=32)	6 mg/kg (N=29)	10 mg/kg (N=30)	Titration (N=22)
Baseline CDR-SB						
n	48	31	32	30	32	23
Mean	2.69	3.40	3.50	3.32	3.14	3.24
Change from Baseline in CDR-SB at Week 54						
n	39	23	27	26	23	21
Adjusted mean	1.88	1.82	1.42	1.20	0.80	1.14
Standard error	0.337	0.424	0.403	0.420	0.424	0.481
95% CI	(1.210, 2.541)	(0.981, 2.653)	(0.627, 2.218)	(0.367, 2.2025)	(-0.041, 1.633)	(0.193, 2.094)

Difference from placebo		-0.06	-0.45	-0.68	-1.08	-0.73
95% CI for difference		(-1.122, 1.005)	(-1.483, 0.578)	(-1.730, 0.371)	(-2.141, -0.018)	(-1.866, 0.403)
% difference vs. placebo		-3%	-24%	-36%	-57%	-39%
p-value (compared with placebo)		0.9136	0.3869	0.2035	0.0464	0.2047
Baseline MMSE						
n	48	31	32	30	32	23
Mean	24.7	23.6	23.2	24.4	24.8	24.7
Change from Baseline in MMSE at Week 52						
n	40	25	26	26	25	21
Adjusted mean	-2.5	-2.2	-1.2	-2.0	-0.5	-1.1
Standard error	0.58	0.73	0.71	0.72	0.72	0.83
95% CI	(-3.60, -1.32)	(-3.64, -0.75)	(-2.59, 0.23)	(-3.45, -0.61)	(-1.94, 0.88)	(-2.74, 0.56)
Difference from placebo		0.3	1.3	0.4	1.9	1.4
95% CI for difference		(-1.56, 2.11)	(-0.52, 3.09)	(-1.37, 2.23)	(0.13, 3.74)	(-0.59, 3.34)
% difference vs. placebo		-12%	-52%	-16%	-76%	-56%
p-value (compared with placebo)		0.7708	0.1615	0.6366	0.0356	0.1680

Source: Table 9 and Table 10 in ISE appendix G5
p-values are nominal

Reviewer Comment: Differences between the two methods are minor. The MMRM results will be considered in the integration of efficacy to maintain consistency with the statistical approach in Studies 301 and 302.

Data Quality and Integrity

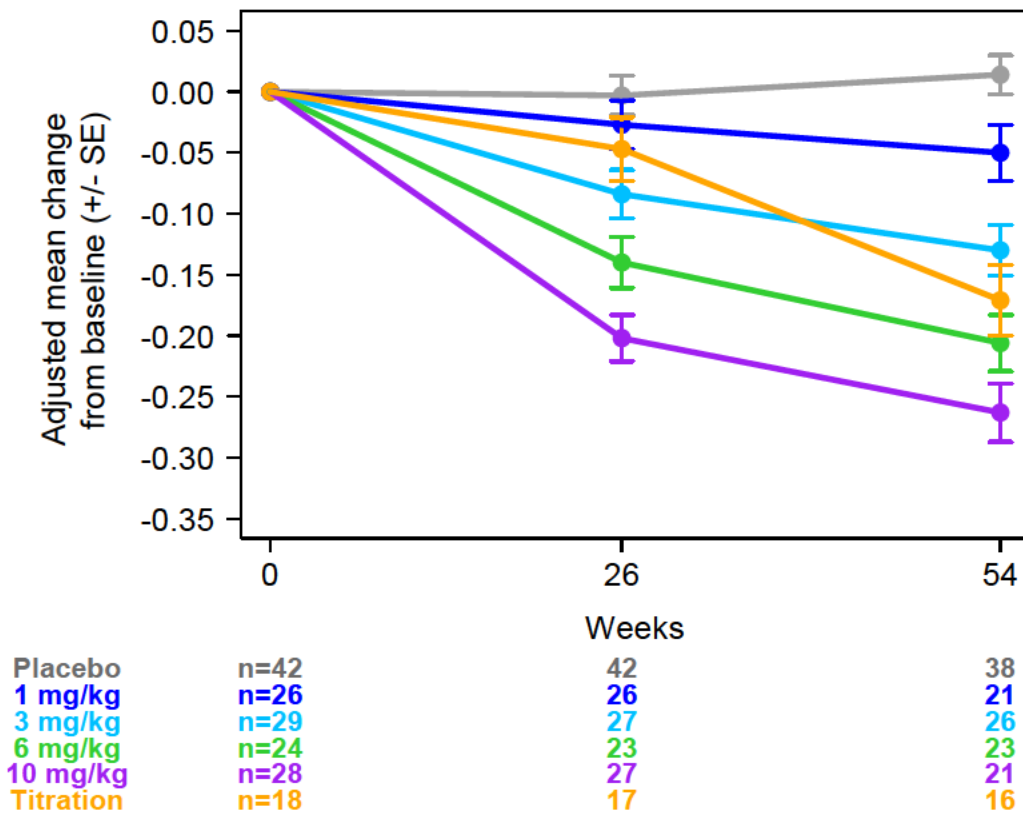
There were no major data quality issues identified during the review of Study 103.

Efficacy Results – Pharmacodynamic Endpoints

Change from baseline in brain amyloid signal as measured by SUVR was analyzed with an ANCOVA model adjusting for baseline SUVR and ApoE ε4 status. Statistically significant reductions in brain amyloid were achieved in the 3 mg/kg, 6 mg/kg and 10 mg/kg fixed-dose treatment groups at Week 26 and all aducanumab treatment groups at Week 54 (Figure 11). The adjusted mean change from baseline at Week 26 and Week 54 for the fixed-dose 10 mg/kg treatment arm were -0.202 and -0.262, respectively. The effect of the titration group was consistent with the effect anticipated for the average expected dose in that group (2.9 mg/kg at

Week 26 and 5.3 mg/kg at Week 54). Based on analyses (linear contrast in ANCOVA) of the aducanumab fixed-dose treatment arms only, a dose-dependent relationship was found.

Figure 11: Study 103 Change from Baseline in A β PET Composite SUVR



Created by reviewer from Table 22 in Study 103 CSR

Reviewer Comment: The reduction in A β PET composite SUVR compared to placebo for the fixed-dose 10 mg/kg arm at Week 54 (-0.277) was similar to the reduction at Week 78 in Study 302 (-0.278). Therefore, it is not unreasonable to compare clinical outcomes from Study 103 at Week 54 to those of Studies 301 and 302 at Week 78.

6. Integrated Review of Effectiveness

6.1. Assessment of Efficacy Across Trials

6.1.1. Primary Endpoints

The primary efficacy endpoint in Studies 301 and 302 was the change from baseline in CDR-SB at Week 78. Study 103 was designed primarily as a safety and tolerability study but included CDR-SB as an exploratory endpoint. Studies 301 and 302 were large, contemporaneous, and identically designed studies and therefore directly comparable. Study 103 included many of the elements of the other two studies with some minor differences: the sample size was considerably smaller, patient recruitment was entirely within the United States, and the entry criteria allowed for enrollment of patients later in the disease continuum. Differences in study design and patient demographics are provided in Section 5.

The aducanumab fixed-dose 10 mg/kg treatment arm in Study 103 is the relevant arm to compare to the high-dose arms of Studies 301 and 302. Although the primary endpoint was assessed earlier in Study 103 than Studies 301 and 302 (54 weeks vs. 78 weeks), the planned number of 10 mg/kg doses received during the treatment period was the same. The appropriateness of the comparison is also supported by the similar reduction in amyloid burden observed in the studies (Table 31).

A summary of the results for CDR-SB across studies is contained in Table 27. The statistically significant result in favor of high-dose aducanumab in Study 302 is consistent with the result from the fixed-dose 10 mg/kg treatment arm in Study 103. Direct comparison of the magnitude of the treatment effect between the two studies is complicated by the differences described earlier as well as the greater degree of decline in placebo patients in Study 103 in a shorter time period. Study 302 and Study 103 are also consistent in their demonstration of an apparent dose-response relationship. Importantly, the dose-response relationships are consistent with the dose- and exposure-response relationships for changes in amyloid burden and other biomarkers. The fixed-dose 1 mg/kg arm is the only treatment in the two studies that does not statistically change brain amyloid load and is also the only dose that has marginal numerical effects on CDR-SB.

The results of Studies 301 and 302 were only partially discordant. The low-dose aducanumab treatment arms provided very similar and numerically favorable estimates of change from baseline in CDR-SB at Week 78. Across the 3 studies, 7 out of 8 treatment arms that resulted in a reduction of amyloid burden numerically favored aducanumab treatment. It may be posited that aducanumab treatment itself causes a small bias in favor of treatment perhaps due to functional unblinding to ARIA. This possibility was explored (Section 6.1.7) and ultimately found to be lacking. Of course, the high-dose treatment in arm in Study 301 is an important finding and in isolation raises questions about the effectiveness of aducanumab. But in the context described above, it is not unreasonable to consider whether the high dose in Study 301 is an aberration. Section 6.1.7 will explore the discordant results in Studies 301 and 302 in detail. The treatment effect in Study 302 is consistent with the effect size used to power the study (-25%). Furthermore, the treatment effect would be expected to accrue over time because aducanumab targets the underlying pathophysiology of the disease.

Table 27: Summary of Findings for CDR-SB in Studies 103, 301 and 302

	Study 301			Study 302			Study 103					
Dose group	Placebo	Low Dose	High Dose	Placebo	Low Dose	High Dose	Placebo	1 mg/kg	3 mg/kg	6 mg/kg	Titration	10 mg/kg
N, ITT population	545	547	555	548	543	547	48	31	32	30	23	32
n at Week 78 or 54	333	331	295	288	290	299	39	23	27	26	21	23
Change from baseline	1.56	1.38	1.59	1.74	1.47	1.35	1.88	1.82	1.42	1.20	1.14	0.80
Difference from placebo (%)		-0.18 (-12%)	0.03 (2%)		-0.26 (-15%)	-0.39 (-22%)		-0.06 (-3%)	-0.45 (-24%)	-0.68 (-36%)	-0.73 (-39%)	-1.08 (-57%)
p-value		0.2250	0.8330		0.0901	0.0120		0.9136	0.3869	0.2035	0.2047	0.0464

Source: Table 13 in the Summary of Clinical Efficacy

6.1.2. Secondary and Other Endpoints

MMSE was the first secondary endpoint in the hierarchy for Studies 301 and 302 and the only other clinical endpoint that was collected in all three studies. Results are presented in Table 28 and are generally similar to those for CDR-SB except that the low-dose treatment effects in Studies 301 and 302 do not consistently favor aducanumab. Also, the 6 mg/kg treatment arm in Study 103 is not consistent with a dose-response relationship.

Table 28: Summary of Findings for MMSE in Studies 103, 301 and 302

	Study 301			Study 302			Study 103					
Dose group	Placebo	Low Dose	High Dose	Placebo	Low Dose	High Dose	Placebo	1 mg/kg	3 mg/kg	6 mg/kg	Titration	10 mg/kg
N, ITT population	545	547	555	548	543	547	48	31	32	30	23	32
n at Week 78 or 52	332	334	297	288	293	299	40	25	26	26	21	25
Change from baseline	-3.5	-3.3	-3.6	-3.3	-3.3	-2.7	-2.5	-2.2	-1.2	-2.0	-1.1	-0.5
Difference from placebo (%)		0.2 (-6%)	-0.1 (3%)		-0.1 (3%)	0.6 (-18%)		0.3 (-12%)	1.3 (-52%)	0.4 (-16%)	1.4 (-56%)	1.9 (-76%)
p-value		0.4795	0.8106		0.7578	0.0493		0.7708	0.1615	0.6366	0.1680	0.0356

Source: Table 14 in the Summary of Clinical Efficacy

Results for ADAS-Cog 13 and ADCS-ADL-MCI are presented in Table 29 and Table 30, respectively. The findings are consistent with those of CDR-SB in two respects: (1) the high-dose

arm in Study 302 demonstrated a statistically significant and clinically meaningful reduction in decline and (2) the low-dose arms in both studies show a consistent numerical trend in favor of aducanumab. The principal components analysis presented by the applicant demonstrates that each of the 3 secondary endpoints captures distinct information regarding cognitive decline. Therefore, the strongly positive results on the secondary endpoints reinforce the persuasiveness of the findings on the primary endpoint.

Table 29: Summary of Findings for ADAS-Cog 13 in Studies 301 and 302

	Study 301			Study 302		
Dose group	Placebo	Low Dose	High Dose	Placebo	Low Dose	High Dose
N, ITT population	545	547	555	548	543	547
n at Week 78 or 54	331	332	294	287	289	293
Change from baseline	5.140	4.558	4.552	5.162	4.461	3.763
Difference from placebo (%)		-0.583 (-11%)	-0.588 (-11%)		-0.701 (-14%)	-1.400 (-27%)
p-value		0.2536	0.2578		0.1962	0.0097

Source: Table 16 in the Summary of Clinical Efficacy

Table 30: Summary of Findings for ADCS-ADL-MCI in Studies 301 and 302

	Study 301			Study 302		
Dose group	Placebo	Low Dose	High Dose	Placebo	Low Dose	High Dose
N, ITT population	545	547	555	548	543	547
n at Week 78 or 54	331	330	298	283	286	295
Change from baseline	-3.8	-3.1	-3.1	-4.3	-3.5	-2.5
Difference from placebo (%)		0.7 (-18%)	0.7 (-18%)		0.7 (-16%)	1.7 (-40%)
p-value		0.1225	0.1506		0.1515	0.0006

Source: Table 17 in the Summary of Clinical Efficacy

Although the placebo decline for CDR-SB was numerically greater in Study 302 (1.74) than Study 301 (1.56), differences in placebo response do not appear to explain why Study 302 was successful and Study 301 was not. First, the placebo decline for CDR-SB in both studies was consistent with expectation (i.e., placebo decline was assumed to be 2 for power calculations) and not markedly different from other studies in similar populations. Second, placebo decline

was actually lower for MMSE in Study 302 compared to Study 301 and placebo decline was similar for ADAS-Cog 13 between the two studies. Consistently positive outcomes were observed across endpoints in Study 302, not just ones with a greater placebo decline. Third, similar treatment effect estimates relative to placebo for CDR-SB were obtained for the two low dose arms. Finally, placebo decline cannot account for the apparent dose-response relationship observed in 302. Although the study was not designed to specifically assess dose-response, the trend observed in Study 302 is consistent with a dose-response relationship for brain amyloid observed in Study 302 and consistent with the dose-response relationship for both brain amyloid and clinical outcomes observed in Study 103.

The change from baseline in amyloid signal as measured by PET and quantified by a composite SUVR was a key pharmacodynamic marker collected using the same methods across all three studies. The results of these studies and the accompanying PKPD modeling demonstrate a robust relationship between aducanumab exposure and reduction in amyloid signal. Interestingly, the high-dose arm in Study 301 yielded a reduction in SUVR commensurate with the 6 mg/kg dose in Study 103, rather than the high dose in Study 302 or the 10 mg/kg dose in Study 103. This observation alone cannot explain the negative outcome of Study 301 but provides an avenue for further exploration. Specifically, it supports the hypothesis that dosing in Study 301 was lower than in Study 302 and that this lower dosing contributed in part to the discordant results.

Table 31: Summary of Findings for Amyloid PET Composite SUVR in Studies 103, 301 and 302

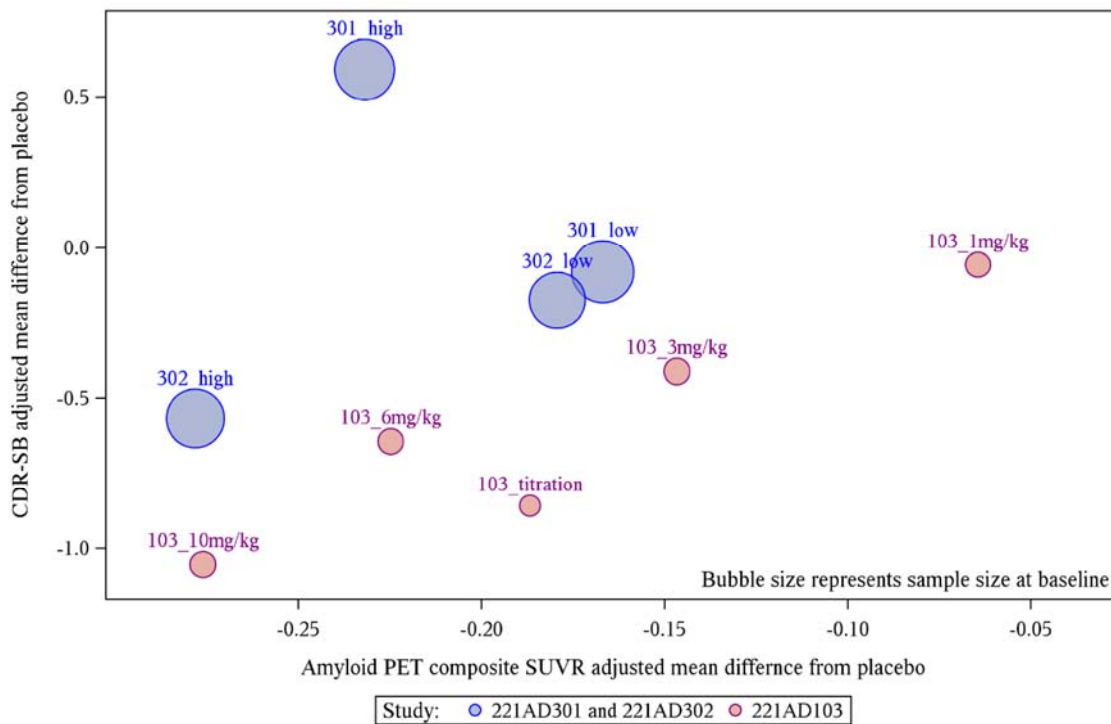
	Study 301			Study 302			Study 103					
Dose group	Placebo	Low Dose	High Dose	Placebo	Low Dose	High Dose	Placebo	1 mg/kg	3 mg/kg	6 mg/kg	Titration	10 mg/kg
N, PET population	204	198	183	159	159	170	46	29	32	30	19	31
n at Week 78 or 54	124	138	112	93	100	109	38	21	26	23	16	21
Change from baseline	-0.003	-0.170	-0.235	0.014	-0.165	-0.264	0.017	-0.047	-0.130	-0.208	-0.170	-0.259
Difference from placebo		-0.167	-0.232		-0.179	-0.278		-0.064	-0.147	-0.225	-0.187	-0.276
p-value		<1e-4	<1e-4		<1e-4	<1e-4		0.0231	<1e-4	<1e-4	<1e-4	<1e-4

Source: Table 15 in the Summary of Clinical Efficacy

The applicant examined the group-level correlation between change from baseline on A β PET and CDR-SB across the three studies (Figure 12). It is noteworthy that due to differences in study design, Week 78 was used for Studies 301 and 302 and Week 54 was used for Study 103. There also other differences in study design, analysis and study populations that affect the interpretation of such as analysis. But from a qualitative perspective, the analysis is consistent with there being a general relationship between brain amyloid reduction and treatment effect

for CDR-SB. Figure 12 also highlights the apparently aberrant performance of the high dose in Study 301.

Figure 12: Group-Level Correlation Between Adjusted Mean Difference from Placebo in A β PET Composite SUVR and CDR-SB



Source: Figure 61 in ISE

6.1.3. Subpopulations

A pooled analysis across trials is not provided because Study 301 is a negative trial and the sample size for Study 103 is small. Subgroup analysis for Study 302 is presented in Section 5.1.2. Treatment effects favored aducanumab for the various subpopulations considered.

One factor of interest is ApoE ϵ 4 carriage. The only subgroup (out of 80 across the primary and secondary endpoints) with a point estimate of the treatment effect that did not favor aducanumab was MMSE in ApoE ϵ 4 non-carriers. In Figure 2, the trend for a lower response in ApoE ϵ 4 non-carriers is also evident for CDR-SB. This might be viewed as unexpected, given that ApoE ϵ 4 non-carriers had the opportunity to be exposed to the 10 mg/kg dose throughout the study. The applicant provided additional analyses using the OTC and uncensored datasets which question the robustness of this finding. For example, in the OTC dataset, the percent decline versus placebo for CDR-SB was -23% for ApoE ϵ 4 carriers and -19% for ApoE ϵ 4 non-carriers. Interestingly, in the disease progression model developed by the applicant, ApoE ϵ 4 carrier status was not a clinically meaningful covariate on the rate of progression after incorporating a

mixture model to describe different subpopulations of progressors. Also, brain amyloid is similarly reduced in ApoE ϵ 4 carriers and ApoE ϵ 4 non-carriers, suggesting the underlying pharmacology is similar in the two populations. In sum, the data suggest there might be a more modest effect in ApoE ϵ 4 non-carriers compared to ApoE ϵ 4 carriers, but a definitive conclusion cannot be drawn.

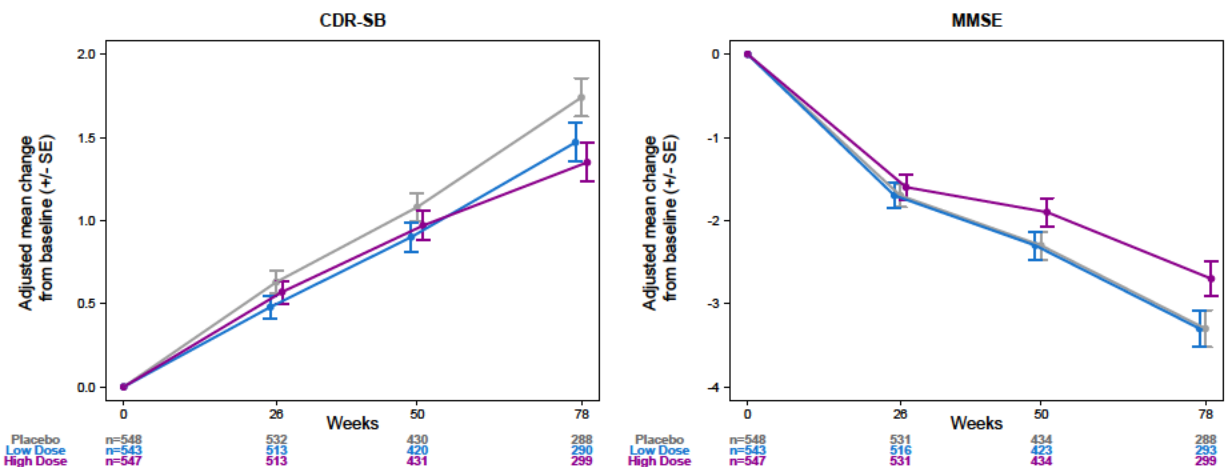
6.1.4. Dose and Dose-Response

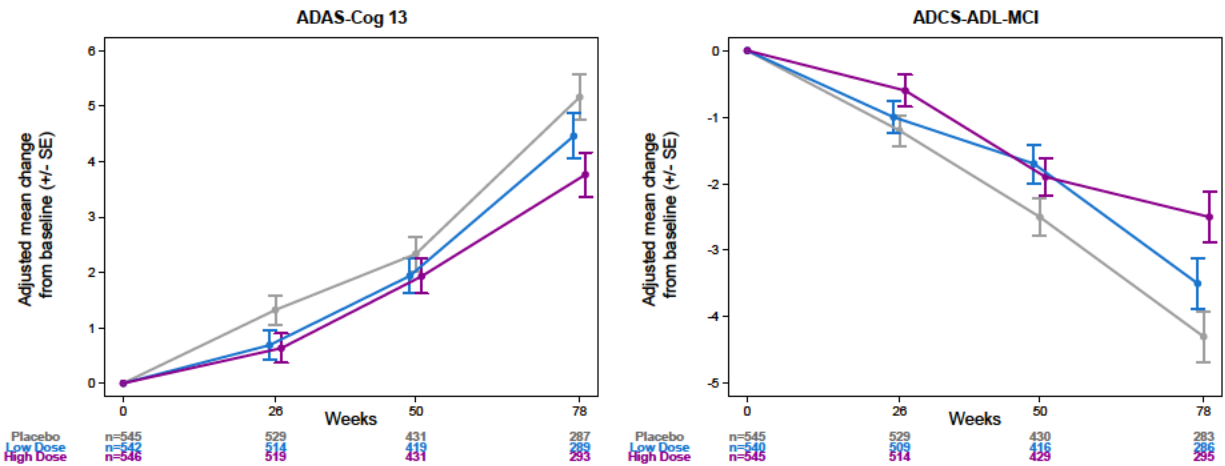
The results of Study 103 and 302 were indicative of a dose-response relationship (see Sections 5.1.2 and 5.2.2). Further considerations of the dose-response relationship are also provided in 6.1.7 and in the clinical pharmacology review.

6.1.5. Onset, Duration, and Durability of Efficacy Effects

Week 78 is the earliest time in Study 302 when statistically convincing evidence of a clinically meaningful effect is observed across all clinical endpoints (Figure 13). There were a few instances at Week 26 or 50 where the adjusted mean change from baseline for ADAS-Cog 13 or ADCS-ADL-MCI crossed the bounds to nominal significance, but the effect was not sustained.

Figure 13: Study 302 Clinical Endpoints Adjusted Mean Change from Baseline over Time

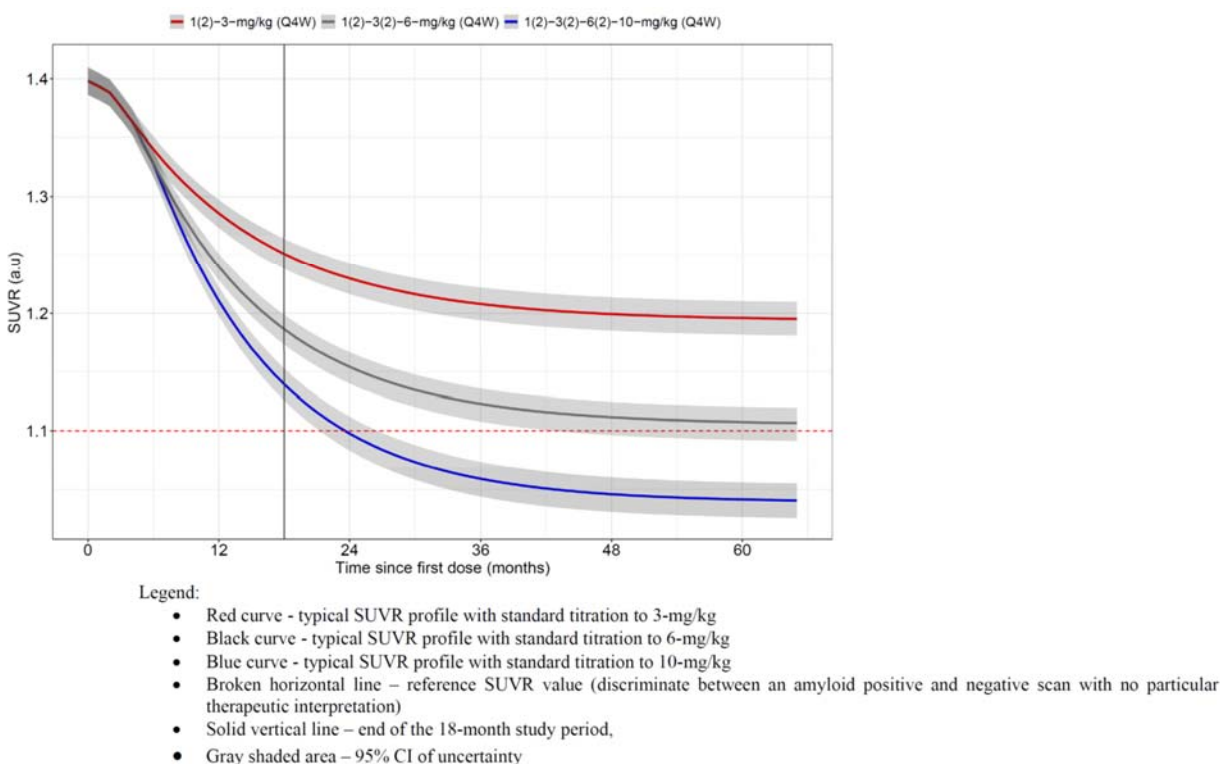




Source: Created from Figures 4, 7, 9 and 11 in Study 302 CSR

Although the precise relationship between reduction in amyloid plaque and clinical outcomes is unknown, consideration of the time-course of SUVR for the proposed dosing regimen is instructive (Figure 14). The first key observation is that the reduction in brain amyloid is delayed with respect to initiation of treatment. Based on the understanding of the disease process, one can reasonably assume that the effect on tau accumulation is even further delayed, although the staging of events is not discernible from the data collected in the studies. Second, the pharmacodynamic effect of aducanumab on amyloid plaque reduction is still increasing during the duration of the trial and does not reach maximum effect until more than a year after the double-blind treatment period. Together, these observations are consistent with an effect on clinical outcomes that did not clearly emerge until the end of the study, with favorable trends observed at earlier time points. Moreover, these findings suggest that the effect should be durable and increase over time. On the other hand, these same observations raise the possibility that patients receiving aducanumab treatment progressed as normal during the early portion of the trial before therapeutic levels of amyloid reduction were reached, thereby making it more difficult to demonstrate effectiveness in the studies.

Figure 14: Typical Profile of SUVR for Planned Titration and No Dose Interruptions



Source: Figure 8 in Summary of Clinical Pharmacology

The persistence of clinical benefit after treatment has been stopped is only indirectly assessed by the efficacy results using the uncensored ITT population (Table 16) because this data set includes assessments after the March 20, 2019, termination of study treatment. The magnitude of the treatment effect is consistent with the estimate derived from the ITT population. These results should be viewed with caution, however, because of the termination of the studies and the subsequent unblinding of some patients.

6.1.6. Early Termination and Interpretation of Study 301 and Study 302

The applicant announced the early termination of Studies 301 and 302 on March 21, 2019 based on the results of a prespecified futility analysis of interim data. A detailed timeline of the interactions leading to the futility declaration is included in Section 5.1.1. The futility criteria were based on conditional power, which was defined by the applicant as the chance that the primary efficacy endpoint analysis will be statistically significant in favor of aducanumab at the planned final analysis, given the data at the interim analysis. Conditional power was calculated assuming that the future unobserved effect would be equal to the maximum likelihood estimate of the pooled observed interim data from Studies 301 and 302. Pooling of the data assumes that the treatment effect would be similar in the two studies. The studies were to be

considered futile if both studies had conditional power less than 20% for both the high-dose and low-dose treatment groups. The SAP specified that other data, in addition to the prespecified futility criteria, could also be considered. At the June 14, 2019, Type C meeting the applicant asserted that if the conditional power had been less than 20% for CDR-SB, but greater than 90% for either MMSE or ADAS-Cog 13, the studies could have continued.

The futility analysis was planned to occur after approximately 50% of the patients enrolled in the studies had the opportunity to complete the Week 78 visit. Data from patients who were enrolled in the studies but did not have the opportunity to complete the Week 78 visit were not included in the calculation of conditional power.

At the interim data cutoff date of December 26, 2018, 57% of participants from Study 301 and 49% of participants from Study 302 had the opportunity to complete the Week 78 visit. Table 32 provides the results of the futility analysis and shows that the futility criteria were met. The results also clearly demonstrate divergent estimates of treatment effect for the high-dose treatment arm in the two studies for all three endpoints. In other words, the assumption that the treatment effect would be similar in the two studies was not realized.

Table 32: Conditional Power for Change from Baseline in CDR-SB, MMSE and ADAS-Cog13 at Week 78 (Interim Futility Dataset)

	Study 301			Study 302		
	Diff vs. placebo (%) Conditional Power			Diff vs. placebo (%) Conditional Power		
	Placebo decline (N=324)	Low Dose (N=320)	High Dose (N=301)	Placebo decline (N=258)	Low Dose (N=269)	High Dose (N=276)
CDR-SB	1.45	-0.15 (-10%) 12.9%	0.22 (15%) 0%	1.53	-0.10 (-7%) 10.7%	-0.28 (-18%) 11.8%
MMSE	-3.2	0.3 (-9%) 5.7%	-0.4 (13%) 0%	-3.0	-0.1 (3%) 0.3%	0.6 (-20%) 24.9%
ADAS-Cog 13	4.827	-0.472 (-10%) 9.8%	-0.058 (-1%) 3.3%	5.488	-0.311 (-6%) 9.0%	-1.089 (-20%) 55.2%

Source: Table 5 in ISE

Upon review of the futility analysis and prior to the June 14, 2019, Type C meeting, the Division asked the applicant to recalculate the conditional power assuming the future unobserved effect would be equal to the maximum likelihood estimates for each individual study independently. The results are provided in Table 33 and show that the futility criteria would not have been met with this analysis. Estimates of the treatment effect for all three endpoints in the high-dose arm in Study 302 numerically favor aducanumab treatment. For these reasons, the futility decision is not an accurate reflection of the individual studies.

Table 33: Conditional Power (Non-Pooled Analysis) for Change from Baseline in CDR-SB, MMSE and ADAS-Cog13 at Week 78 (Interim Futility Dataset)

	Study 301			Study 302		
	Diff vs. placebo (%) Conditional Power			Diff vs. placebo (%) Conditional Power		
	Placebo decline (N=324)	Low Dose (N=320)	High Dose (N=301)	Placebo decline (N=258)	Low Dose (N=269)	High Dose (N=276)
CDR-SB	1.45	-0.15 (-10%) 11.23%	0.22 (15%) 0%	1.53	-0.10 (-7%) 4.4%	-0.28 (-18%) 58.6%
MMSE	-3.2	0.3 (-9%) 12.4%	-0.4 (13%) 0%	-3.0	-0.1 (3%) 0.02%	0.6 (-20%) 78.2%
ADAS-Cog 13	4.827	-0.472 (-10%) 8.1%	-0.058 (-1%) 0.3%	5.488	-0.311 (-6%) 3.7%	-1.089 (-20%) 72.4%

Source: Table 6 in ISE

It is important to note that rigorous, per-protocol collection of blinded data continued between the time of the data cutoff in December 2018 and the futility announcement in March 2019. Subsequently, the primary analysis per the SAP was conducted on the more complete ITT dataset including assessments collected before March 21, 2019. On face, Study 302 was a positive study which might be considered exceptionally persuasive on several of the instruments used to evaluate efficacy. Study 301 remained a negative study. To a casual observer, the turnabout from futility to a positive study might suggest that the data collected after the futility announcement was completely incongruent with the data used to make the futility decision or that some new analysis was performed on the data. Although the treatment effect did improve over time (e.g., -18% to -22% for the high dose in Study 302), one can see from Table 33 that Study 302 was trending positively and had a reasonable likelihood of success if run to conclusion. It is not surprising that the treatment effect would improve over time, as protocol amendments implemented during the studies had the effect of increasing the exposure to aducanumab, although other factors also clearly contributed.

Before further consideration could be given to Studies 301 and 302 it was imperative to evaluate the effect of early termination of the studies on the interpretability of the observed efficacy data and associated analyses. Virtual completion of the studies using modeling and simulation was used to explore the range of plausible outcomes had the studies been run to completion. Two approaches were used to virtually complete the studies. Both approaches estimated parameters from a target dataset, simulated patient-level data, and fit an MMRM model to the simulated data. The primary approach supplemented the existing observed data with simulated assessments for the data that were censored due to the early termination of the trials. Another approach fully simulated studies to explore the range of plausible results if many trials like Studies 301 and 302 were run from start to completion. Overall, simulation results were highly consistent with the primary analysis of the observed data. Similar results were

obtained using all data (ITT) or only data in patients who had the opportunity to complete the Week 78 visit.

The results of this exercise established the following:

- The early termination of the aducanumab program does not compromise the interpretability of the efficacy results of Studies 301 and 302
- The appropriate dataset for further consideration is the ITT population consisting of all randomized subjects who received at least one dose of study treatment and excluding data collected after March 20, 2019.

Thus, the final analysis results presented earlier in Table 27 to Table 30 simply provide the final, interpretable results of trials that were terminated early but analyzed in accordance with the prespecified analysis.

6.1.7. Analysis of Discordant Results in Study 301 and Study 302

Once the results of Studies 301 and 302 were deemed interpretable, the review team was met with a unique situation of having a large and complicated, but incomplete and partially discordant, data set suggestive of the possible effectiveness of aducanumab.

Upon initial review, the one positive study (Study 302) and one negative study (Study 301) were given equal weight and consideration. Despite divergent outcomes in the primary endpoint, there were some key similarities between the two studies. As noted earlier, the low-dose aducanumab treatment arms, while not statistically significant, demonstrated consistent numerical effects favoring aducanumab in the studies. Also, aducanumab produced a time- and dose-dependent reduction in brain amyloid burden. Upon closer review of the individual studies, Study 302 appeared to be a strongly positive study on many distinct and important clinical measures, robust to sensitivity analyses, and supported by well-characterized biomarker data. In the context of a positive Study 302, the dose-response relationship observed in Study 103, and the numerically favorable results of similar magnitude in the low-dose groups in both studies, the high dose in 301 tends to stand apart. Also, the magnitude of the effect on CDR-SB in the high-dose group in Study 301 underwent the greatest change from the futility analysis (15%) to the final analysis (2%), suggesting instability in the estimate.

If aducanumab is effective, it follows that Study 302 is reflective of the true effect of aducanumab and that there would be patients in Study 301 who, based on certain characteristics, should show response of similar character to patients in Study 302. Given the possibility that aducanumab is an effective drug for a disease with an enormous unmet medical need, it was agreed by the applicant and Division at the June 14, 2019, Type C Meeting that

extensive resources should be brought to bear on achieving a maximum understanding of the existing data.

By their nature, these analyses are post hoc and exploratory and therefore carry with them the appropriate caveats and caution in their interpretation. To address these concerns, any exploration of the data was to be rigorous, limited in scope, and based on well-defined hypotheses. To the maximum degree possible, the analyses were pre-specified. An important distinction is that these analyses were not aimed at obtaining independent support from Study 301. Study 301 is a negative study. The purpose of these analyses is to provide maximum understanding of the partially discordant results and to determine if this understanding either supports or undermines the outcome of Study 302.

Demographics and Baseline Disease Characteristics

A first step to establish the foundation for future investigation was to rule out any glaring differences in the characteristics of the patients enrolled in the two studies, especially in the high-dose arms. Prior to the June 14, 2019, Type C Meeting, the Division requested the applicant investigate this possibility even though differences were not expected due to the identical design and timing of the studies. The individual studies are reviewed in Sections 5.1.2 and 5.2.2 and no major differences in demographics or baseline disease characteristics between the studies were noted.

A population exposure-CDR-SB model developed by the applicant revealed that observable baseline factors do not explain much of the variability in disease progression. This suggests that even if there were small differences in observed baseline characteristics, they are unlikely to explain the discordant results in the high dose arms. Most of the variability between patients in disease progression is due to unknown factors that should be balanced by randomization.

Influence of Study Participants with Rapid Progression

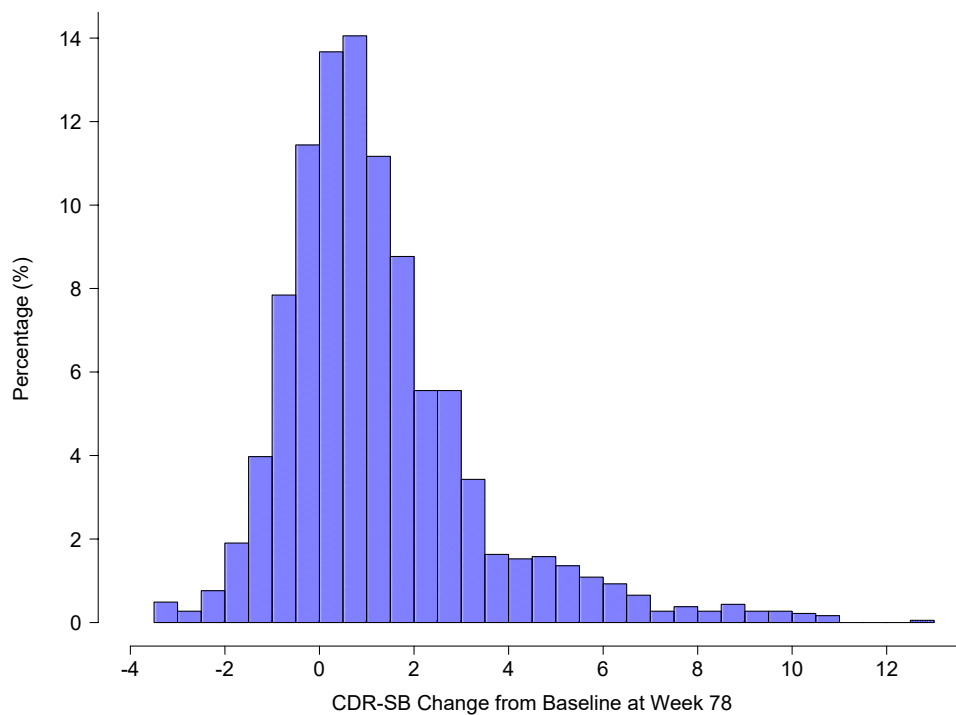
Investigation of individual treatment responses, particularly identification of participants with a rapid rate of disease progression, was an area of focus due to the following considerations:

- Diagnostic plots of the primary endpoint analysis demonstrated that the distribution of the change from baseline in CDR-SB was skewed.
- At the time of the futility analysis, the high-dose treatment arm in Study 301 was unique in that aducanumab-treated patients had greater cognitive decline (15%) than patients receiving placebo. A credible indication that aducanumab treatment was accelerating cognitive decline in Study 301 would undermine the results of Study 302. It was

therefore particularly important to interrogate the nature of the decline in aducanumab high-dose patients in Study 301.

A graphical review of the data revealed that CDR-SB change from baseline at Week 78 was right-skewed with a small percentage of patients having relatively large increases (Figure 15). Some degree of right-skewness is expected as the mean baseline CDR-SB was approximately 2.5 on a scale that ranges from 0 to 18. Also, the existence of a small proportion of patients with a rapid rate of disease progression has been noted in the literature. The question was not whether these rapid progressors existed in the data, but rather what potential impact they had on the primary analysis and whether that impact was similar across treatment groups.

Figure 15: Distribution of Change from Baseline in CDR-SB at Week 78 in Pooled Treatment Arms in Studies 301 and 302



Created by reviewer using adqs.xpt

A cutoff of an 8-unit increase in CDR-SB over 78 weeks was initially chosen to define patients with rapid progression. This increase is 4 times the expected estimate of mean progression used to power Studies 301 and 302. The numbers of patients in each study who met the cutoff were evenly distributed across the treatment groups except for the high-dose group in Study 301 (Table 34).

Table 34: Number of Patients with Rapid Progression (CDR-SB Change from Baseline of >8 at Week 78)

	Placebo	Low Dose	High Dose	Total
Study 301	4	5	9	18
Study 302	4	4	5	13

Source: Table 34 in ISE

Analyses for the primary and secondary endpoints were conducted using a dataset with all rapid progressors removed and compared to the results from the ITT population. Results for the primary and secondary endpoints were most sensitive to rapid progressors in the high-dose treatment arm in Study 301. For example, the results for CDR-SB went from a 2% increase relative to placebo, to a -6% decrease when the rapid progressors were removed (Table 35). Minor differences were observed at the lower dose in Study 301. Likewise, results of Study 302 were not sensitive to exclusion of rapid progressors (Table 36). The analysis was repeated for other cutoff values to define rapid progressors and the results were similar.

Table 35: Study 301 Change from Baseline in CDR-SB, MMSE and ADAS-Cog13 at Week 78 Excluding Rapid Progressors (CDR-SB Change from Baseline > 8)

	Study 301 (excluding rapid progressors)			Study 301 (ITT)		
	Placebo decline (N=541)	Diff vs. placebo (%)		Placebo decline (N=545)	Diff vs. placebo (%)	
		Low Dose (N=542)	High Dose (N=546)		Low Dose (N=547)	High Dose (N=555)
CDR-SB	1.47	-0.19 (-13%)	-0.09 (-6%)	1.56	-0.18 (-12%)	0.03 (2%)
MMSE	-3.5	0.2 (-6%)	0.1 (-3%)	-3.5	0.2 (-6%)	-0.1 (3%)
ADAS-Cog 13	5.033	-0.607 (-12%)	-0.818 (-16%)	5.140	-0.583 (-11%)	-0.588 (-11%)
ADCS-ADL-MCI	-3.7	0.8 (-22%)	1.0 (-27%)	-3.8	0.7 (-18%)	0.7 (-18%)

Source: Table 3 in Appendix C in ISE

Table 36: Study 302 Change from Baseline in CDR-SB, MMSE and ADAS-Cog13 at Week 78 Excluding Rapid Progressors (CDR-SB Change from Baseline > 8)

	Study 302 (excluding rapid progressors)			Study 302 (ITT)		
	Placebo decline (N=544)	Diff vs. placebo (%)		Placebo decline (N=548)	Diff vs. placebo (%)	
		Low Dose (N=539)	High Dose (N=542)		Low Dose (N=543)	High Dose (N=547)

CDR-SB	1.66	-0.29 (-17%)	-0.42 (-25%)	1.74	-0.26 (-15%)	-0.39 (-22%)
MMSE	-3.2	-0.1 (3%)	0.6 (-19%)	-3.3	-0.1 (3%)	0.6 (-18%)
ADAS-Cog 13	4.996	-0.729 (-15%)	-1.335 (-27%)	5.162	-0.701 (-14%)	-1.400 (-27%)
ADCS-ADL-MCI	-4.1	0.7 (-17%)	1.9 (-46%)	-4.3	0.7 (-16%)	1.7 (-40%)

Source: Table 2 in Appendix C in ISE

The applicant investigated potential group- and individual-level factors that may have contributed to the rapid clinical decline in the 31 patients defined as rapid progressors. Demography, baseline disease characteristics, comorbidities, concomitant medication use, imaging biomarkers, and exposure were explored at the group level, but a distinct pattern for rapid progressors did not emerge. There was a trend for lower baseline whole cortex volume, increased lateral ventricle volume, and higher baseline ADA-Cog 13 scores in rapid progressors but there was considerable overlap with the overall population.

Even though the high-dose arm in Study 301 was the only aducanumab-treatment arm with a relatively higher number of rapid progressors, it was important to investigate whether aducanumab treatment itself was the cause of rapid progression. Of the 23 rapid progressors who received aducanumab, 10 had an adverse event of ARIA-E, ARIA-H, or superficial siderosis, although most were mild or moderate and asymptomatic. For the two patients in Study 301 who received high-dose aducanumab and had symptomatic ARIA, the symptoms occurred late in the treatment after significant clinical decline had already occurred. Other AEs were also reviewed, but there was no consistent pattern with respect to rapid progression.

Taken together, these analyses suggest that (1) small imbalances in the number of rapid progressors can have a relatively large impact on the magnitude of the primary and secondary endpoints and (2) the high-dose arm in Study 301 was disproportionately affected by such an imbalance in rapid progressors. With rapid progressors excluded from the analysis, the high-dose aducanumab arm in Study 301 suggests a numerical trend in favor of aducanumab, consistent with every other aducanumab treatment arm. It is also important to note that 8 of the 9 rapid progressors from the high-dose arm in Study 301 were included in the futility data cut. This further casts doubt on the appropriateness of the futility determination, as the numerical trend in favor of placebo over high-dose aducanumab in Study 301 may be due to a small imbalance in rapid progressors rather than a systematic worsening across aducanumab-treated patients.

Influence of ARIA and ARIA Management

The potential impact of ARIA and ARIA management on efficacy results was investigated for the following reasons:

- The occurrence of ARIA has the potential to cause functional unblinding of investigators, patients, and caregivers because ARIA is associated with aducanumab treatment and it prompts differential management of patients
- ARIA management evolved over the course of the studies (see Section 5.1.1)
- An initial review of the data revealed that the number of patients with symptomatic ARIA was higher for the high-dose arm in Study 301 (65) than for the high-dose arm in Study 302 (45)

The incidence and radiographic severity of ARIA was generally similar between Study 301 and Study 302. A small imbalance in symptomatic ARIA was noted for patients in the high-dose arms and a post hoc analysis suggested that the primary endpoint was sensitive to exclusion of patients with symptomatic ARIA in Study 301. Upon further review, it was determined that this sensitivity was almost entirely due to two patients who were rapid progressors and experienced symptomatic ARIA. As noted in the previous section, rapid clinical decline occurred prior to the occurrence of ARIA in these patients. Therefore, differences in the occurrence or severity of ARIA are not likely to explain the discordant results at the high dose in the two studies.

Given the similar incidence of ARIA between the two studies, potential unblinding due to ARIA is likely to affect both studies equally. As such, the following analyses are not likely to directly address the discordant results in the two studies but are presented more as sensitivity analyses. It is important to reiterate that steps were taken in the protocol to minimize functional blinding, specifically the use of an independent rater who was blinded to patient management, including occurrence of ARIA and dose modifications.

To address the potential effect of functional unblinding due to ARIA, the applicant compared the results of the primary analysis using the ITT dataset to results using a reduced ITT dataset in which all assessments after occurrence of ARIA were excluded. For the low-dose and high-dose arms, approximately 30% and 40% of the observations were excluded, respectively. Overall, the results do not suggest a systematic bias due to functional unblinding (Table 37). The treatment difference for the high-dose arm in Study 302 was higher (33% vs. 22% for CDR-SB) after excluding observations post-ARIA. It is worth noting that all 4 patients in Study 302 who were rapid progressors and experienced ARIA also received the high-dose. The occurrence of ARIA in these patients tended to happen early in the course of treatment, so exclusion of post-ARIA observations likely influenced this finding. It also demonstrates that rapid progressors can have an outsized effect when interpreting subgroups.

Table 37: Change from Baseline in CDR-SB, MMSE, ADAS-Cog 13 and ADCS-ADL-MCI at Week 78: All Observations vs. Exclusion of Observations Post-ARIA

	Study 301			Study 302		
	Placebo decline (N=545)	Diff vs. placebo (%)		Placebo decline (N=548)	Diff vs. placebo (%)	
		Low Dose (N=547)	High Dose (N=555)		Low Dose (N=543)	High Dose (N=547)
CDR-SB						
All observations	n=333 1.56	N=331 -0.18 (-12%)	N=295 0.03 (2%)	N=288 1.74	N=290 -0.26 (-15%)	N=299 -0.39 (-22%)
Post-ARIA observations excluded	n=298 1.55	N=240 -0.11 (-7%)	N=181 0.00 (0%)	N=254 1.72	N=194 -0.19 (-11%)	N=172 -0.57 (-33%)
MMSE						
All observations	N=332 -3.5	N=334 0.2 (-6%)	N=297 -0.1 (3%)	N=288 -3.3	N=293 -0.1 (3%)	N=299 0.6 (-18%)
Post-ARIA observations excluded	N=297 -3.6	N=242 0.2 (-6%)	N=182 -0.1 (3%)	N=254 -3.4	N=195 -0.1 (3%)	N=172 0.8 (-24%)
ADAS-Cog 13						
All observations	N=331 5.140	N=332 -0.583 (-11%)	N=294 -0.588 (-11%)	N=287 5.162	N=289 -0.701 (-14%)	N=293 -1.400 (-27%)
Post-ARIA observations excluded	N=296 5.143	N=240 -0.507 (-10%)	N=180 -0.669 (-13%)	N=254 5.306	N=193 -0.628 (-12%)	N=169 -2.193 (-41%)
ADCS-ADL-MCI						
All observations	N=331 -3.8	N=330 0.7 (-18%)	N=298 0.7 (-18%)	N=283 -4.3	N=286 0.7 (16%)	N=295 1.7 (-40%)
Post-ARIA observations excluded	N=296 -3.6	N=239 0.5 (-14%)	N=183 0.3 (-8%)	-4.3	N=190 0.5 (-12%)	N=171 2.6 (-60%)

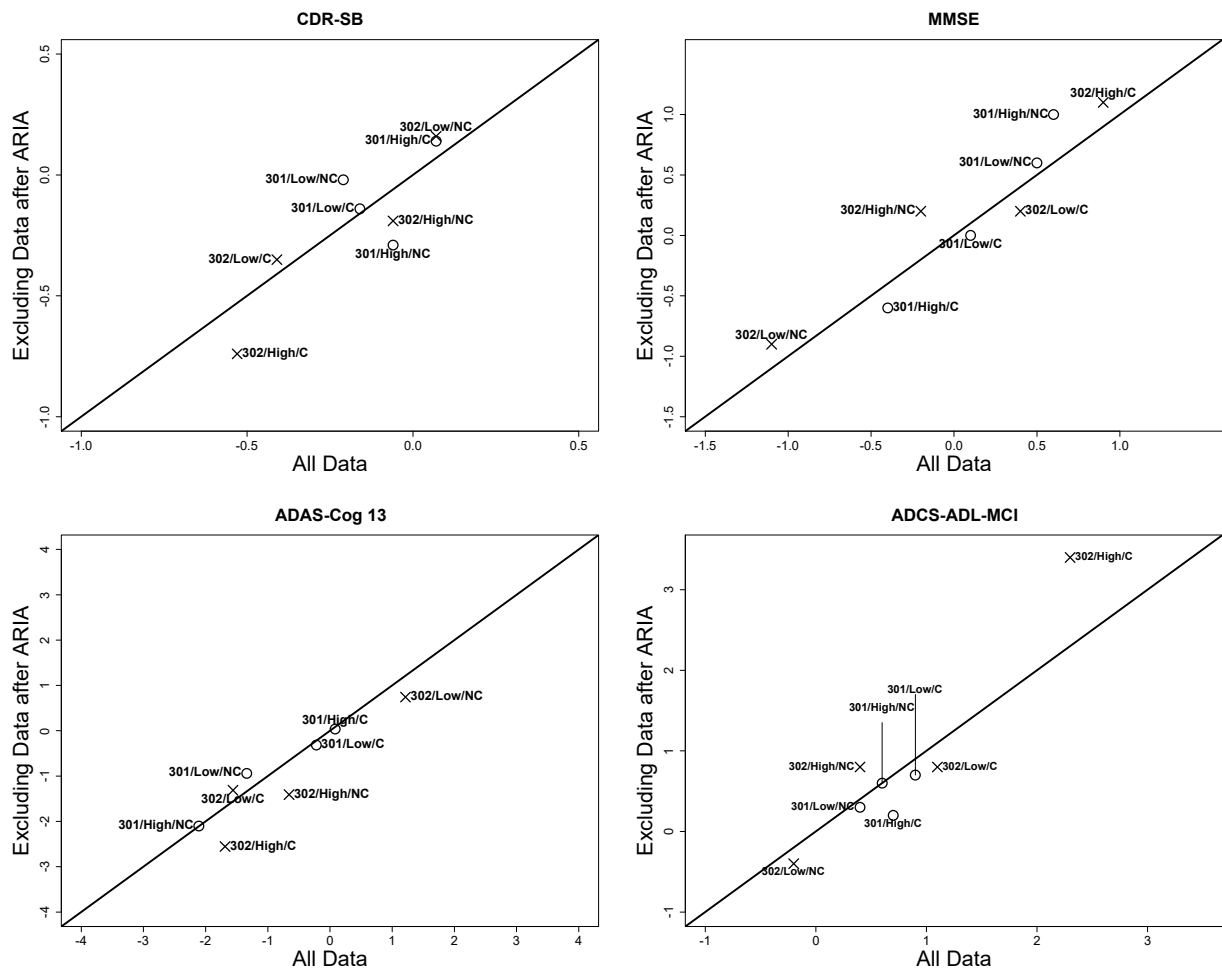
Source: Table 33 in ISE and Output # 128, 129 and 130 in ISE Appendix F

The applicant also assessed the treatment difference at Week 78 for groups defined by dose, study, and ApoE ε4 carrier status including and excluding observations after occurrence of ARIA. The results are presented graphically in Figure 16. If a systematic bias were present, one would expect the points to consistently fall on one or other side of the line of unity. Instead, the results show an even scatter around the line of unity, suggesting no systematic bias due to ARIA.

The potential impact of ARIA is inherently confounded with dosing through protocol-specified dose modifications due to ARIA. Therefore, it is challenging to independently assess the impact

of ARIA and dosing. Approaches to address the potential effect of dosing on outcomes will be further discussed in the following section.

Figure 16: Treatment Difference of Change from Baseline in CDR-SB, MMSE, ADAS-Cog 13 and ADCS-ADL-MCI at Week 78 Grouped by Study, Dose and ApoE Carrier Status and Including or Excluding Post-ARIA Observations. (× - Study 302, o – Study 301, C – carrier, NC – con-carrier; solid line is line of unity)



Source: Output # 219-126 in ISE Appendix F

Influence of Dosing

Dosing was an obvious area of investigation to understand the discordant results for the high dose arms between Study 301 and Study 302. The critical role of dosing for amyloid-targeting therapies became increasingly evident during the aducanumab clinical program. The lack of

adequate dosing was acknowledged as a contributing factor to the failure of previous clinical trials.

The importance of dose was also directly established in the aducanumab program by Study 103 which demonstrated a dose-dependent reduction in brain amyloid and reduction of decline on clinical outcome measures.

The motivation to investigate the potential importance of dosing in the interpretation of Studies 301 and 302 is further supported by the conduct of these clinical studies, including the following considerations:

- The protocols for Studies 301 and 302 were amended twice during the studies to increase the dose or to allow for patients to continue dosing:
 - Protocol Version 3 modified ARIA management rules so that patients were more likely to be able to continue dosing or achieve target dose levels.
 - Protocol Version 4 increased the target dose for ApoE ε4 carriers in the high-dose arm from 6 mg/kg to 10 mg/kg.
- The timing of the studies and pace of enrollment was such that Study 302 was more likely to benefit from changes to the protocol than Study 301.
- The timing of the protocol amendments was such that patients enrolled later in the studies were more likely to achieve higher aducanumab exposures.

In fact, before any investigation of the results of Studies 301 and 302 began, the Division asked the applicant about the role of dosing in preliminary comments to the June 14, 2019, Type C Meeting.

Dosing in Studies 301 and 302

One challenge in comparing dosing over time or between studies is identifying an appropriate dosing metric. Steady-state is achieved after 4 consecutive doses and dose interruptions can have long-term consequences on amyloid reduction (Figure 14). There was also heterogeneity in the dosing profiles due to changes in the protocol and dose modifications due to adverse events. It may be possible that achieving consistent exposure early in the study is more important than reaching the target dose later in the study. It is worth noting that patients receiving lower doses had more consistent dosing compared to patients receiving higher doses. For these reasons, multiple measures of dosing were characterized, including cumulative dose, number of doses at target dose, and number of uninterrupted doses at target dose.

Through the implementation of protocol amendments, exposure to aducanumab increased over the course of Studies 301 and 302. Protocol Version 3 modified ARIA management rules so that patients who experienced ARIA were more likely to continue dosing or achieve target dose levels. Protocol Version 4 increased the target dose for all ApoE ϵ 4 carriers in the high-dose arm from 6 mg/kg to 10 mg/kg. To quantify the actual increase in dose due to these protocol amendments, patients were dichotomized into groups based on the timing of their consent to Protocol Version 4 (PV4) or higher prior to Week 16 (i.e., Pre-PV4 or Post-PV4). Week 16 was chosen because patients who consented by this time had the opportunity to receive all 14 doses of 10 mg/kg according to the protocol amendment. The mean cumulative dose received up to Week 78 for the high-dose arms was lower in the Pre-PV4 group (104.4 mg/kg in Study 301 and 109.5 mg/kg in Study 302) compared to the Post-PV4 group (129.8 mg/kg in Study 301 and 127.8 mg/kg in Study 302). As expected, the mean cumulative dose received up to Week 78 was similar in the Pre-PV4 (58.8 mg/kg in Study 301 and 59.9 mg/kg in Study 302) and Post-PV4 (59.8 mg/kg in Study 301 and 57.8 mg/kg in Study 302) groups for the low-dose arms because Protocol Version 4 did not change dosing for the low-dose arms. In the high-dose arms, the mean number of 10 mg/kg doses received up to Week 78 was lower in the pre-PV4 groups (6.5 in Study 301 and 7.3 in Study 302) compared to the post-PV4 groups (10.8 in Study 301 and 10.5 in Study 302).

Due to the timing of the initiation of the studies and the pace of enrollment, more patients in Study 302 were able to benefit from the amendments in Protocol Versions 3 and 4 (81.3% and 55.9%, respectively) compared with patients in Study 301 (73.4% and 49.1%, respectively). For example, the mean cumulative dose received up to Week 78 in the high-dose arms was 114.7 mg/kg in Study 302 compared with 110.6 mg/kg in Study 301. The difference between studies is more evident at the extremes where 100 patients in Study 302 received all 14 doses of 10 mg/kg compared to 82 in Study 301. The difference in dosing between studies is smaller than the difference in dosing over time in the studies as described in the previous paragraph.

The potential impact of differences in dosing between studies is evident in the amyloid PET analysis populations. The mean cumulative dose received up to Week 78 was higher in the high-dose arm of Study 302 (118.3 mg/kg) compared to Study 301 (109.1 mg/kg). This relatively small difference in dosing corresponded with a 16% smaller reduction in brain amyloid at Week 78 in Study 301 (-0.232) compared to Study 302 (-0.278). It is also noteworthy that the biomarker changes for the high-dose arm in Study 301 were consistently smaller than the changes observed in Study 302.

Response in Dosing Subgroups Based on Stratified Randomization

Modifications in dosing due to protocol changes, ApoE ϵ 4 status, and ARIA management suggest that randomized dose (i.e., low or high) does not fully capture the level of aducanumab exposure. To further investigate the relationship between dose and response, subgroups were

formed for each study based on stratified randomization (low/high dose and ApoE ϵ 4 carrier/non-carrier) and by enrollment (Pre-PV4 vs. Post-PV4) to form a total of 16 distinct groups (Table 38). The advantage of this approach is that it largely preserves randomization and the primary analysis method can be used when estimating treatment differences. Mean cumulative dose at Week 78 was calculated for patients in each subgroup. The expected cumulative dose at Week 78 assuming no dose interruptions is 56 mg/kg, 98 mg/kg, and 160 mg/kg for randomized target dose levels of 3 mg/kg, 6 mg/kg, and 10 mg/kg. The 16 randomized groups were categorized as low, medium, or high in Table 38 and roughly correspond with the target levels. Using this categorization, it becomes evident that some groups in the randomized high-dose arms (i.e., pre-PV4 ApoE ϵ 4 carriers) achieved actual aducanumab dosing similar to groups randomized to the low dose (i.e., ApoE ϵ 4 noncarriers).

Table 38: Dosing in Subgroups Defined by Study, Randomized Dose Group, ApoE ϵ 4 Status and Pre- and Post-Protocol Version 4

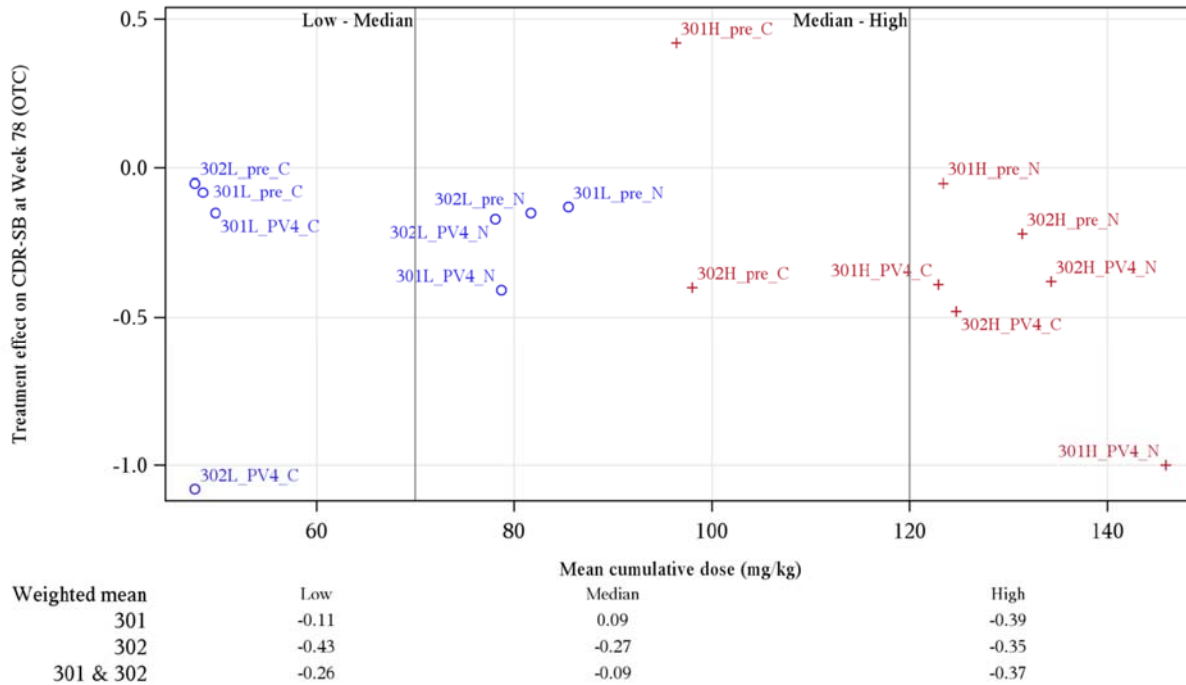
Dosing Regimen (Study/Dose/Protocol Version/ApoE ϵ 4 status)	Mean Cumulative Dose at Week 78 (mg/kg)	Exposure Categorization
301 high pre-PV4 carrier	96.4	Medium
301 high post-PV4 carrier	122.9	High
301 high pre-PV4 non-carrier	123.4	High
301 high post-PV4 non-carrier	145.9	High
302 high pre-PV4 carrier	98.0	Medium
302 high post-PV4 carrier	124.7	High
302 high pre-PV4 non-carrier	131.4	High
302 high post-PV4 non-carrier	134.3	High
301 low pre-PV4 carrier	48.5	Low
301 low post-PV4 carrier	49.8	Low
301 low pre-PV4 non-carrier	85.5	Medium
301 low post-PV4 non-carrier	78.7	Medium
302 low pre-PV4 carrier	47.7	Low
302 low post-PV4 carrier	47.7	Low
302 low pre-PV4 non-carrier	81.7	Medium
302 low post-PV4 non-carrier	78.1	Medium

Adapted from Table 38 in ISE
PV=Protocol Version

Mean differences of CDR-SB between aducanumab treatment and placebo were calculated for each of the 16 subgroups listed in Table 38 and are plotted against mean cumulative dose in Figure 17. The relationship between mean cumulative dose and treatment effect is weak with considerable variability between the groups. On the other hand, it is noteworthy that the weighted mean treatment effect among randomized groups from Study 301 in the “high” exposure category is -0.39. This estimate is consistent with the weighted mean estimate from Study 302 in the “high” exposure category (-0.35) and with the primary analysis of the high-dose arm Study 302 in (-0.39). These results suggest that groups of patients in Study 301 who

had the opportunity to receive higher exposure to aducanumab show a response to aducanumab treatment similar in character to Study 302.

Figure 17: Treatment Difference in CDR-SB at Week 78 Vs. Mean Cumulative Dose by Study, Randomized Dose Group, ApoE ε4 Carrier Status and Protocol Version



Source: Figure 41 in ISE

Notes: Blue circles are groups randomized to low dose. Red + are groups randomized to high dose. The label for each point represents the study, dose (L = low, H = high), Protocol Version (pre = pre-PV4, PV4 = post-PV4), ApoE ε4 status (C = carrier, N = non-carrier).

Response in Dosing Subgroups Based on Individual Dosing

Within the randomized subgroups presented in the previous section there may still be meaningful heterogeneity with respect to individual aducanumab dose or exposure. Therefore, in a separate analysis, subgroups were defined by levels of individual aducanumab dosing. Analyses were conducted on the subset of patients who had the opportunity to complete 78 weeks of treatment so as not to confound cumulative dosing with time in the study. A disadvantage of this approach, however, is that it primarily includes patients enrolled early in the study and does not fully address changes in exposure that occurred over time. The purpose of the following analyses was to test the hypothesis that patients in Study 301 who were exposed to higher levels or doses of aducanumab were more likely to show benefit. An initial view of the data (Table 39) from the high-dose arm in Study 301 suggests a trend between increasing cumulative dose and numerically greater treatment difference vs. placebo. Placebo here corresponds to the overall placebo response.

Table 39: Treatment Difference in CDR-SB at Week 78 by Cumulative Dose Threshold for High Dose in Study 301

Cumulative Dose Threshold (mg/kg)	Number of Subjects	Treatment Difference vs. Placebo
< 100*	401	-4%
≥ 100	223	4%
≥ 110	206	1%
≥ 120	178	-6%
≥ 130	157	-10%
≥ 140	133	-14%
≥ 150	98	-24%
≥ 160	75	-30%

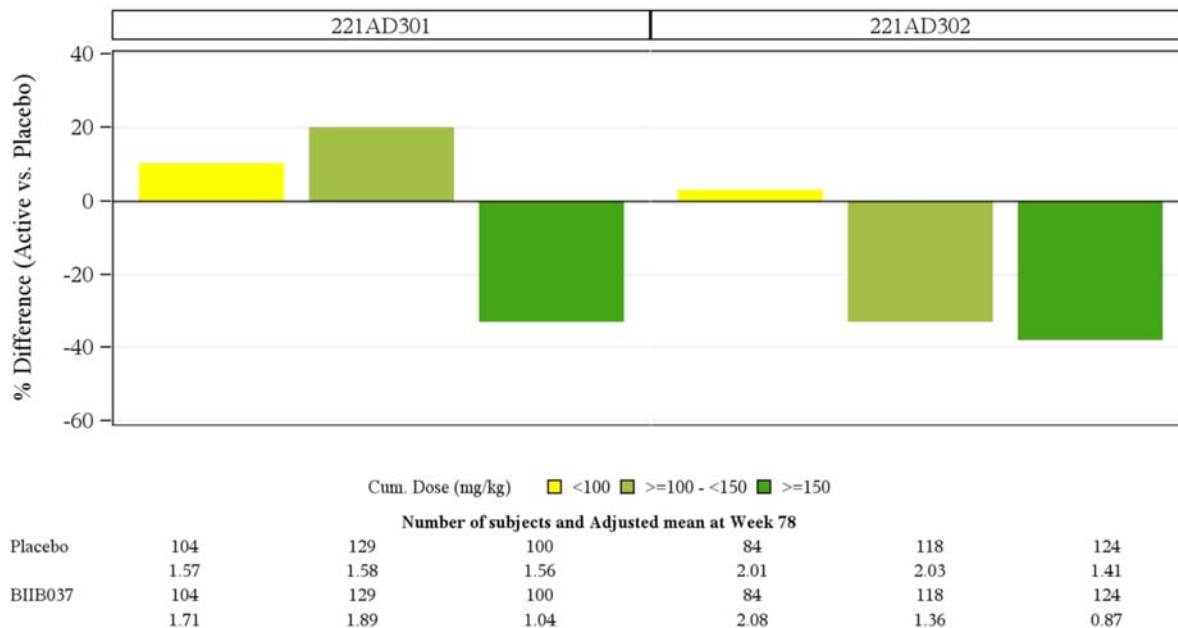
Source: Appendix G7 Table 16 in ISE

* Patients in high dose were combined with low dose group

A limitation of this approach is that subgroups are created by post-randomization outcomes (i.e., dose) and may not be balanced on other relevant factors. For example, the ≥160 mg/kg group in Table 39 is made up of approximately 75% ApoE ε4 non-carriers whereas the overall placebo group is approximately 33% ApoE ε4 non-carriers. Therefore, a propensity score matching analysis was used to balance placebo and aducanumab groups on baseline demographic and disease characteristics. To explore the relationship between individual aducanumab dosing within the high-dose arm and treatment response, three independent groups of low, intermediate, and high exposure were formed. Three groups were chosen so that each group included approximately ≥100 aducanumab-treated patients.

The treatment differences at Week 78 for CDR-SB between aducanumab and propensity score matched placebo subgroups based on cumulative dose are illustrated in Figure 18. The results for Study 302 show the expected pattern with subgroups exposed to higher cumulative doses exhibiting the largest treatment effect and the subgroup with limited exposure showing a response similar to placebo. For Study 301, the high cumulative dose subgroup performs similarly to Study 302. The discrepancy is primarily in the intermediate exposure subgroup which appears to show an advantage of placebo over aducanumab. Dosing cannot explain this observation.

Figure 18: CDR-SB Adjusted Mean Change from Baseline % Difference from Propensity-Matched Placebo at Week 78 in Subgroups by Cumulative Dose in Studies 301 and 302

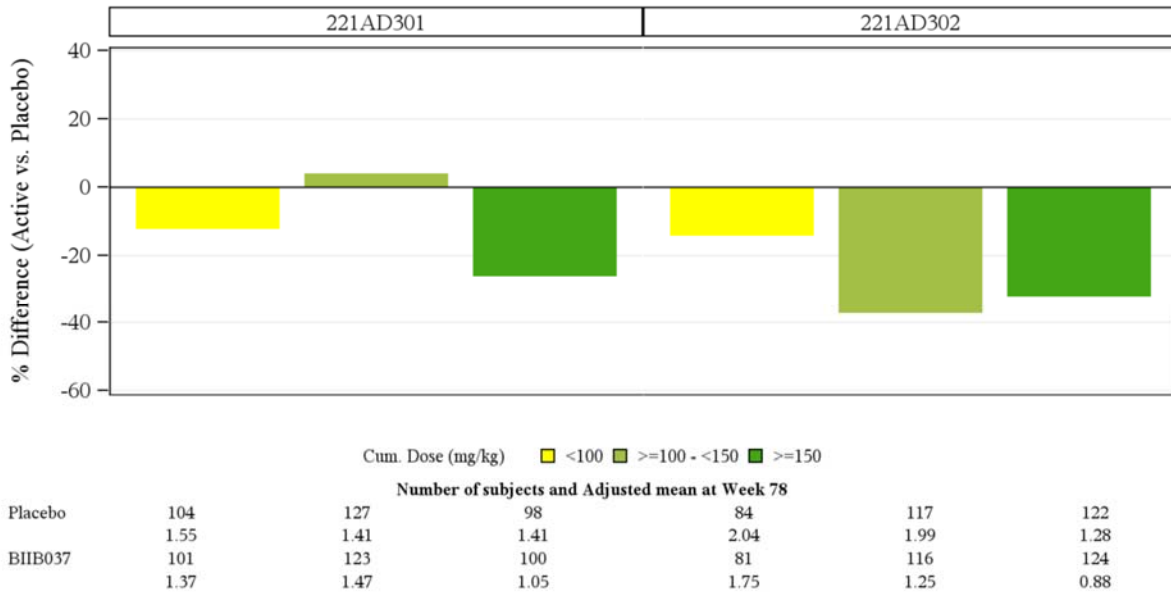


Source: Figure 52 in ISE

* BIIB037 refers to aducanumab

Previous analyses have demonstrated the outsized effect rapid progressors can have on interpretation of treatment differences. This effect is also notable when interpreting subgroups of data. Figure 19 shows the results of the same analyses, but with the rapid progressors removed. Again, the high cumulative dose group in Study 301 continues to indicate a favorable response to aducanumab treatment in line with the corresponding group in Study 302. This finding is robust to the existence of rapid progressors. Instead of showing a favorable response for placebo compared to aducanumab, the intermediate exposure group indicates a response similar to placebo.

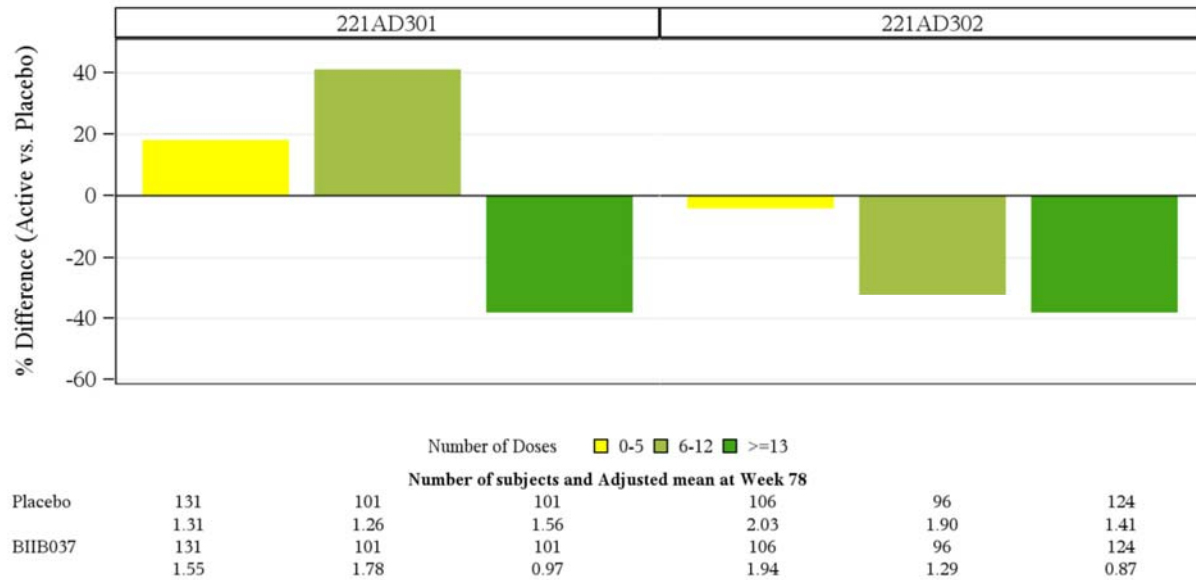
Figure 19: CDR-SB Adjusted Mean Change from Baseline % Difference from Propensity-Matched Placebo at Week 78 in Subgroups by Cumulative Dose in Studies 301 and 302 (Rapid Progressors Removed)



Source: Figure 53 in ISE
* BIIB037 refers to aducanumab

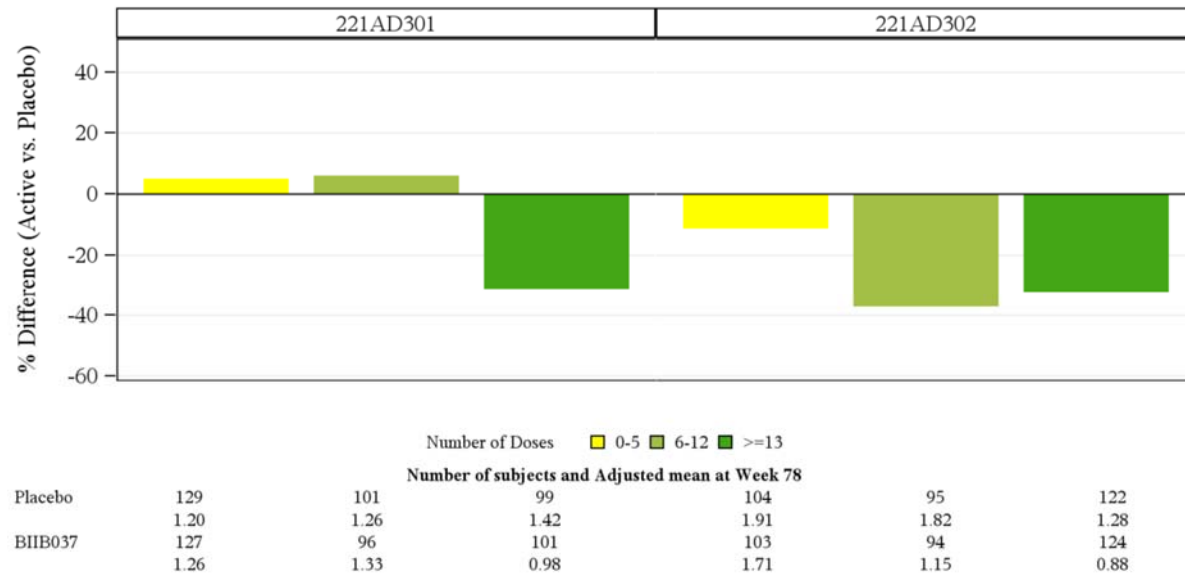
The same analyses were also performed using the number of uninterrupted steady-state doses at 10 mg/kg and the number of doses at 10 mg/kg without regard for interruption as the dosing metrics. The results using the number of doses at 10 mg/kg without regard for interruptions are illustrated in Figure 20 and Figure 21 (with rapid progressors removed) and show a similar pattern to results using cumulative dose.

Figure 20: CDR-SB Adjusted Mean Change from Baseline % Difference from Propensity-Matched Placebo at Week 78 in Subgroups by Number of 10 mg/kg Doses in Studies 301 and 302



Source: Figure 49 in ISE
* BIIB037 refers to aducanumab

Figure 21: CDR-SB Adjusted Mean Change from Baseline % Difference from Propensity-Matched Placebo at Week 78 in Subgroups by Number of 10 mg/kg Doses in Studies 301 and 302 (Rapid Progressors Removed)



Source: Figure 50 in ISE
* BIIB037 refers to aducanumab

It is noteworthy that the treatment difference observed in the high exposure group in Study 301 does not appear to be driven by the placebo response. The adjusted mean change from baseline in this group in Study 301 is approximately 1 unit, similar to the approximately 0.88 change observed in Study 302. It is still possible that patients in the high exposure group in Study 301 have other characteristics not considered in the propensity score analysis that make them more likely to have a favorable response. For example, the fact that they remained in the study and received most of their expected doses might also make them more likely to be responders. As a supplementary analysis, only placebo patients in the per-protocol population (i.e., received $\geq 70\%$ expected infusion) were included in the propensity matching pool. A favorable treatment effect was still observed in the high exposure group in Study 301.

The difference in dosing between the two studies is also illustrated in Figure 20. A total of 124 patients in Study 302 received ≥ 13 doses of 10 mg/kg compared to 101 patients in Study 301. Conversely, more patients (131) in Study 301 received 0-5 doses of 10 mg/kg compared to Study 302 (106 patients). As noted before, this difference in dosing between studies is modest. Simple arithmetic calculations suggest that if dosing in Study 301 had been similar to Study 302, the overall treatment effect from the primary analysis in Study 301 would only improve by 1-2%.

The primary discrepancy between Study 301 and Study 302 in these analyses is the group of patients with cumulative dose or number of 10 mg/kg doses in the intermediate range. It is worth noting that this group includes the most heterogeneity in terms of dosing profiles. Even with this heterogeneity, pharmacokinetic and PET data indicate that the aducanumab exposure and amyloid reduction in this subgroup is in excess of the low-dose arm in the study. Therefore, if one accepts that the low doses in Studies 301 and 302 consistently demonstrate a numerical effect in favor of aducanumab, one would expect to observe at least a numerically favorable outcome in this intermediate exposure subgroup in Study 301. Given the multitude of findings suggesting a dose-response relationship and a directionally favorable effect of aducanumab, the fact that this is not the case in this one subset raises the possibility that this observation may simply be due to chance or some other unknown factor unrelated to aducanumab.

Overall, the analyses on dose indicate the following:

- Dosing is an important consideration for interpretation of the efficacy result of aducanumab in Studies 301 and 302
- Patients in Study 301 with higher exposure to the 10 mg/kg dose demonstrated treatment effects similar to patients in Study 302
- Lower exposure to the target dose of 10 mg/kg in Study 301 was a small but contributing factor for the discordant results with Study 302. Factors other than dosing may also contribute to the difference

It may be tempting to misinterpret the considerations presented in Section 6.1.7 as “explaining why Study 301 was negative.” This was not the intention of this work and such an explanation is not necessary to establish the effectiveness of aducanumab. At the June 14, 2019, Type C Meeting, the Division clearly stated that, “available data do not suggest the future use of Study 301 as an efficacy study providing independent evidence of effectiveness supporting the approval of aducanumab.” Rather, the Division noted that analyses “may have a role in supporting the results of Study 302,” or “may be understood well enough... to not represent evidence that the drug is ineffective.” The analyses presented in this section are exploratory by design but limited in scope and focused on pre-defined areas of interest. The rapid progressor analysis indicated that a small imbalance in the number of rapid progressing patients in the high-dose arm in Study 301 had a disproportionate impact on the estimate of the treatment effect using the primary analysis method. An examination of dosing in Study 301 indicates that patients with higher exposure to the 10 mg/kg dose in Study 301 had similar responses to patients in Study 302. These two factors contribute to the overall understanding of Study 301 and together indicate that the results of Study 301 do not meaningfully detract from the persuasiveness of Study 302. There were no findings from the exploration that represented evidence that aducanumab is clearly not effective.

6.2. Integrated Assessment of Effectiveness

Taken together, the evidence supporting the effectiveness of aducanumab is highly persuasive. In short, this evidence is provided in large part by Study 302, which provides important context for an additional and not insubstantial contribution from Study 103. The results of Study 301 are sufficiently well understood that they do not preclude independent consideration of the results of Study 302 and 103.

In order to consider the evidence provided by Study 302 and Study 301, it was first necessary to assess the impact of early termination of the nearly completed studies on interpretability of the observed efficacy data and associated analyses. Modeling and simulation analyses were used to rigorously evaluate the likely outcomes of the studies had they been run to completion. The results of these analyses indicated that the simulation results were highly consistent with the primary analysis of the observed data, establishing the interpretability of the studies and the appropriateness of the observed data prior to early termination for further consideration. Accordingly, Study 302, analyzed in accordance with the prespecified analysis, and supported by multiple additional analyses, is a strongly positive study.

The effect of the high dose of aducanumab in Study 302 is robust and exceptionally persuasive on several of the instruments used to evaluate efficacy. It provides highly persuasive results on the prespecified primary endpoint intended to establish effectiveness and on the prespecified secondary endpoints and exploratory tertiary endpoint, including particularly persuasive results on secondary endpoints which themselves represent an acceptable approach to establishing

effectiveness. Its findings are robust to numerous sensitivity analyses and are strongly supported by highly consistent findings in prespecified subgroups and well-characterized biomarker data. Beneficial effects on clinical measures are further supported by evidence suggesting a dose-response relationship on clinical outcomes and by evidence of a dose- and time-dependent relationship on biomarkers of fundamental Alzheimer's disease pathophysiology, including brain amyloid burden, the primary direct marker of aducanumab's intended mechanistic effect. In short, when considered individually, Study 302 is an exceptionally persuasive study and has many characteristics indicating that it is capable of independently providing substantial evidence of effectiveness.

The results of Study 103 provide additional data addressing the effectiveness of aducanumab. Although designed to primarily assess safety and tolerability rather than effectiveness and uncontrolled for multiplicity, the 10 mg/kg dose arm was able to achieve statistical significance according to the prespecified analysis plan on both of the clinical efficacy outcomes assessed. A supplementary analysis of Study 103 using the same primary analysis method that was used in Studies 301 and 302 provides similar efficacy results for the 10 mg/kg dose arm and is supportive of the prespecified findings. Further, the dose-response relationship for A β reduction provides support for the positive 10 mg/kg efficacy findings and is consistent with the dose-response relationship observed for the efficacy outcome measures. The design of Study 103 was consistent with Study 302 in several aspects, including choice of clinical and biomarker outcome assessments. Clinical effects in the 10 mg/kg dose arm in Study 103 and the high dose arm in Study 302 favored aducanumab and biomarker effects in those arms were similar. In the context of the results of Study 302, the results of Study 103 are appropriately viewed as an additional and supportive contribution to substantial evidence of effectiveness.

With Study 302 capable of providing the primary contribution and Study 103 capable of providing an additional contribution to a demonstration of substantial evidence of effectiveness, it is important to understand Study 301 in depth in order to decide if Study 301 detracts from the persuasiveness of that evidence, with particular attention to the evidence provided by Study 302, given that it shares its design with Study 301. An important initial observation of Study 301 is that primary and secondary endpoints for the low dose had responses that were numerically favorable and similar in magnitude to those in Study 302. The response of the high dose was the notable difference between the studies. Explorations of demographic and baseline characteristics and the incidence of ARIA did not reveal differences that could be responsible for the difference in high dose results. An exploratory analysis of rapid progressors indicated that a small imbalance in the number of rapid progressing patients in the high-dose arm in Study 301 had a disproportionate impact on the estimate of the treatment effect using the primary analysis method. An examination of dosing in Study 301 indicates that patients with higher sustained exposure to the 10 mg/kg high dose in Study 301 had similar responses to patients in Study 302. The results of these exploratory analyses

contribute to the overall understanding of Study 301 and together do not meaningfully detract from the persuasiveness of Study 302.

Based on the considerations above, the applicant has provided substantial evidence of effectiveness to support approval. Study 302 provides the primary evidence of effectiveness as a robust and exceptionally persuasive study demonstrating a treatment effect on a clinically meaningful endpoint and reinforced by effects on secondary endpoints, biomarkers, and in relevant subgroups. Study 103 was an adequate and well-controlled study which included design components consistent with Study 302 and demonstrated a persuasive treatment effect on both clinical endpoints, providing additional and supportive evidence of effectiveness. The dose-response relationship for A β reduction provides support for the positive finding in the 10 mg/kg treatment arm and to the apparently dose-related effects observed on clinical outcomes in Studies 103 and 302.

8. Appendix 2



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 #: 761178

Drug Name: Aducanumab

Indication(s): Alzheimer's

Applicant: Biogen

Date(s): July 7, 2020

Review Priority: 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 Reviewer: Kevin Krudys, Ph.D., Brian Trummer, M.D., Ph.D., Natalie Branagan, M.D.

Project Manager: Emilios (Andrew) Papanastasiou

Keywords: substantial evidence; futility

Table of Contents

1	EXECUTIVE SUMMARY	7
2	INTRODUCTION	11
2.1	OVERVIEW	11
2.2	DATA SOURCES.....	12
3	STATISTICAL EVALUATION	12
3.1	DATA AND ANALYSIS QUALITY	12
3.2	EVALUATION OF EFFICACY	13
3.2.1	<i>Study 302</i>	13
3.2.1.1	Study Design and Endpoints	13
3.2.1.2	Statistical Methodologies.....	14
3.2.1.3	Patient Disposition, Demographic and Baseline Characteristics	19
3.2.1.4	Results and Conclusions.....	22
3.2.1.4.1	Sponsor's Results	22
3.2.1.4.2	Reviewer's Results	24
3.2.2	<i>Study 301</i>	52
3.2.2.1.1	Sponsor's Results	52
3.2.2.1.2	Reviewer's Results	52
3.2.3	<i>Study 103</i>	60
3.2.3.1.1	Sponsor's Results	66
3.2.3.1.2	Reviewer's Results	67
3.3	EVALUATION OF SAFETY	73
4	FINDINGS IN SPECIAL/SUBGROUP POPULATIONS	74
4.1	GENDER, RACE, AGE, AND GEOGRAPHIC REGION.....	74
4.1.1	<i>Gender, Race, and Age</i>	74
4.1.2	<i>Geographic Region</i>	77
4.1.2.1	Individual Sites	83
4.2	OTHER SPECIAL/SUBGROUP POPULATIONS	85
5	SUMMARY AND CONCLUSIONS	88
5.1	STATISTICAL ISSUES	88
5.2	COLLECTIVE EVIDENCE	95
5.3	CONCLUSIONS AND RECOMMENDATIONS	96
I.	APPENDIX.....	96
I.1	ADDITIONAL ISSUES	96

LIST OF TABLES

Table 1 Efficacy Study Characteristics.....	11
Table 2 Patient Disposition in the Phase 3 Trials	20
Table 3 Baseline Demographics for Studies 301 and 302	21
Table 4 Baseline Disease Characteristics for Studies 301 and 302	22
Table 5 Phase 3 Primary and Key Secondary Results.....	23
Table 6 Study 302 Demographic Characteristics by pre-PV4 and post-PV4	26
Table 7 Study 302 Week 78 CDRSB Analyses Exploring Exclusion of Data after post-Baseline starting of AD medications	27
Table 8. Study 302 Estimated Treatment Differences at Week 78 by APOE for Primary and Key Secondary Endpoints	42
Table 9. Study 301 Post-PV4 subset High Dose vs Placebo CDRSB Treatment Effect Estimates by Age Group	55
Table 10. Study 301 Post-PV4 High -Placebo Estimated CDRSB Difference at Week 78 by APOE subgroups	55
Table 11 Sponsor's Completer's Analysis of Study 103	68
Table 12 Study 103 Randomization Supported and Sponsor's non fully randomization supported Analyses	69
Table 13 Adverse Events With at Least 5% Incidence in BIIB037 10 mg/kg and 2% Higher Incidence Than Placebo –Pool A1	74
Table 14 Differences on CDRSB by Sex Subgroup	76
Table 15 CDRSB Hi Placebo Differences on CDRSB at Week 78 by Age Group and Study	77
Table 16 Pooled and By Study Analysis of Estimated High Dose Treatment Effects by Race Groups for CDRSB at Week 78	78
Table 17 Enrollment by Country across Studies 301 and 302.....	78
Table 18 CDRSB Hi vs. Placebo Results at Week 78 Averaged over 301 and 302 by Country	84

Table 19 Study 302 Hi vs. Placebo LSMean differences on CDRSB at Week 78 after excluding a site of interest.....	85
Table 20 Study 302 Estimated CDRSB Change Differences from Placebo at Week 78 by Baseline Stage Diagnosis	87
Table 21 CDRSB at Week 78 for High vs Placebo by APOE and Study and Pooled	88
Table 22 Study 302 Correlations at Week 78 between changes from baseline on Primary and Key secondary endpoints.....	93

DRAFT

LIST OF FIGURES

Figure 1 Probability of Randomization of Mild (1) vs. Prodromal (0) Baseline Disease Stage as Study 302 progressed	25
Figure 2. CDRSB changes at Week 78 in those with opportunity to complete (studies 301 and 302 pooled)	28
Figure 3. Study 301 CDRSB Changes at Week 78 (excluding missing data but deaths coded as missing)..	29
Figure 4. Study 302 CDRSB changes at Week 78 (excluding missing data).....	30
Figure 5. Study 302 Placebo, Low Dose, and High dose CDRSB profiles by Pre-PV4 and-Post PV4 status.	32
Figure 6. Study 302 Placebo, Low Dose, and High Dose CDRSB Profiles by Pre-PV4 and Post-PV4	33
Figure 7. Local Trend in CDRSB Week 78 Changes over Study Duration and by PV4	34
Figure 8. Local Trend in CDRSB Week 78 Changes over Study Duration and by PV4	35
Figure 9. Probability of Asian Enrollment over Time in Study 302	36
Figure 10. Study 302 Week 78 LS Mean CDRSB Change by ARIA Dose Modification Status and pre/post PV4 in APOE+	38
Figure 11. Differences in estimated Week 78 CDRSB High Dose effects by Country in Study 302	39
Figure 12. Study 302 Key Endpoints APOE subgroup treatment effects estimates by Dose	41
Figure 13 Bar plot of CDR Sum of Boxes Adjusted Mean Change from Baseline Percent Difference from Placebo at Week 78 by Number of 10 mg/kg Doses, with Placebo Selected by Propensity Score Matching - ITT Population that have had Opportunity to Complete Week 78 by 20Mar2019: Placebo-controlled Period Excluding Data after 20Mar2019 – Nested Categories.....	44
Figure 14. Study 302 CDRSB Tipping Point Analysis.....	45
Figure 15. Assessment of Correlation between Dose Achieved and Week 78 CDRSB across Low and High Doses (not placebo subtracted)	46
Figure 16. Assessing Correlation of Amyloid Pet and CDRSB in High Dose at Week 78	48

Figure 17. Correlation between Week 78 CSF P _{Tau} change and Week 78 CDRSB change within High Dose in 302	50
Figure 18. Week 78 Change in Medial Temporal Tau by Mean Dose.....	52
Figure 19. Study 301 High Treatment Effect Estimates at Week 78 pre-PV4 and post-PV4	54
Figure 20. Study 301 Placebo, Low Dose, and High Dose CDRSB Profiles in Study 301 APOE non-carriers Stratum	56
Figure 21 Study 301 Placebo, Low Dose, High DOse CDRSB profiles in Study 301 APOE carrier Stratum .	57
Figure 22. Study 301 LS Mean Change from Baseline in CDR-sob at Week 78 in APOE+ by PV4 and ARIA dose modification	58
Figure 23 Study 301 local trend in CDRSB change at Week 78 over randomization time.....	60
Figure 24 Study 103 Staggered Arm Design.....	63
Figure 25 Study 103 Patient Disposition	67
Figure 26 Study 103 CDRSB profile for 10 mg/kg groups and Overall Pooled placebo	71
Figure 27 Baseline CDRSB by Arm showing Staggered Design	73
Figure 28 Study 302 Placebo and High Dose LS means for CDRSB by Country.....	81
Figure 29 Forest Plot of Change in CDRSB at Week 78 by Country and Other Subgroups	83
Figure 30 Study 302 CDRSB Change at Week 78 by Site.....	85
Figure 31 Study 302 Estimated High vs Placebo Differences at Week 78 on CDRSB by Baseline Disease Stage	87
Figure 32 Scree Plot of Principal Components of Primary and Key Secondary Endpoints	94

1 EXECUTIVE SUMMARY

The two phase 3 studies were stopped early for futility (March 21, 2019 press announcement) when both studies had reached 50% completion (thus, in a sense, together equivalent in information to a completed study) since it was estimated based on the interim study-pooled estimate of the treatment effects that both studies had <20% chance of success for either dose if completed. Following a futility press release announcement and collection of subsequent study closeout follow up data, the sponsor requested a meeting to discuss the two trials final data after discovering that despite the futility conclusion, the final analysis on face showed a statistically significant effect for the high dose in one of the two trials ($p=0.01$) but not the other ($p=0.83$).

Inconsistency on many levels summarizes the final clinical efficacy data from these trials. Because the two phase 3 studies were terminated for futility the NDA package doesn't contain a single phase 3 study that was fully completed according to the plan. In fact, almost 50% are missing the Week 78 time point assessment of CDRSB which is the only timepoint that shows any significance and that is only significant in one of the two studies (the first study high dose is numerically worse than placebo at Week 78 on the primary endpoint). A worse placebo response in study 302 than was observed in study 301 could explain the significance of study 302 ($p=0.01$).

There is a reason why two positive studies has been the standard in Alzheimer's, e.g., the need for reproducibility and adequate strength of evidence, in a disease with soft (more subjective and variable than mortality) endpoints. This BLA submission does not have a situation such as just one study in existence and for which that study is strong. We have a second large adequate well controlled study that directly contradicts the first and is not even close to significance $p=.8252$. Under the null hypothesis (no drug effect), there is a .0975 chance of at least one type I error across 2 studies. If one has two studies and takes the best and pretends like it's the only study, one's estimate is most likely biased and misleading. In the opportunity to complete subset of 302 the high dose vs. placebo has a p-value of 0.0368 for the CDRSB at Week 78 and even in the ITT there was no significance before Week 78. It is not justifiable to search for patients in 301 who are similar to 302 because that may have selection bias and presumes that 302 is right and 301 is wrong, for which there is no justification not relying on post-hoc analyses. Any selection of patients would need a proper placebo control, that is the regulatory standard in Alzheimer's. The overall 301 primary result is the only valid well controlled, multiplicity adjusted, randomization validated analysis of 301 (and it had a substantial sample size).

The sponsor tries to discount study 301 due to post-hoc defined "rapid progressors". Rapid progressors are part of the reality of Alzheimer's and after the fact it is too late to address them in a completed large randomized study. A highly effective drug would not be likely to fail because of rapid progressors especially in the early stages of a disease. Study 302 could just as well be the outlier relative to the true proportion of outliers in the natural progression. In fact, the range of CDRSB changes in Study 301 at 18 months appears consistent with the ADNI study ([adnimerge_May15.2014 data](#)). There are slightly more outliers in the high dose in 301 but that

is worrisome in itself since they are consistent with the ADNI data and so should again raise doubts about the representativeness of the 302 result. Furthermore, robust regression, techniques (M estimation, least trimmed squares, MM estimation, S estimation) designed to be resistant to and downweight outliers, applied to the 301 Week 78 data still suggest no effect of the high dose compared to placebo and that it was numerically worse than the low dose (vs. Pl $+0.0024 \pm 0.1159$ $p=0.9836$; vs. Low $+0.0322 \pm 0.1154$ $p=0.7802$). Without the worst change of +13 in the high dose group the 301 high dose vs. placebo result is $+0.0267 \pm 0.1495$ $p=0.8581$ as compared to $+0.0316 \pm 0.1499$ $p=0.8330$ including it. This shows that Study 301 is a big study and one outlier patient has limited influence. Totally excluding the patient instead of just the Week 78 observation the result is $+0.0072 \pm 0.1487$ $p=0.9615$, still in the wrong direction. Even excluding the 3 worst outliers for the high dose, the high dose is still nowhere near significant in Study 301. More than one outlier in the high dose is more of a systemic problem and should be more worrisome and harder to discount. The sponsor also tries to use 301 to find a subgroup similar to Study 302, i.e., a subgroup showing efficacy in 301 but this relies on post-hoc non-randomized comparisons (Figure 13, page 44). These analyses hide the fact that the post-hoc matched placebo progresses faster as the number of 10 mg/kg doses increases in these post-randomization event defined subgroups and such post-hoc matching can never equal a true randomization backed analysis. The only valid analysis of Study 301 is the prespecified randomization supported analysis of study 301 which failed for the high dose ($p=0.83$) and this Study outcome should not be discounted without an extremely compelling reason (which there is not).

The sponsor's argues, relying on non-randomized comparisons, that the high dose arm was challenged by intermediate dosing rather than full dosing in some patients. This can be countered by the fact that the low dose was numerically better than the high dose in Study 301, a comparison supported by randomization. Furthermore, the APOE non-carriers having less treatment effect on all four primary and key secondary efficacy endpoints despite having 10 mg/kg dosing from study start and less ARIA than APOE carriers, so fewer dose reductions due to ARIA. In study 302 the estimated effect in APOE- non-carriers on the primary endpoint is -0.08 with $95\%CI=[-0.598, +0.435]$. The high dose APOE non-carriers were also numerically worse than placebo in Study 302 on the first key secondary endpoint MMSE (treatment by APOE interaction term $p=0.0065$). In the APOE+ subgroup which seems to drive study 302 the high dose was in the wrong direction overall in study 301 (APOE+ 0.07014 ± 0.1802 [std err], $p=0.6971$). The study 302 success could be explained by a higher placebo progression after the implementation of protocol amendment 4 while the study was ongoing (Figure 5, page 32). This amendment increased the dose for APOE carriers from 6 mg/kg to 10 mg/kg but resulted in more drug related ARIA adverse events with attendant individual patient dose titration modifications and unblinding (including some sponsor personnel) for the sake of dose managing (up to 35% of high dose patients had dose modifications). In 302 the occurrence of ARIA

adverse events in the high dose was 1.4 times higher for APOE+ vs. APOE- prior to amendment 4 and 2.3 times higher post PV4. Limitation of dose titration in the high dose was 2.1 times higher for APOE+ prior to PV4 and 3.7 times higher post PV4. Thus, unblinding for dose managing would have been higher after PV4. The APOE- stratum high dose had more 10 mg/kg doses but was worse on average than APOE+ in 4 out of the 4 primary and key secondary endpoints (and the low dose shows the same pattern).

In the original “final” data presented to the Agency in June 2019, in Study 302 the MMSE had a p-value for the high dose of 0.062 which would mean that no secondary endpoints in 302 would be significant following the prespecified hierarchy and multiplicity adjustment plan. In particular, the analysis plan suggests that the testing sequence was to compare all doses before moving to the next endpoint, which results in the same conclusion for secondary endpoints because the low dose is not significant for any of the primary of key secondary endpoints (SAP excerpt: “for **each** of the secondary endpoints, a sequential (closed) testing procedure”).

Given the large amount of missing data in the final ITT dataset (>40% per group) and much lower rate missing in the Opportunity to Complete dataset, some different demographics and disease characteristics in those without the opportunity to complete (due to futility stopping) that are related to outcome and not incorporated in the primary model (see discussion in section 3.2.1.4.2), the latter OTC dataset seems more relevant and reliable. The result for the Opportunity to Complete Dataset (total N=953) in Study 302 for the high dose on the primary endpoint, CDRSB at Week 78, was -0.36 , 95% CI= $[-0.70, -0.02]$, $p=0.0368$.

The sponsor’s analysis of Study 103, 10 mg/kg vs. pooled placebo arms, is not supported by the randomization (3 of the placebo arms had no chance of receiving 10 mg/kg and one was entirely APOE carriers, which 10 mg/kg was not). Outside of rare diseases there is no justification for an analysis involving the pooling of staggered arms that is not supported by the study’s overall randomization scheme. The comparisons that are supported by randomization (10 mg/kg [arm 4] vs. corresponding placebo [arm 5] $p=0.12$ and titration [arm 8] vs. corresponding placebo [arm 9] $p=0.60$ are not significant). A very small study without a proper randomization supported analysis should never have more weight than a much larger phase 3 randomized (parallel) placebo controlled group trial (e.g., Study 301).

In summary, the totality of the data does not seem to support the efficacy of the high dose. There is only one positive study at best and a second study which directly conflicts with the positive study. Both studies were not fully completed as they were terminated early for futility and had sporadic unblinding for dose management of ARIA cases which was much higher in the drug group(s). The Amyloid PET substudy data suggested a larger effect in APOE- (non-carriers) which is the opposite of what was observed for the clinical outcome data. Within the high dose group at the patient level there is no correlation between the Week 78 change in the primary biomarker A β in the Cerebellum and the Week 78 Change from baseline in CDRSB. In Study 302, the on-face positive study, the raw correlation had the wrong +/- sign to support a realistic link between biomarker and long term clinical change in cognition/function as measured by

CDRSB. For these reasons, the reviewer believes there is no compelling substantial evidence of treatment effect or disease slowing and that another study is needed to confirm or deny the positive study and the negative study.

DRAFT

2 INTRODUCTION

2.1 Overview

The associated IND for the drug development was 106230. BIIB037 is a recombinant fully human antibody expressed in a CHO cell line, purified to a high degree of purity and formulated as a frozen liquid. BIIB037 is an IgG1 consisting of two heavy and two light chains connected by inter-chain disulfide bonds. BIIB037 has 1 carbohydrate moiety linked to Asn-304 in each heavy chain. The key studies intended to support efficacy are summarized in Table 1.

Table 1 Efficacy Study Characteristics

Study Name	Phase and Design	Treatment Period	# of Subjects per Arm	Study Population
301	3 placebo controlled parallel study	78 Weeks	N=1647 total placebo/low/high APOE depended high dose	MMSE 24-30 CDR global of 0.5 RBANS \leq 85
302	3 placebo controlled parallel study	78 Weeks	N=1638 total Placebo/low/high APOE depended high dose	MMSE 24-30 CDR global of 0.5 RBANS \leq 85
103	1B Staggered Multiple Dose Design	54 Weeks	N=180 total 9 arms Arms 1:3 1mg/kg 3mg/kg placebo Arms 4-5 :10 mg/kg /placebo 3:1 randomization	Prodromal MMSE 24-30 Mild AD CDR global of 0.5 or 1.0 and MMSE 20-26

			Arms 6-7: 6 mg/kg:placebo 3:1 Arms 8-9: (Apoet+ only) titration to 10 mg/kg: placebo	
--	--	--	---	--

2.2 Data Sources

The primary efficacy ADAM and SDTM data for Study 302 were located in the following directory at the time of review.

```
\\cdsesub1\evsprod\bla761178\0003\m5\datasets\221ad302\analysis\adam\datasets\adqs.xpt
```

```
\\cdsesub1\evsprod\bla761178\0003\m5\datasets\221ad302\tabulations\sdm\qs.xpt
```

The primary efficacy ADAM and SDTM data for Study 301 were located in the following directory at the time of review.

```
\\cdsesub1\evsprod\bla761178\0003\m5\datasets\221ad301\analysis\adam\datasets\adqs.xpt
```

```
\\cdsesub1\evsprod\bla761178\0003\m5\datasets\221ad301\tabulations\sdm\qs.xpt
```

The primary efficacy data for Study 103 were located in the following directory at the time of review.

```
\\cdsesub1\evsprod\bla761178\0003\m5\datasets\221ad103\analysis\adam\datasets\adqs.xpt
```

```
\\cdsesub1\evsprod\bla761178\0003\m5\datasets\221ad103\tabulations\sdm\qs.xpt
```

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

The submitted data and analysis quality appear adequate for review.

3.2 Evaluation of Efficacy

3.2.1 Study 302

3.2.1.1 Study Design and Endpoints

Study 221AD301 (ENGAGE) [and similarly designed study 221AD302] was a multicenter, randomized, double-blind, placebo-controlled, parallel-group study in subjects with early AD, including mild cognitive impairment (MCI) due to AD and a subset of mild AD, followed by an optional dose-blinded long-term extension (LTE) period of up to 5 years. Approximately 1605 subjects were to be enrolled across approximately 150 centers globally. The primary study objective is to evaluate the efficacy of monthly doses of aducanumab on the CDR-SB relative to placebo. Secondary objectives include assessment of the effect of monthly doses of aducanumab on clinical progression as compared to placebo.

Subjects were to be randomized in a 1:1:1 ratio to 1 of the 3 treatment groups: aducanumab high dose, aducanumab low dose and placebo, with stratification based upon their apolipoprotein E4 (ApoE ϵ 4) carrier status (carrier/non-carrier) and site. During the placebo-controlled period, subjects were to receive infusions of aducanumab or placebo approximately every 4 weeks for approximately 18 months (a total of 20 doses). Dose levels may be different in the same treatment group based upon subjects' ApoE ϵ 4 carrier status, and specifically, ApoE ϵ 4 carriers were to receive placebo, aducanumab 3 mg/kg, or aducanumab 10 mg/kg whereas ApoE ϵ 4 non-carriers were to receive placebo, aducanumab 6 mg/kg, or aducanumab 10 mg/kg.

Aducanumab will be titrated for up to 6 doses prior to reaching the target. Note: As of Protocol Version 4 (mid-study implementation), 10 mg/kg is the target dose for all ApoE ϵ 4 carriers in the high-dose group. ApoE ϵ 4 carriers who were randomized to the high dose group when the target dose was 6 mg/kg (under protocol versions prior to Version 4) must have received 2 or more doses at 6 mg/kg prior to being titrated up to 10 mg/kg. At the end of the double-blind, placebo-controlled treatment period, subjects who met the extension entry criteria could enter a long-term safety and efficacy extension period, with all subjects receiving aducanumab approximately every 4 weeks (up to a total of 65 doses over 5 years).

Additionally, participants who developed ARIA (except those with asymptomatic, radiographically mild ARIA-H microhemorrhage) were to have a follow-up MRI performed every 4 weeks until the ARIA resolved (ARIA-E) or stabilized (ARIA-H), per the centrally read MRI. These participants were also to have MoCA assessments at these follow-up visits as well as

biomarker, PK, and peripheral blood mononuclear cells samples collected at the first unscheduled visit following an episode of ARIA. Note: Participants with asymptomatic, mild ARIA-H microhemorrhages were exempt from these follow-up visits as mild ARIA-H microhemorrhage was observed at a similar incidence in aducanumab and placebo-treated participants in Study 221AD103.

The total duration of study participation for each subject only participating in the placebo-controlled period of the study was to be up to approximately 102 weeks, including a series of Screening visits within 8 weeks before administration of the first dose, a 76-week placebo-controlled treatment period, and a safety follow-up (FU) period of approximately 18 weeks after the final dose.

The primary is the Change from baseline in CDRSB at Week 78.

Secondary endpoints have been rank prioritized, in the order shown

- o Change from baseline in MMSE score at Week 78.
- o Change from baseline in ADAS-Cog 13 score at Week 78.
- o Change from baseline in ADCS-ADL-MCI score at Week 78.

3.2.1.2 Statistical Methodologies

Analysis Plan

Considerations for multiple comparison adjustments

A sequential (closed) testing procedure was to be used to control the overall Type I error rate due to multiple comparisons for the primary endpoint. The order of treatment comparisons was as follows: aducanumab high-dose versus placebo and aducanumab low-dose versus placebo. All comparisons after the initial comparison with $p > 0.05$ were not to be considered statistically significant.

Secondary endpoints have been rank prioritized, in the order shown

- o Change from baseline in MMSE score at Week 78.
- o Change from baseline in ADAS-Cog 13 score at Week 78.
- o Change from baseline in ADCS-ADL-MCI score at Week 78.

Note that the key endpoints are also assessed at Week 26 and Week 50.

In order to control for a Type I error for the secondary endpoints, a sequential closed testing procedure was to be used and was to include both the order of the secondary endpoints and treatment comparisons. Specifically, for each of the secondary endpoints, a sequential (closed) testing procedure, as for the primary endpoint, will be used to control the overall Type I error rate due to multiple treatment comparisons. If statistical significance is not achieved for 1 or 2 treatment comparisons, all endpoint(s) of a lower rank will not be considered statistically significant for that 1 or 2 treatment comparisons, respectively.

Reviewer's Comment: The closed testing for **each** of the secondary endpoints suggests that if the low dose is not significant the following tests and p-values for the high dose for lower endpoints in the hierarchy would not be allowable without inflating type I error.

Primary analysis

The estimand of the primary analysis is the mean difference of the change from baseline CDR-SB scores at Week 78 between treatment groups in the ITT population [ICH E9 (R1) Addendum 2014, 2017]. All observed data was to be included in the primary analysis, including data collected after intercurrent events [ICH E9 (R1) Addendum 2017], i.e., treatment discontinuation or a change in concomitant use of AD symptomatic medication. The change from baseline CDR-SB scores was to be summarized by treatment group at each post-baseline visit. A mixed model repeated measures (MMRM) model was to be used as the primary analysis to analyze change from baseline CDR-SB using fixed effects of treatment group, time (categorical), treatment group-by-time interaction, baseline CDR-SB, baseline CDR-SB by time interaction, baseline MMSE, AD symptomatic medication use at baseline(yes/no), region, and laboratory ApoE ε4 status (carrier/non-carrier). An unstructured covariance matrix was to be used to model the within-patient variance-covariance errors. If the unstructured covariance structure matrix results in a lack of convergence, the heterogeneous Toeplitz covariance structure followed by the heterogeneous first-order autoregressive covariance structure was to be used. The Kenward-Roger approximation was to be used to estimate the denominator degrees of freedom. In the primary analysis, missing data are assumed to be missing at random [Rubin 1976].

Sample Size Justification

A sample size of 450 subjects per treatment group (1350 in total) was planned to have approximately 90% power to detect a true mean difference of 0.5 in change from baseline CDR-SB at Week 78 between the 2 treatment groups. This power calculation was based on a 2-sided t-test assuming equal variance with a final significance level of 0.05, a standard deviation (SD) of 1.92 and a drop-out rate of 30%. The SD estimate of 1.92 for Week 78 reflected a 39% increase over the SD from the protocol-specified interim analysis of 1-year data. The assumed true mean difference of 0.5 between the 2 treatment groups represents an approximately 25% reduction in the placebo mean change from baseline at Week 78 if the placebo mean change is estimated to be 2. As defined in the protocol, the sample size for this study (and for the identically designed

Study 221AD302) was reassessed in a blinded manner in November 2017 (approximately 3 months before enrollment completion and with about 10.6% of the data available on the primary endpoint from Studies 221AD301 and 221AD302 combined). At this timepoint, the SD of the primary endpoint was estimated based on the pooled blinded data from two studies using a modified version of Gould-Shih simple-adjustment one sample variance (Zucker et al. 1999):

$$s_{adj}^2 = s_{os}^2 - \frac{2N}{9(N-1)} \delta^2,$$

where N denotes the number of subjects included in the analysis for blinded sample size reestimation (subjects with both baseline and Week 78 CDR-SB available at the time of sample size re-estimation), δ is the assumed true treatment effect (same treatment effect assumed for both the high dose group and low dose group in this analysis), and s_{os}^2 is the unadjusted one sample variance of the primary endpoint estimate from the pooled blinded data. As a result of this analysis, the sample size was increased from 1350 to 1605 (450 to 535 per treatment group) to assure adequate power for detecting a mean treatment effect of 0.5.

INTERIM ANALYSIS

An interim analysis was planned to occur after approximately 50% of the subjects had the opportunity to complete the Week 78 visit for both 221AD301 and 221AD302. To maintain the integrity of the study in the event of the interim analysis, an independent group external to Biogen that was not to be involved in the conduct of the study after unblinding was to perform the interim analysis. The IDMC was to review the unblinded results of the interim analysis provided by the independent group and was to make a recommendation to Biogen based on pre-specified criteria.

An interim analysis for futility of the primary endpoint was to be performed to allow early termination of the studies if it was evident that the efficacy of aducanumab is unlikely to be achieved. The futility criteria were to be based on conditional power, which is the chance that the primary efficacy endpoint analysis will be statistically significant in favor of aducanumab at the planned final analysis, given the data at the interim analysis. The conditional power is calculated assuming that the future unobserved effect is equal to the maximum likelihood estimate of what is observed in the interim data:

$$CP(Z(1) \geq Z_\alpha | Z(t)) = 1 - \phi \left(\frac{Z_\alpha \sqrt{n_2} - Z(t) \sqrt{n_1}}{\sqrt{n_2 - n_1}} - \frac{Z(t) \sqrt{n_2 - n_1}}{\sqrt{n_1}} \right)$$

where t is the fraction of information and $Z(t)$ is the observed Z-statistic at the interim analysis, $Z(1)$ is the Z-statistic and α is the type I error at the final analysis, n_1 and n_2 are the number of subjects at the interim and at the final analysis, respectively.

The futility decision was to primarily be based on the conditional power for the primary efficacy endpoint. The study was not to be considered as futile unless both studies 221AD301 and 221AD302 had conditional power for the primary efficacy endpoint less than 20% in both the high-dose and low-dose treatment groups. Given the insufficient knowledge of aducanumab's potential effects on various functional/cognition endpoints or in certain subgroups at the present time, other data in addition to the pre-specified futility criteria was to be considered as well, and the IDMC may recommend the studies to be continued as planned based on the weight of the evidence.

An interim analysis for superiority may be performed, to allow the possibility to demonstrate the treatment effect early. If an interim analysis for superiority was performed, the O'Brien-Fleming stopping boundary was to be used. If an interim analysis for superiority was not performed, then no alpha adjustment would be used for the final analysis after all subjects have had the chance to complete the Week 78 visit.

Reviewer's Comment: The Statistical analysis plan and the Unblinding plan do not definitively state whether the interim analysis was to include data from ongoing subjects who had not had the opportunity to complete Week 78.

Responder analysis

To further assess whether subjects on aducanumab progress differently from those on placebo, responder analysis will be conducted. The responders will be determined by a threshold of the primary endpoint, i.e., subjects whose change from baseline CDR-SB at Week 78 is smaller than or equal to the threshold will be classified as responders and otherwise will be classified as non-responders. All subjects with missing data at Week 78 will be classified as non-responders. The responder analysis will be conducted for two threshold values: 0.5 or 1.5, i.e., subjects whose change from baseline CDR-SB at Week 78 ≤ 0.5 or ≤ 1.5 . The number of responders and the response rate will be summarized by treatment group. The dichotomized response, responder vs. non-responder, will be modeled using a logistic regression with the following covariates: treatment group, baseline CDR-SB, baseline MMSE, AD symptomatic medication use at baseline (yes/no), region, and laboratory ApoE $\epsilon 4$ status (carrier/non-carrier). In addition to the two selected threshold values, the continuous responder curve that displays the percentage of responders under a wide range of threshold values will be presented by treatment group.

Amyloid PET Analysis

Amyloid PET substudy

Every subject enrolled into the study must have a positive amyloid PET scan by visual read either at screening or obtained within 12 months of screening. Subjects enrolled into the amyloid PET substudy will have the quantitative standard uptake value ratio (SUVR) scores at screening and at each planned post-baseline visit. The amyloid PET substudy will include a subset of approximately 400 subjects in countries other than Japan where PET scans will be performed using ^{18}F -florbetapir

ligand, and a small subset of subjects in Japan where either ^{18}F -florbetapir ligand or ^{18}F -flutemetamol ligand will be used. In the placebo-controlled period, amyloid PET assessments are scheduled at screening, Week 26, and Week 78.

Amyloid PET SUVR regions-of-interest and reference regions

Amyloid PET standardized uptake value ratio (SUVR) is a quantitative measure of cerebral amyloid plaque burden. The SUVR will be calculated for the following target brain regions of-interest (ROIs): composite ROI, frontal cortex, parietal cortex, lateral temporal cortex, sensorimotor cortex, anterior cingulate cortex, posterior cingulate cortex, medial temporal cortex, occipital cortex, striatum, and statistical ROI normalized to reference region activity.

Additionally, SUVR ROIs including pons and deep subcortical white matter which are believed to be least affected by amyloid pathology will also be evaluated. The composite ROI will comprise of major cortical regions part of the frontal, parietal, lateral temporal, sensorimotor, anterior, posterior cingulate and occipital cortices to serve as a summary measure of global cerebral amyloid burden. The statistical ROI is a region of interest consisting of the posterior cingulate cortex, precuneus and medial frontal cortex that has been demonstrated to yield optimal group separation between subjects with low and high amyloid burden across different reference regions. A negative change from baseline in composite ROI SUVR indicates a reduction in amyloid burden and a negative treatment difference (aducanumab minus placebo) favors aducanumab. The composite ROI will serve as the ROI of primary focus.

The following reference regions will be employed: cerebellum, cerebellum cropped, cerebellar white matter, cerebellar grey matter, deep subcortical white matter, pons, cerebellum + pons, cerebellar white matter + pons, deep subcortical white matter + cerebellum, deep subcortical white matter + pons and deep subcortical white matter + cerebellum + pons. Cerebellum will serve as the reference region of primary focus. The composite ROI SUVR using cerebellum as the reference region will be used as the primary endpoint for amyloid PET analysis.

Amyloid PET analysis population

There are two amyloid PET analysis population: ^{18}F -florbetapir amyloid PET analysis population and ^{18}F -flutemetamol amyloid PET analysis population.

The following background characteristics tables will be generated for the ^{18}F -florbetapir amyloid PET analysis population and will be presented by treatment group: number of subjects enrolled by region and country, demography, baseline disease characteristics, medical history.

By visit summary and MMRM model

The baseline and change from baseline amyloid PET SUVR values will be summarized by treatment groups (placebo, low dose and high dose) and by visit for each of the target ROIs using cerebellum as the reference region for each of the amyloid PET analysis populations. In addition, the baseline and change from baseline amyloid composite ROI values will be summarized by treatment groups by visit for each of the reference regions for each of the amyloid PET analysis populations.

For the ^{18}F -florbetapir amyloid PET analysis population, an MMRM model will be used to analyze change from baseline SUVR for each target ROI with cerebellum as the reference region. Fixed effects of the model will include treatment groups (placebo, low dose and high dose), visit (Week 26 and Week 78), treatment group-by-visit interaction, baseline SUVR (continuous), baseline SUVR by visit interaction, baseline MMSE (continuous), laboratory ApoE $\epsilon 4$ status (carrier and non-carrier), and baseline age (continuous). Visit and treatment group will be treated as categorical variables in the model along with their interactions. An unstructured covariance matrix will be used to model the within-patient variance-covariance errors. If the unstructured covariance structure matrix results in a lack of convergence in any of the parameters, the heterogeneous Toeplitz covariance structure followed by the heterogeneous first-order autoregressive covariance structure will be used for all the

parameters. The Kenward-Roger approximation will be used to estimate the denominator degrees of freedom. Adjusted means for each treatment group, pairwise adjusted differences with placebo, 95% confidence intervals for the differences and associated p-values will be presented at week 26 and week 78. The same MMRM model will also be used to analyze the change from baseline SUVR for the composite ROI with each of the reference regions. No multiple comparison adjustment will be used for amyloid PET analysis.

3.2.1.3 Patient Disposition, Demographic and Baseline Characteristics

Subject Accountability

Subject flow through the study is shown in Table 2 Patient Disposition. In both studies, incidence of discontinuation due to AEs was highest in the aducanumab high dose group.

Reviewer's Comment: The High dose had a larger proportion of patients discontinuing treatment, mostly due to AEs. This might affect the efficacy assessment because the missing data may be missing not at random, i.e., worse than completers.

Table 2 Patient Disposition in the Phase 3 Trials

Randomized	Study 301 (n=1653)			Study 302 (n=1643)		
	n=1647			n=1638		
Dosed	Placebo 545	Low Dose 547	High Dose 555	Placebo 548	Low Dose 543	High Dose 547
Discontinued treatment*, n (%)	96 (17.6)	105 (19.2)	148 (26.7)	82 (15.0)	108 (19.9)	131 (23.9)
Adverse event	26 (4.8)	43 (7.9)	64 (11.5)	16 (2.9)	41 (7.6)	46 (8.4)
Lost to follow-up	1 (0.2)	1 (0.2)	0	2 (0.4)	0	3 (0.5)
Disease progression	0	0	1 (0.2)	1 (0.2)	0	2 (0.4)
Consent withdrawn	14 (2.6)	11 (2.0)	15 (2.7)	6 (1.1)	22 (4.1)	18 (3.3)
Investigator decision	2 (0.4)	2 (0.4)	1 (0.2)	2 (0.4)	3 (0.6)	6 (1.1)
Death	0	3 (0.5)	1 (0.2)	5 (0.9)	0	5 (0.9)
Withdrew from study*, n (%)	58 (10.6)	60 (11.0)	78 (14.1)	39 (7.1)	54 (9.9)	66 (12.1)
Adverse event	16 (2.9)	23 (4.2)	26 (4.7)	10 (1.8)	11 (2.0)	18 (3.3)
Lost to follow-up	1 (0.2)	3 (0.5)	2 (0.4)	2 (0.4)	0	3 (0.5)
Disease progression	1 (0.2)	0	1 (0.2)	2 (0.4)	0	1 (0.2)
Consent withdrawn	21 (3.9)	14 (2.6)	23 (4.1)	8 (1.5)	28 (5.2)	22 (4.0)
Investigator decision	3 (0.6)	2 (0.4)	0	2 (0.4)	2 (0.4)	5 (0.9)
Death	0	3 (0.5)	2 (0.4)	5 (0.9)	0	6 (1.1)
Completed study (PC), n (%)	319 (58.5)	314 (57.4)	275 (49.5)	275 (50.2)	274 (50.5)	285 (52.1)
Active at the time of futility announcement, n (%)	168 (30.8)	173 (31.6)	202 (36.4)	234 (42.7)	215 (39.6)	196 (35.8)

* due to space limitation, only a subset of reasons are displayed.

Data source: t-acct-sub-pc-new-301/Output 5, t-acct-sub-pc-new-302/Output 6

Note: This table was copied from page 18 of the sponsor's 6/14/19 briefing package

Baseline Disease Characteristics

Subject demographics (Table 3) and baseline disease characteristics (Table 4) were balanced across groups in both phase 3 Studies.

Table 3 Baseline Demographics for Studies 301 and 302

Dosed	Study 301			Study 302		
	Placebo (n=545)	Low Dose (n=547)	High Dose (n=555)	Placebo (n=548)	Low Dose (n=543)	High Dose (n=547)
Age in years, mean \pm SD	69.8 \pm 7.72	70.4 \pm 6.96	70.0 \pm 7.65	70.8 \pm 7.40	70.6 \pm 7.45	70.6 \pm 7.47
Gender, Female n (%)	287 (52.7)	284 (51.9)	292 (52.6)	290 (52.9)	269 (49.5)	284 (51.9)
Race						
Asian n (%)	55 (10.1)	55 (10.1)	65 (11.7)	47 (8.6)	38 (7.0)	41 (7.5)
White n (%)	413 (75.8)	412 (75.3)	413 (74.4)	415 (75.7)	418 (77.0)	405 (74.0)
Education years, mean \pm SD	14.7 \pm 3.66	14.6 \pm 3.77	14.6 \pm 3.72	14.5 \pm 3.82	14.5 \pm 3.63	14.6 \pm 3.74
AD medications used, n (%)	293 (53.8)	307 (56.1)	307 (55.3)	279 (50.9)	277 (51.0)	277 (50.6)
ApoE ϵ 4, n (%)						
Carriers	376 (69.0)	391 (71.5)	378 (68.1)	367 (67.0)	362 (66.7)	365 (66.7)
Homozygote E4	104 (19.1)	101 (18.5)	104 (18.7)	92 (16.8)	97 (17.9)	77 (14.1)
Heterozygote E4	272 (49.9)	290 (53.0)	274 (49.4)	275 (50.2)	265 (48.8)	288 (52.7)
Non-carriers	167 (30.6)	156 (28.5)	176 (31.7)	178 (32.5)	178 (32.8)	181 (33.1)
Clinical stage, n (%)						
MCI due to AD	443 (81.3)	440 (80.4)	442 (79.6)	446 (81.4)	452 (83.2)	438 (80.1)
Mild AD	102 (18.7)	107 (19.6)	113 (20.4)	102 (18.6)	91 (16.8)	109 (19.9)
PET SUVR, mean composite \pm SD (n) – PET substudy only	1.38 \pm 0.198 (203)	1.39 \pm 0.186 (198)	1.41 \pm 0.177 (181)	1.37 \pm 0.175 (157)	1.39 \pm 0.181 (157)	1.38 \pm 0.183 (171)

Note: Table Copied from page 19 of 6/14/19 briefing package

Table 4 Baseline Disease Characteristics for Studies 301 and 302

Dosed	Study 301			Study 302		
	Placebo (n=545)	Low Dose (n=547)	High Dose (n=555)	Placebo (n=548)	Low Dose (n=543)	High Dose (n=547)
RBANS delayed memory score, mean \pm SD	60.0 \pm 13.65	59.5 \pm 14.16	60.6 \pm 14.09	60.5 \pm 14.23	60.0 \pm 14.02	60.7 \pm 14.15
MMSE, mean \pm SD	26.4 \pm 1.73	26.4 \pm 1.78	26.4 \pm 1.77	26.4 \pm 1.78	26.3 \pm 1.72	26.3 \pm 1.68
CDR Global Score, n (%)						
0.5	544 (99.8)	546 (99.8)	554 (99.8)	544 (99.3)	543 (100)	546 (99.8)
1	1 (0.2)	1 (0.2)	0	3 (0.5)	0	1 (0.2)
CDR-SB, mean \pm SD	2.40 \pm 1.012	2.43 \pm 1.014	2.40 \pm 1.009	2.47 \pm 0.999	2.46 \pm 1.011	2.51 \pm 1.053
CDR cognitive subscore, mean \pm SD	1.73 \pm 0.623	1.75 \pm 0.615	1.72 \pm 0.612	1.75 \pm 0.644	1.76 \pm 0.643	1.78 \pm 0.650
CDR functional subscore, mean \pm SD	0.68 \pm 0.574	0.68 \pm 0.558	0.68 \pm 0.585	0.72 \pm 0.554	0.71 \pm 0.558	0.73 \pm 0.577
ADAS-Cog 13, mean \pm SD	22.5 \pm 6.56	22.5 \pm 6.30	22.4 \pm 6.54	21.9 \pm 6.73	22.5 \pm 6.76	22.2 \pm 7.08
ADCS-ADL-MCI score, mean \pm SD	43.0 \pm 5.55	42.9 \pm 5.73	42.9 \pm 5.70	42.6 \pm 5.73	42.8 \pm 5.48	42.5 \pm 5.82

Data source: t-bl-char-pc-new-301/Output 7, t-bl-char-pc-new-302/Output 8

Note: Table Copied from page 20 of 6/14/19 briefing package

3.2.1.4 Results and Conclusions

3.2.1.4.1 Sponsor's Results

Table 5 shows the sponsor's results for both Study 301, study 302 and a pooled study analysis by the sponsor. The sponsor's final analysis of the first key secondary endpoint, MMSE, had a p-value of 0.0493 (down from .0620 as presented to the Agency in June 2019) after the addition of a few more records that were collected after unblinding.

Table 5 Phase 3 Primary and Key Secondary Results

	Study 301			Study 302			Studies 301+302		
	Diff vs PBO ^a			Diff vs PBO ^a			Diff vs PBO ^a		
	(%)			(%)			(%)		
	p-value			p-value			p-value		
	PBO decline (n=545)	Low dose (N=547)	High dose (N=555)	PBO decline (n=548)	Low dose (N=543)	High dose (N=547)	PBO decline (n=1093)	Low dose (N=1090)	High dose (N=1102)
CDR-SB	1.55	-0.18 (-12%)	0.03 (2%)	1.74	-0.25 (-14%)	-0.40 (-23%)	1.64	-0.21 (-13%)	-0.18 (-11%)
		0.2362	0.8252		0.1171	0.0101		0.0513	0.0974
MMSE	-3.5	0.2 (-6%)	-0.1 (3%)	-3.3	-0.1 (3%)	0.5 (-15%)	-3.4	0 (0%)	0.2 (-6%)
		0.4875	0.7961		0.6900	0.0620		0.8346	0.2621
ADAS-Cog13	5.171	-0.590 (-11%)	-0.605 (-12%)	5.171	-0.747 (-14%)	-1.395 (-27%)	5.171	-0.657 (-13%)	-0.989 (-19%)
		0.2475	0.2446		0.1672	0.0098		0.0764	0.0083
ADCS-ADL-MCI	-3.8	0.7 (-18%)	0.7 (-18%)	-4.3	0.7 (-16%)	1.7 (-40%)	-4.0	0.7 (-18%)	1.2 (-30%)
		0.1345	0.1520		0.1556	0.0009		0.0347	0.0005

^a: difference vs placebo at week 78. Negative percentage means less progression in the treated arm.

N: numbers of all randomized and dosed subjects that were included in the ITT analysis.

Note: This table was copied from page 23 of the 6/14/19 briefing package

Reviewer's Comment: Week 78 High vs Placebo Confidence Intervals: 302: -.401 +/- .156 (-.706 -.095) ; 301: +.03 +/- .150 (-.262,.328). The Placebo LS Mean at Week 78 was 1.56 and 1.74 in 301 and 302, respectively, and Aducanumab hi was 1.59 and 1.35.

A large proportion of the ITT population (~45%) did not have the opportunity to complete Week 78 due to the futility stopping of the trials. The result for the Opportunity to Complete Dataset (N=953) in Study 302 for the high dose on the primary endpoint, CDRSB at Week 78, was -.36, 95% CI=[-.70,-.02], p=0.0368.

Reviewer's Comment: The sponsor's highlighting of percent reduction from placebo masks the placebo effect which is highly variable and doesn't represent the analysis scale or acknowledge the standard error of the percent reduction which is needed for proper context. The p-values reflect the primary analysis model which evaluated the simple difference $\mu_d - \mu_p$, not the percent difference. Percent reduction should really be estimated using a different model, e.g., $\log(\text{CDRSB})$, and p-values and standard errors would be different.

3.2.1.4.2 Reviewer's Results

The sponsor published the placebo controlled period results of study 103 in September 2016 and presented the 48 month analysis of study 103 results at an Alzheimer's meeting in October 2018. The date of first treatment in study 302 was 15 September 2015.

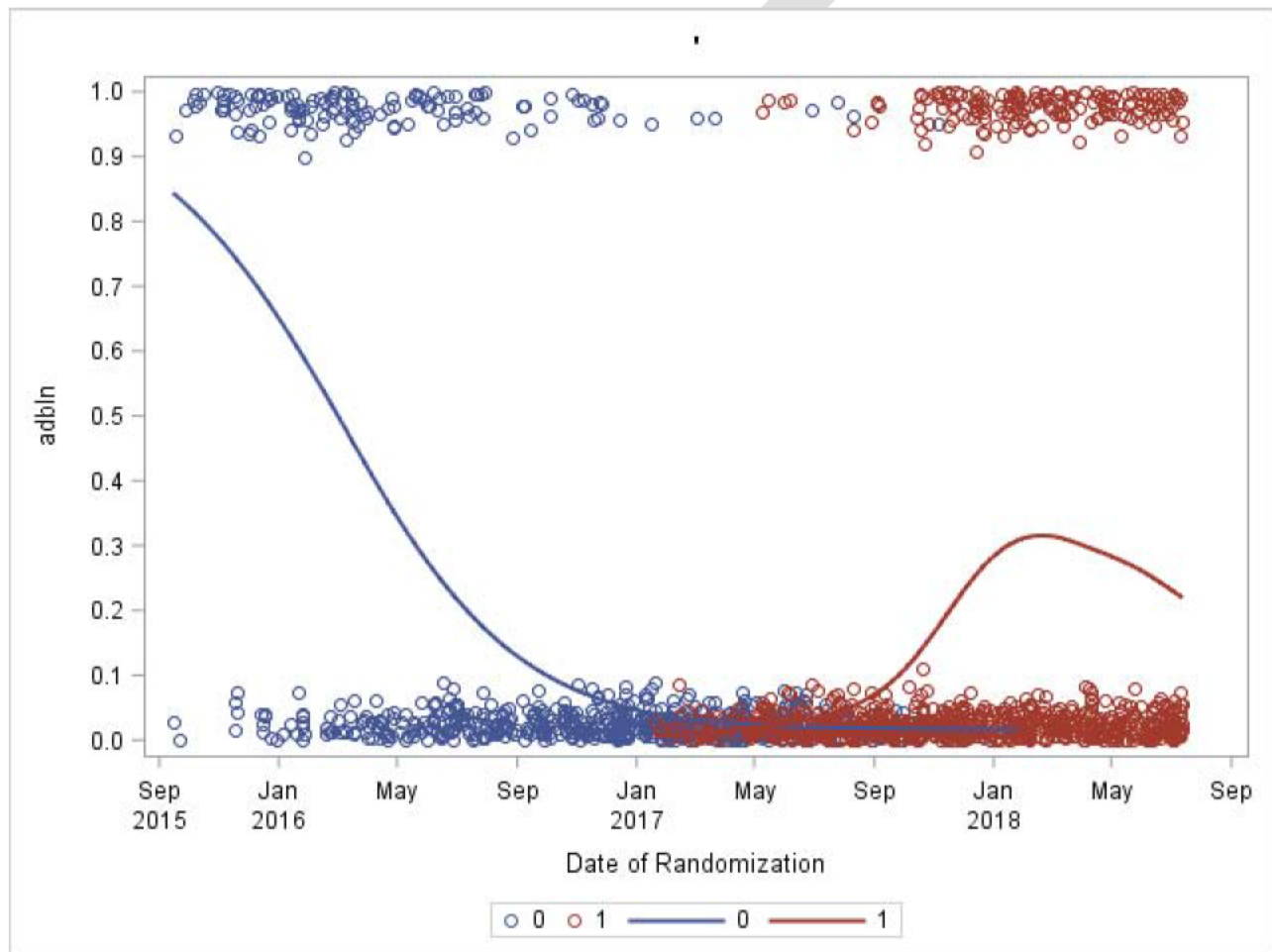
It seems to be an uncommon situation to continue collecting and/or cleaning additional efficacy data after most of data has been unblinded and analyzed (BLA final: nsub=1581 nobs=3716; June 2019 final : nsub=1580 nobs=3712 [876 Wk 78 out of 1637 patients: 258 placebo and 276 high with Opportunity to Complete -OTC]).. [Study 301 BLA final: nsub=1602 nobs=3901 ; June 2019 final: nsub= 1602 nobs=3897]. With the addition of this data the first key secondary, MMSE, went from non-significance to nominal significance in study 302 ($p=.0620$ to $.0493$).

There is a lot of missing data in study 302 (and 301) at Week 78 (>40%) caused by early stopping due to futility. Week 78 is the only Visit with apparent efficacy for CDRSB in 302. Unblinding due to ARIA and high dose titration limitations seem to have increased after protocol amendment 4 and these vary significantly by APOE+ vs. APOE- (as do to some extent CDRSB changes). Therefore, missing at random may not be a reasonable assumption without stratifying by APOE.

Those missing Week 78 due to late enrollment and early stopping have some differences in baseline characteristics compared to those who had the opportunity to complete: Baseline use of symptomatic medications overall was increased by 6.7% ,Region 1 decreased from 46 to 29% (region 2 increased from 50 to 57% and region 3 increased from 3 to 13%), the percentage Mild diagnosis increased from 15 to 25% , baseline CDRSB for the high dose is about 0.20 points higher and .05 lower for placebo , and the oldest age group proportion increased by 5% after Sep 21, 2017 as compared to those randomized before the same date (i.e., with the opportunity to complete by March 21, 2019). This indicates the missing data is not missing completely at random, so the analysis must rely on the missing at random assumption in this dataset, i.e., missingness could depend on covariates and/or earlier observed post-baseline CDRSBs. In study 301 also, the proportion of region 1 decreased steadily over the course of the study and was decreased by more than 15% and the proportion mild increased by 10% and there were 8% more placebo than high in the youngest group and 9% more high than placebo in the middle age group for those randomized after Sep 21, 2017. Similarly, in the PV4 subset (those who consented to protocol version 4 [dated March 2017] by Week 16 of their participation), 97% of those with the opportunity to complete had baseline stage Prodromal, whereas among those without the opportunity to complete 74% were Prodromal and 26% were Mild. This is illustrated in Figure 1 which shows Study 302 enrollment probability of Mild (rather than Prodromal) at baseline over the duration of enrollment. It started high then decreased to near zero and began to increase again slightly as PV4 patients were enrolled. This indicates that PV4 consented patients with the opportunity to complete were more Prodromal than pre-PV4 and that those PV4 without the

opportunity to complete were more Mild. The opportunity to complete population analysis seems the most appropriate since it relies least on the model being correctly specified and considering the various suggestions from the data that the model failed to include various important interaction effects (Country*VISIT, Baseline Disease Stage*Visit, Baseline AD meds*VISIT, etc.) which could cause the ITT analysis to be biased given the large amount of missing data.

Figure 1 Probability of Randomization of Mild (1) vs. Prodromal (0) Baseline Disease Stage as Study 302 progressed



Pv4 compared to pre-PV4 also had some of these baseline characteristic differences as shown in Table 6.

Table 6 Study 302 Demographic Characteristics by pre-PV4 and post-PV4

		PV4		All
		Pre-PV4	Post-PV4	
Baseline Alz Dis Med Use Flag		323	467	790
N	N			
	PctN	44.74	50.98	48.23
Y	N	399	449	848
	PctN	55.26	49.02	51.77
Baseline Alzheimer Disease Stage		129	173	302
MILD	N			
	PctN	17.87	18.89	18.44
PRODROMAL (MCI)	N	593	743	1336
	PctN	82.13	81.11	81.56
Geographic Region 1		12	109	121
Asia	N			
	PctN	1.66	11.90	7.39
Europe/Canada/Australia	N	383	482	865
	PctN	53.05	52.62	52.81
United States	N	327	325	652
	PctN	45.29	35.48	39.80
All	N	722	916	1638

In fact there was a significant three way interaction between country treatment and visit; comparing these additional model terms to the primary analysis model the three way interaction was highly significant $F=1.63$ (num df=106,den df=3024) $p<.0001$ and tested individually:

(COUNTRY	12	1473	4.45	<.0001.		
COUNTRY*TR01PG1N	24	1460	0.78	0.7644		
COUNTRY*AVISITN	24	1751	2.00	0.0028		
COUNTR*AVISIT*TR01PG	48	1840	1.55	0.0095).		

The primary analysis of 302 contained 1581 patients and 3716 observations. About 14% of randomized patients in study 302 started concomitant AD medications post-baseline(the highest use was in the high dose 15.2% [15.6% post -PV4]). The rate of starting increased slightly after the protocol version 4 implementation 11.5% to 16.0% after 12 to 17.8% for the high dose [study 301 overall (10.0% pre to 13.3% post)]. Analyses addressing post-baseline starting of concomitant AD medications involved the following number of patients and results.

Table 7 Study 302 Week 78 CDRSB Analyses Exploring Exclusion of Data after post-Baseline starting of AD medications

Handling	Subjects	Records	Hi Vs. Placebo LS Mean Difference	Std. Error of Difference	p-value
Censoring impacted Data	N=1524	3466	-0.410	0.158	0.0098
Censoring Impacted Patients	N=1358	3206	-0.431	0.161	0.0074

In study 302, the estimated high dose effect was smaller in those who started concomitant AD medications (0.1993+/-0.4209 vs. 0.4302 +/- 0.1651). In study 301 the subgroup that started concomitant AD medications was in the right direction but the subgroup that did not was in the wrong direction for the high dose compared to placebo (.269 +/- .452 vs. -.088 +/- .158).

The sponsor makes an argument about outliers being influential more in study 301 than 302. However, a robust regression, which is by design less affected by outliers (least trimmed squares /M estimation), of study 301 Week 78 data shows no effect of the high dose on CDRSB (Placebo-High=0.0265 S.E.=0.1247 p=0.8315). The low dose is also still numerically better than the high dose in the robust analysis of study 301 (-0.0628 SE=0.1249 p=0.6153). The Opportunity to Complete Subset as defined by the Sponsor consists of those randomized patients who were randomized early enough in the study timeline in order to have had the Opportunity to have a Week 78 assessment before March 20, 2019. Note that missing data is fairly limited in the Opportunity to Complete subset (91.7% complete in 301 and 93.7% complete in 302) so that analyzing just the Week 78 Visit in this subset seems not unreasonable, especially since there is no indication of an earlier effect (Week 26 or Week 50) in the primary MMRM analysis.

Figure 2 shows histograms of Week 78 changes from baseline in CDRSB for placebo and the high dose with the bars side by side at each level of change in the Opportunity to Complete Population.

Figure 2. CDRSB changes at Week 78 in those with opportunity to complete (studies 301 and 302 pooled)

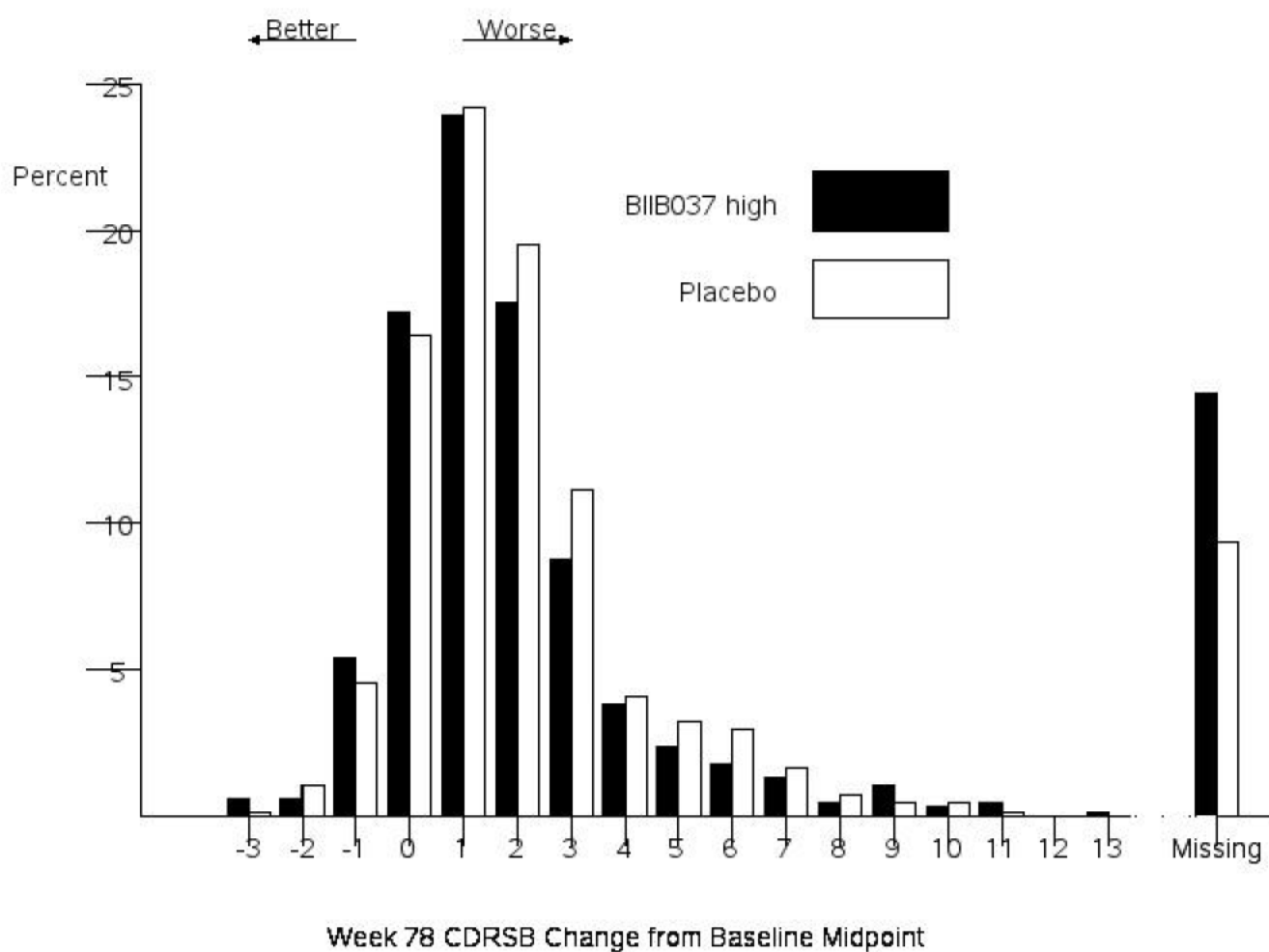
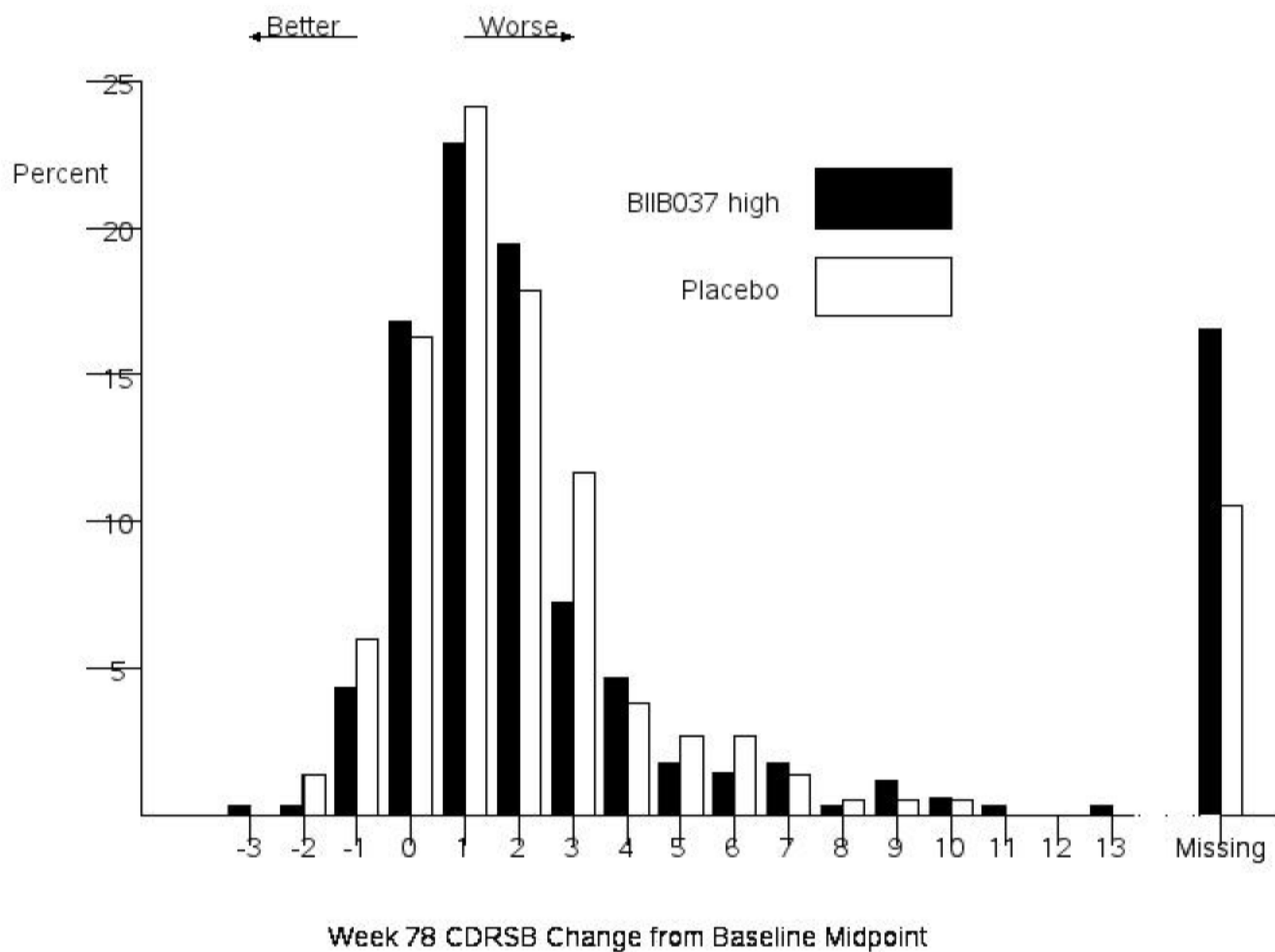


Figure 3 shows histograms for CDRSB Changes at Week 78 in Study 301 for the high dose and placebo groups.

Figure 3. Study 301 CDRSB Changes at Week 78 (excluding missing data but deaths coded as missing)

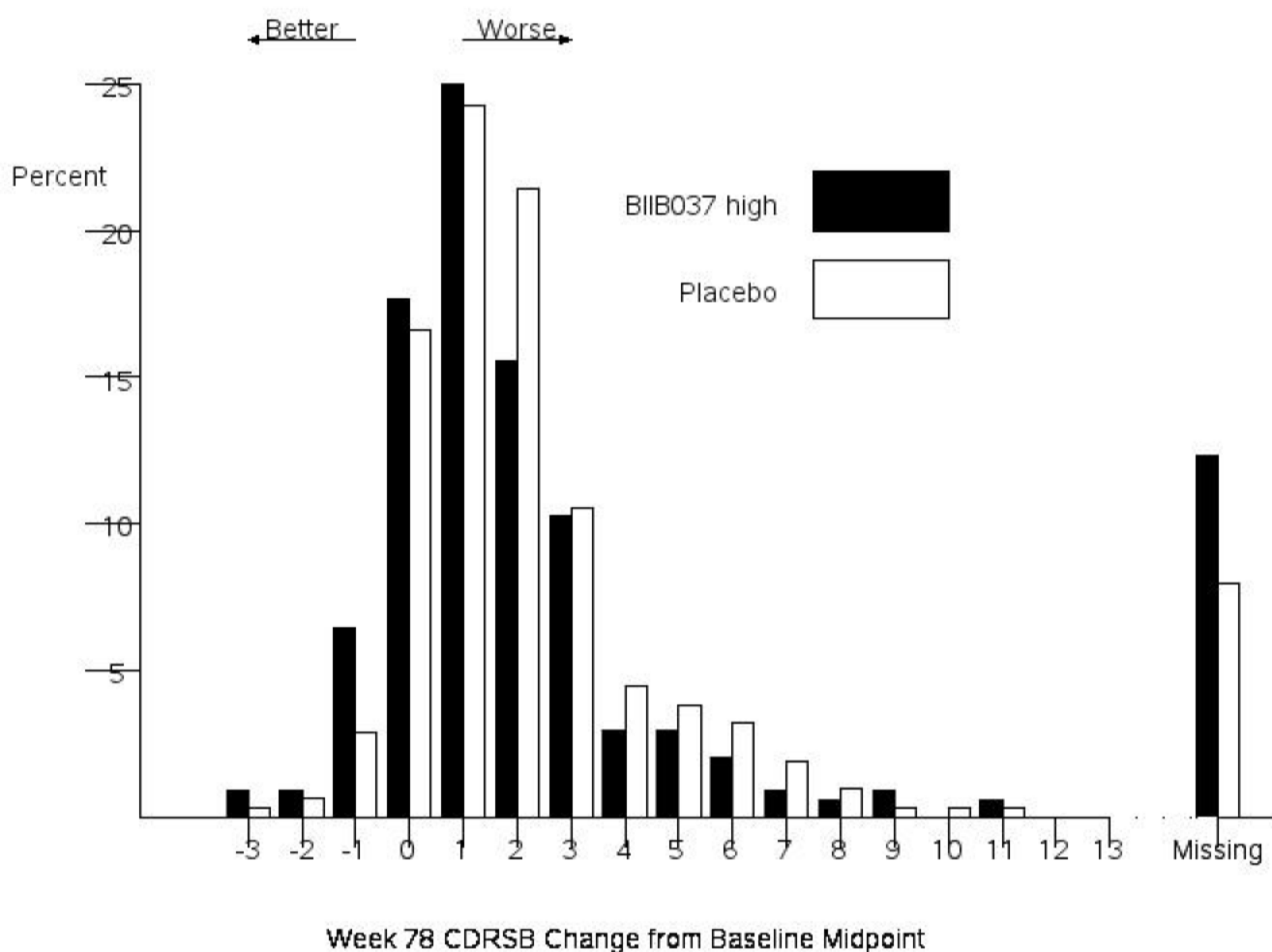


The sponsor prespecified two responder thresholds for secondary analyses of CDRSB, 1.5 and 0.5 thresholds. In study 302 there is a 3.24 % higher responder rate with 1.5 CDRSB threshold for response (high dose compared to placebo) and 4.41% with the 0.5 threshold in the

opportunity to complete week 78 group (N=803 with 83 missing). These differences are 5.60%, and 9.39% respectively in ITT excluding missing Week 78.

In **Figure 4** one can see that Study 301 only had one worse change observed for the high dose, i.e., a 13 point worsening, (and 301 had 1 11 point worsening while study 302 had 2 11 point worsenings).

Figure 4. Study 302 CDRSB changes at Week 78 (excluding missing data)



The sponsor also asserts that there is no bias due to ARIA because excluding data after ARIA doesn't markedly change the primary result. However, one can't conclusively rule out an impact

of those experiencing ARIA on the result because it requires making a comparison based on differential exclusions between the randomized groups (drug patients and/or censoring of drug arm data) and the resultant groups without ARIA to be compared are no longer as randomized and/or have differential follow-up and selection bias due to conditioning on a post-randomization event.

Figure 5 shows from left to right a comparison of pre-to post PV4 CDRSB profiles LSmeans by group: placebo (left), low dose(middle), and high dose (right). It shows a dramatic worsening of placebo post-PV4 (red) relative to pre-PV4 (blue). The high dose LS Mean at Week 78 is virtually unchanged between pre-PV4 and post-PV4 and equivalent to the low dose post-PV4. This suggests that it may not be high dose improvement due to the dose increase but rather placebo worsening that explains the apparent high dose effect in study 302. This theory is further supported by a lack of correlation between high dose patient CDRSB changes at Week 78 and the same patients' Week 78 cerebellum A β SUVR changes in the PET subset (where available). In the APOE non carriers the high dose was numerically worse than placebo in the post-PV4 subset and only slightly better pre-PV4 (and no better than the low dose pre-PV4).

One can see in Figure 5 that placebo was dramatically worse in the APOE+ stratum post-PV4 for CDRSB as compared to pre-PV4, while the other treatment groups were more consistent from pre-PV4 to post-PV4.

Figure 5. Study 302 Placebo, Low Dose, and High dose CDRSB profiles by Pre-PV4 and-Post PV4 status

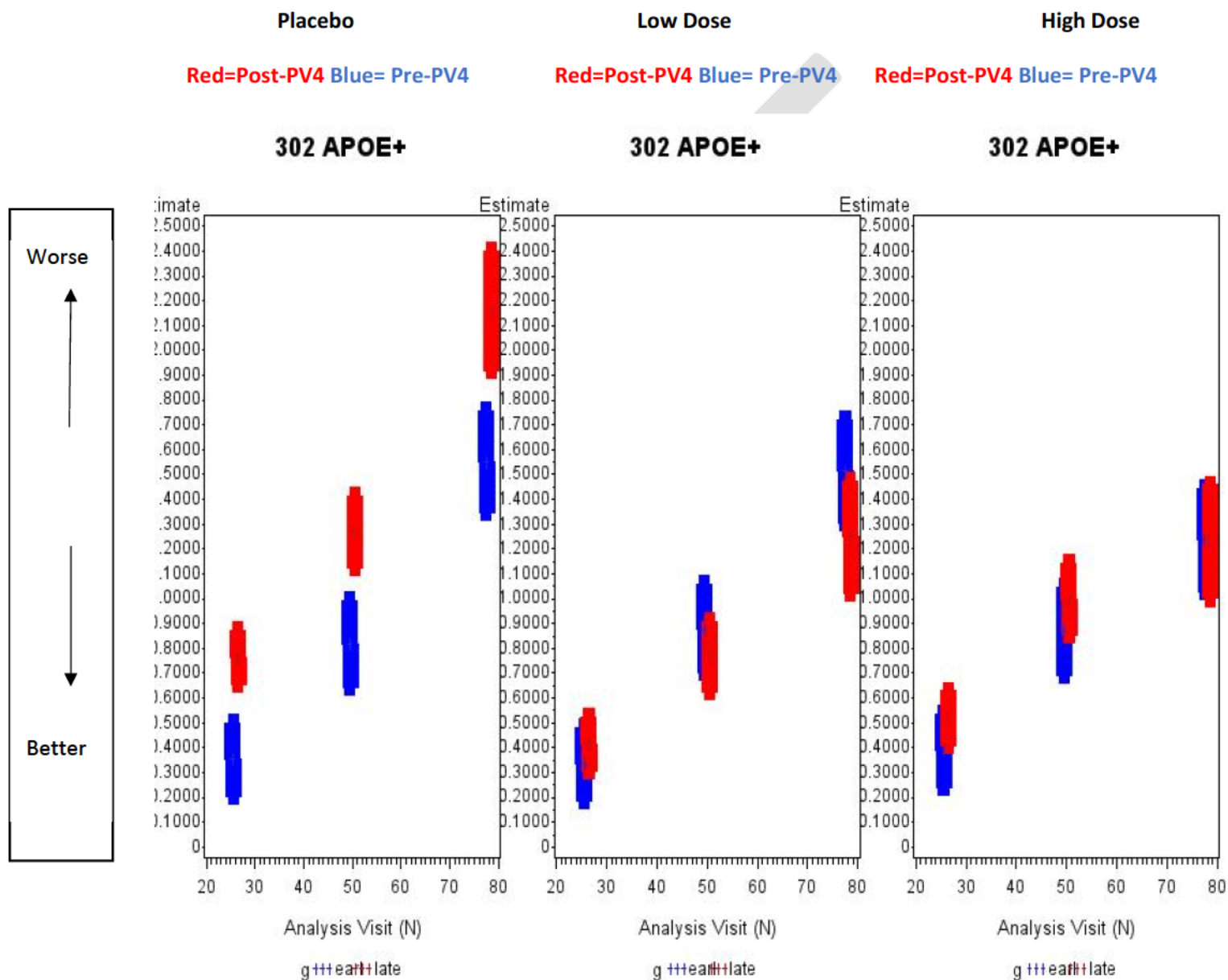


Figure 6 shows the corresponding CDRSB profiles for the APOE non-carrier stratum. The high dose is numerically worse than placebo at Week 78 in the protocol amendment 4 subgroup in the non-carrier stratum (compare red on the left at Week 78 with red on the right at Week 78).

Figure 6. Study 302 Placebo, Low Dose, and High Dose CDRSB Profiles by Pre-PV4 and Post-PV4

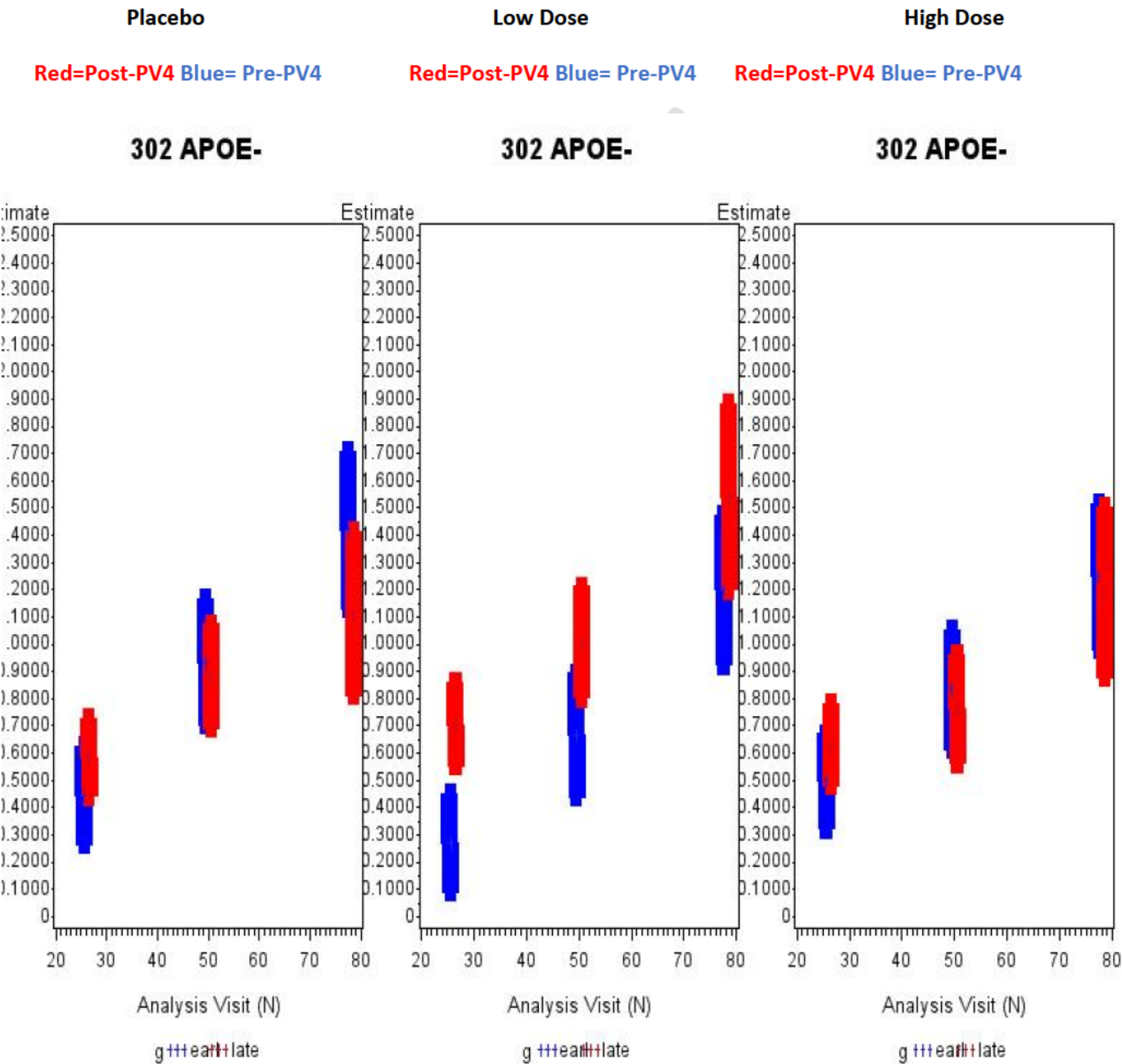


Figure 7 shows the local trend (adjusted for primary model covariates other than visit since this is all Week 78) in study time in LS Mean CDRSB Change at Week 78. Blue is non-PV4 placebo, Red is for non-PV4 High Dose, Green is for PV4 placebo and brown is for PV4 high dose. The number of 10 mg/kg doses are shown just above and to the right of the plot symbol. One can see that there were some relatively poor Week 78 changes among those PV4 high dose with the full possible 10 mg/kg dosing. The figure suggests that there was a worsening trend in PV4 placebo as well as to a lesser degree PV4 high dose.

Figure 7. Local Trend in CDRSB Week 78 Changes over Study Duration and by PV4

0=Pre-PV4 Placebo ; 1=Pre-PV4 High Dose; 2=Post-PV4 Placebo; 3= Post-PV4 High Dose

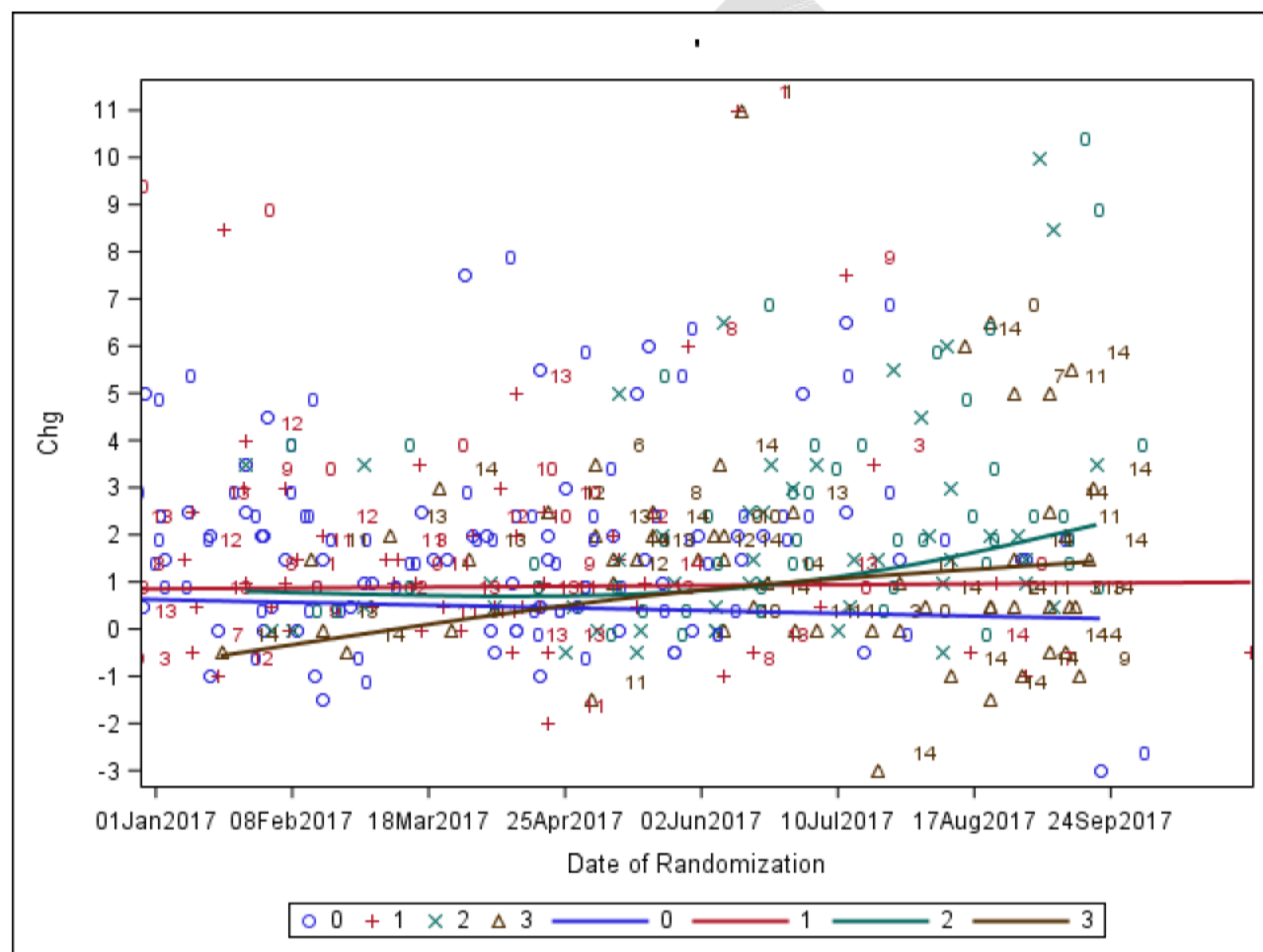


Figure 8 shows the same trends over the full course of the study.

Figure 8. Local Trend in CDRSB Week 78 Changes over Study Duration and by PV4

0=Pre-PV4 Placebo ; 1=Pre-PV4 High Dose; 2=Post-PV4 Placebo; 3= Post-PV4 High Dose

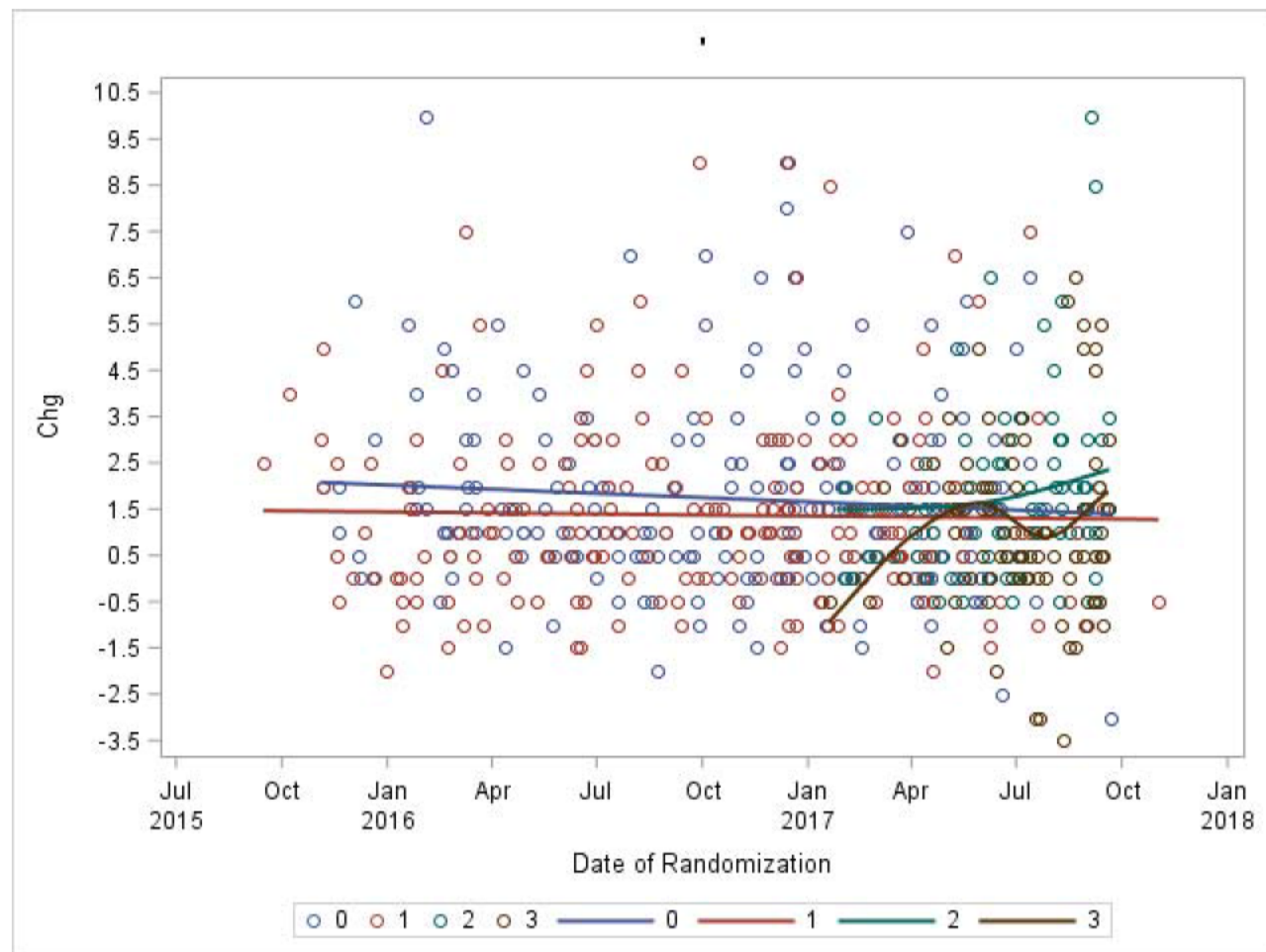


Figure 9 shows that the probability of a randomized patient being Asian increased post-PV4 (red) in Study 302. The blue symbols indicate pre-PV4 subjects and the red symbols indicate post-PV4 subjects. In the figure shown 0 y-values represent non-Asian patients and 1.0 values represent Asian (a small amount of noise was added to avoid obscuring many coincident points). The curves show the trend in the probability over time as estimated by a generalized additive binomial model (smooth curve) for the probability of a randomized patient being Asian.

Figure 9. Probability of Asian Enrollment over Time in Study 302

0= Non-Asian ; 1=Asian

Blue=Pre-PV4 Red=Post-PV4

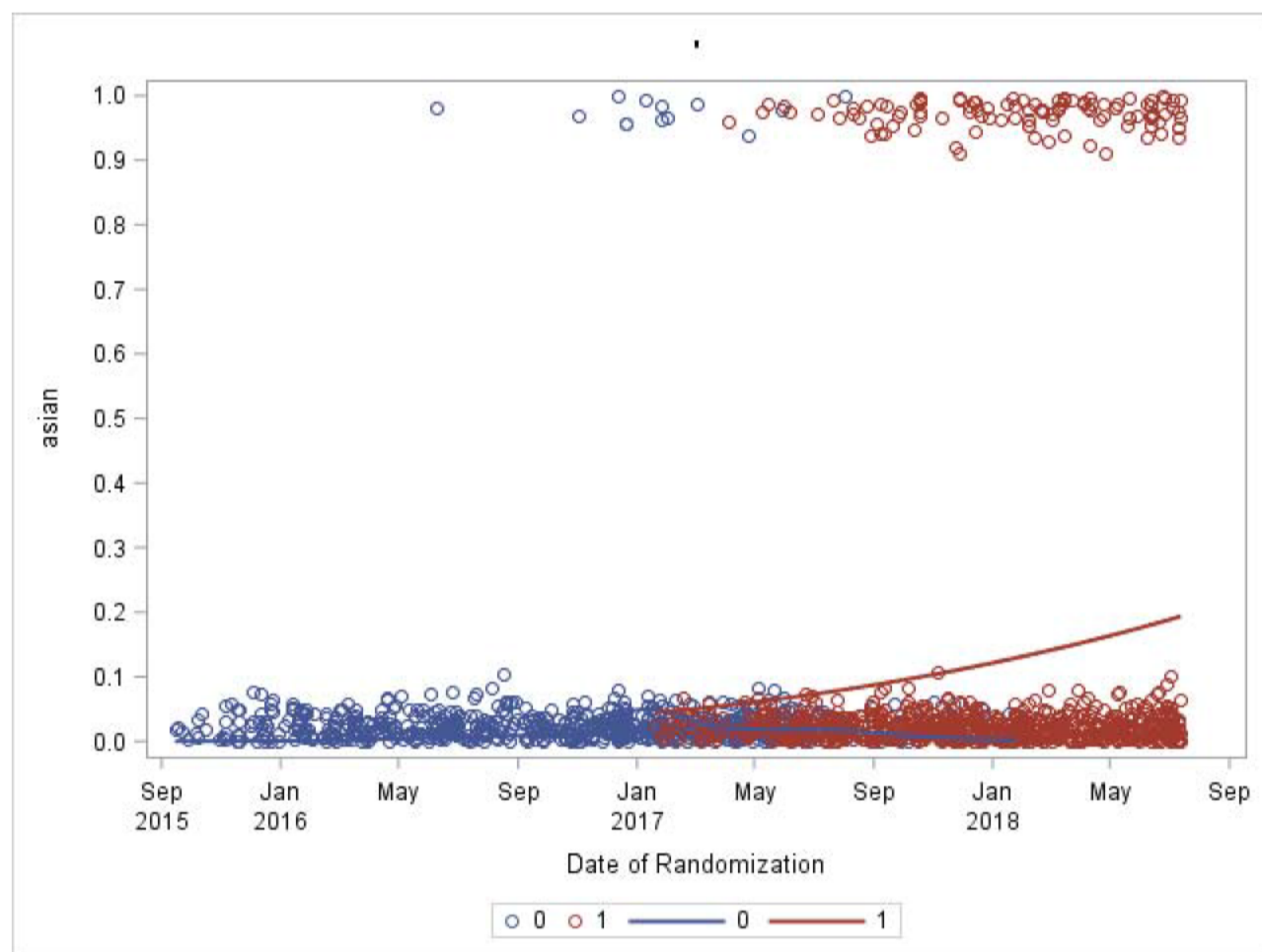
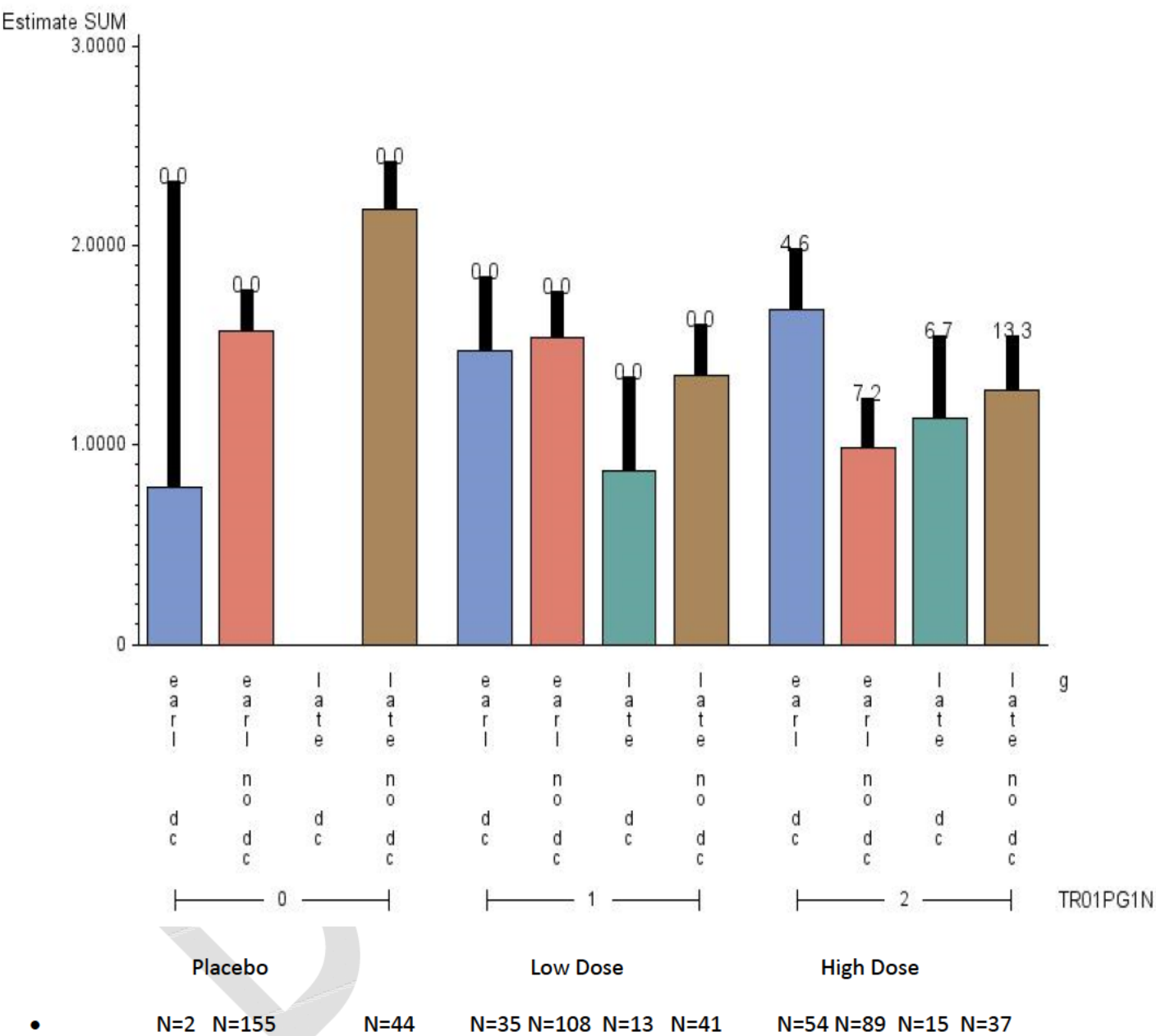


Figure 10 shows the surprising result that the high dose subgroup with dose titration change due to ARIA had a numerically better Week 78 LS mean CDRSB change than the unmodified subgroup that had more 10 mg/kg doses. The statistic over the bar indicates the average number of 10 mg/kg doses which is obviously zero for the low and placebo doses (ranging from 4.6 to 13.3 for high dose). DC stands for dose titration change or at least modification due to an adverse event of ARIA. Those classified as early by not meeting the sponsor's definition of PV4, which requires consent by Week 16 could still possibly have consented to PV4 later than Week 16 (depending on enrollment time). The same is true for the low dose and it is interesting to note that the late high dose subgroups are no better than the corresponding low dose groups. Overall, in PV4 APOE+ the low dose LS mean at week 78 is 1.25 and the high dose LS mean is 1.24.

Worsening of placebo from early to late alone can explain the high doses improved outcome relative to placebo, i.e., the high dose is not improved relative to the corresponding low dose and placebo worsened. In study 302, it is also numerically worse than the Early High Dose non-dose-reduced subgroup (red) that also had significantly fewer 10 mg/kg doses by Week 78 and only .07 better than the low dose. Overall in study 302 APOE+ post-PV4, regardless of dose titration changes or lack thereof, at Week 78 the high dose LS mean change from baseline in CDRSB was only .018 points better than the low dose calling into question the supposed importance of 10 mg/kg doses (and only .11 better in study 301 post-PV4). The sample sizes shown below the figure are admittedly small but the reviewer is surprised that the subgroup requiring dose titration adjustments due to ARIA (requiring some unblinding) and so getting fewer 10 mg/kg doses is numerically better in the post-PV4 subset. What did change dramatically in study 302 was the placebo response was considerably worse post-PV4 (pre: 1.55 +/- .20 vs. post: 2.16 +/- .24 post this would be a nominally significant change; compare leftmost brown and blue and red bars). There also appears to be very little difference between the low (middle brown) and high dose (right brown) LS means post-PV4 after the dose of the high dose had been allowed to increase from 6 to 10 mg/kg.

DRAFT

Figure 10. Study 302 Week 78 LS Mean CDRSB Change by ARIA Dose Modification Status and pre/post PV4 in APOE+

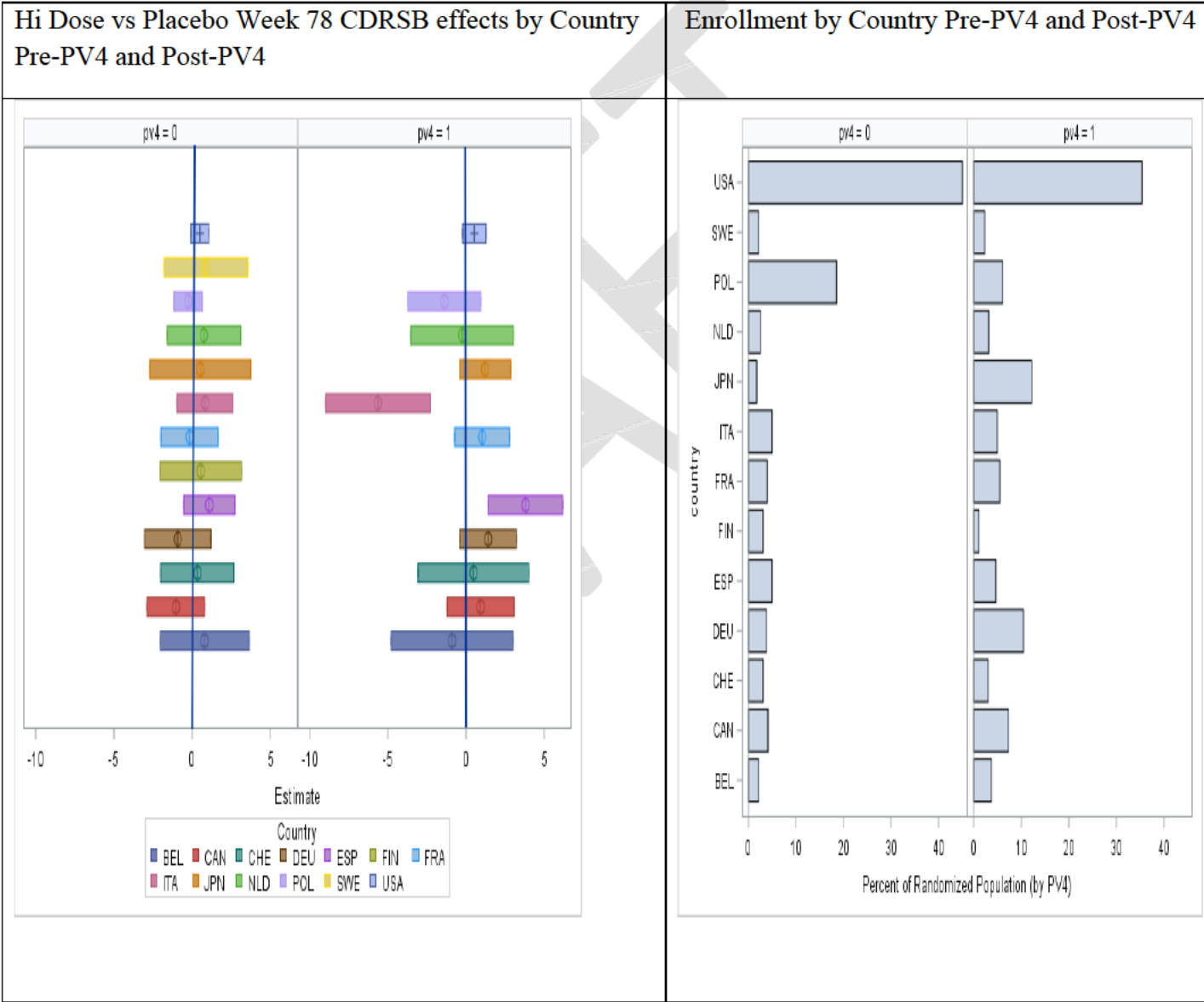


• Note: earl=Pre-PV4 late=Post-PV4; dc= dose titration slowing or reduction due to ARIA

Overall, not subsetting by PV4, the country by Visit by treatment group interaction was significant, $p=0.01$, suggesting that CDRSB profiles over time and treatment group differences varied significantly by Country.

Figure 11 shows the distribution of enrollment by Country pre-PV4 and post-PV4 as well as the differences in estimated high dose Week 78 effects by Country pre-PV4 and post-PV4.

Figure 11. Differences in estimated Week 78 CDRSB High Dose effects by Country in Study 302



There is almost surely less happening in APOE non-carriers in **study 302** although they got higher dosing from the beginning. Consider the forest plot below.

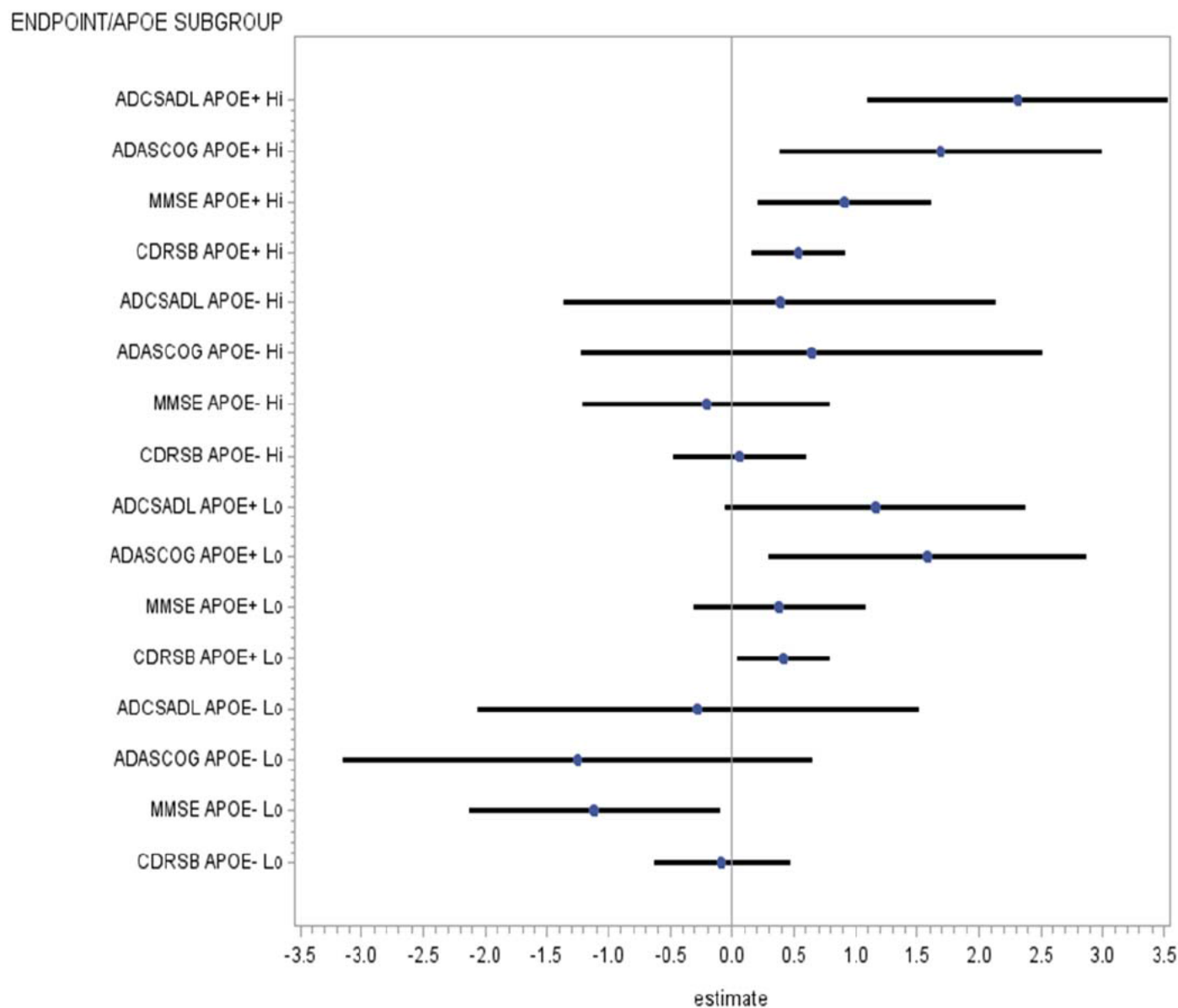
There are eight possible comparisons of Week 78 treatment effect estimates between APOE+ to APOE-, across the four key endpoints and two doses. The forest plot (**Figure 12**), cycles through the four endpoints from top to bottom first for APOE + high dose, then for APOE- high dose, then for APOE+ low dose, and finally the four endpoints for APOE- low dose. One can see that for each horizontal line the point estimate at the center for the opposite APOE subgroup for the same endpoint and the same dose (four bars below) is lower. Eight out of eight cases, APOE+ is better. Note that the signs of MMSE and ADCSADL-MCI were changed, so that for all four endpoints a positive difference would favor the drug. All four endpoints for APOE- non-carrier low dose are in the wrong direction numerically.

There is a statistically significant interaction for MMSE between APOE and treatment groups, $p=0.0065$, i.e., **the treatment effect(s) is not the same across APOE subgroups**.

Also, a multivariate test for interaction over the four key endpoints between APOE and treatment at Week 78 adjusted for the non-Visit based primary model covariates has $p=0.0495$.

DRAFT

Figure 12. Study 302 Key Endpoints APOE subgroup treatment effects estimates by Dose



For all of the four key endpoints the APOE- estimated effect is lower than APOE+ for both low and high doses (see also Table 8). This is despite the fact that APOE- high dose group had 10mg/kg as maximum dose both pre-PV4 and post-PV4 (higher average exposure than APOE+). On the first key secondary (MMSE) the treatment by APOE interaction test has a p-value of .0065.

A test for a difference in MMSE treatment effect at Week 78 in either low or high between APOE+ and APOE- has a p-value of .0454. If the doses are combined the p-value is .0157.

Table 8. Study 302 Estimated Treatment Differences at Week 78 by APOE for Primary and Key Secondary Endpoints

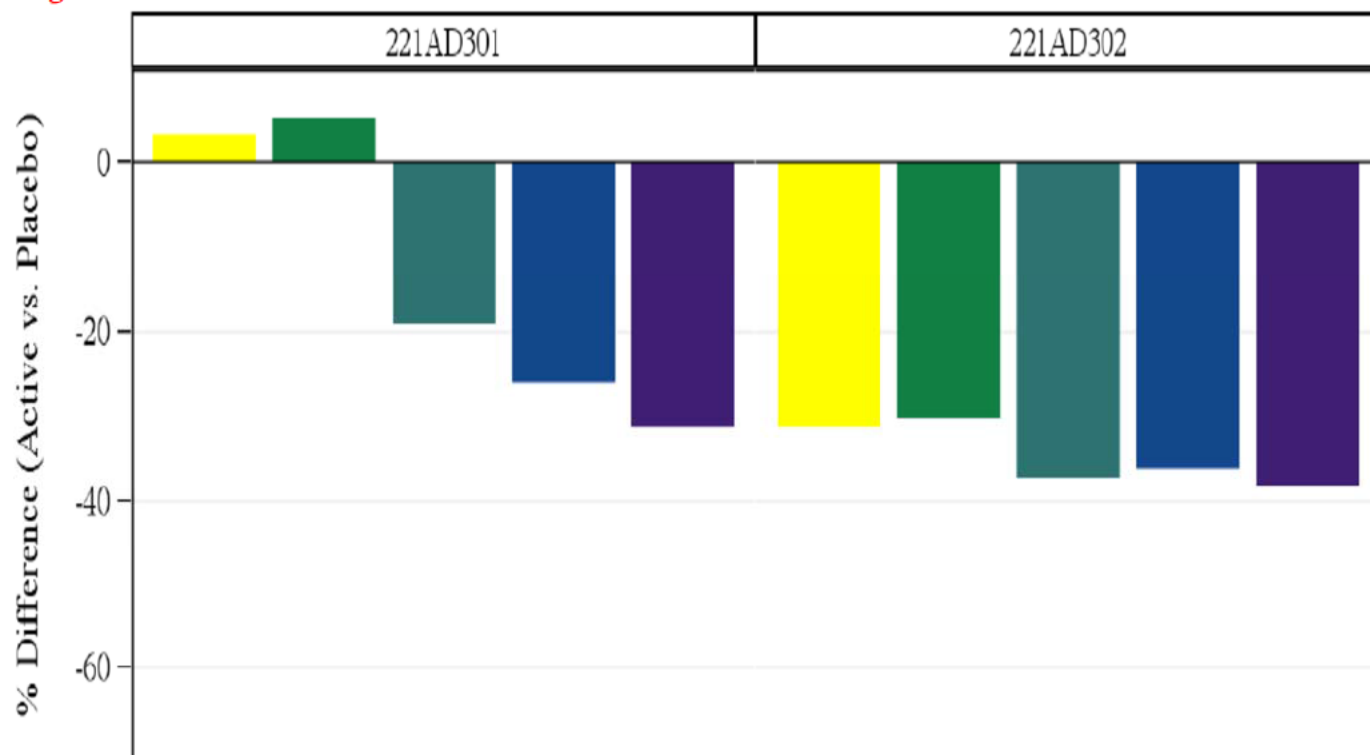
Endpoint	Dose	APOE+ Diff vs Pl LSMean (Std.Error)	p-value	APOE- Diff vs Pl LSMean (Std.Error)	p-value	APOE- - APOE+ diff LSMean (Std.Error)	p-value
CDRSB	Hi	0.53 (0.19)	0.0053	0.06 (0.27)	0.8366	-0.47 (0.33)	0.1537
CDRSB	Lo	0.42 (0.19)	0.0268	-0.09 (0.28)	0.7523	-0.51 (0.34)	0.1324
MMSE	Hi	0.91 (0.35)	0.0103	-0.21 (0.51)	0.6789	-1.12 (0.62)	0.0710
MMSE	Lo	0.38 (0.35)	0.2747	-1.12 (0.52)	0.0314	-1.50 (0.63)	0.0167
ADASCOG	Hi	1.68 (0.66)	0.0108	0.65 (0.95)	0.4966	-1.04 (1.16)	0.3697
ADASCOG	Lo	1.58 (0.66)	0.0164	-1.25 (0.97)	0.1968	-2.83 (1.17)	0.0158
ADCSADL	Hi	2.31 (0.62)	0.0002	0.39 (0.89)	0.6651	-1.92 (1.08)	0.0761
ADCSADL	Lo	1.16 (0.62)	0.0602	-0.28 (0.91)	0.7611	-1.44 (1.10)	0.1914

A test for a difference in estimated high dose treatment effect at Week 78 between APOE+ and APOE- has a p-value of .0761 for ADCSADL, .071 for MMSE. For the low dose, a test for a difference in estimated treatment effect at Week 78 between APOE+ and APOE- has a p-value of .0167 for MMSE and .0158 for ADASCOG. All of these trends suggest bigger effects in APOE+ as compared to APOE-.

The sponsor tried to use a post-hoc propensity score matching analysis to suggest that there was a trend in study 301 for the high dose as a function of the number of 10 mg/kg doses (Figure 13).

The sponsor did not include standard errors of the bars a basic requirement for statistical inference, without which statistical inference is impossible, so differences are impossible to judge with high confidence. However, one can notice that study 302 shows no real difference between ≥ 6 , ≥ 8 , ≥ 10 , ≥ 12 , or $= 14$, which would suggest that the number of 10 mg/kg doses doesn't matter in study 302. On the other hand, the sponsor seems to suggest that in 301 the number of doses matters (an apparent linear trend in the bars). However, the linear trend in % difference in 301 may well be attributable to the simultaneous linear trend in placebo responses across the categories (≥ 8 to $= 14$: varying from 1.36 at the second lowest category [≥ 8] to 1.58 at the highest [$= 14$], i.e., the worst placebo response coinciding with the highest dose category). Furthermore, overall, the figure suggests again that the studies are not consistent (with respect to the importance of the number of doses in this case). The sponsor provided no assessment of how well the resulting groups were matched and it could be poor since the number of like patients decreases with the number of matching factors and, for example, enrollment window of every 200 subjects was a matching factor. Also, the reviewer has shown that the CDRSB outcomes in the non-US countries were not sufficiently similar to be pooled but this analysis pooled them. Therefore, this analysis does not incorporate these important regional differences into the matching and therefore the analysis may be confounded with regional differences in estimated Week 78 effect, on top of the unstable placebo response across the categories in 301. There are also differences in estimated Week 78 effect by APOE which this analysis does not properly address. The implications of the figure are also called into question by the fact that the low dose was better than the high dose in Study 301 despite having zero 10 mg/kg doses! In summary this post-hoc subgroup analysis is flawed and such a post-hoc subgroup analysis could never measure up to a prespecified primary analysis supported by randomization.

Figure 13 Bar plot of CDR Sum of Boxes Adjusted Mean Change from Baseline Percent Difference from Placebo at Week 78 by Number of 10 mg/kg Doses, with Placebo Selected by Propensity Score Matching - ITT Population that have had Opportunity to Complete Week 78 by 20Mar2019: Placebo-controlled Period Excluding Data after 20Mar2019 – **Nested Categories**



Number of Doses ■ >=6 ■ >=8 ■ >=10 ■ >=12 ■ =14 (full dose)

Number of subjects and Adjusted mean at Week 78										
Placebo	202	185	157	129	77	220	201	177	144	98
	1.42	1.36	1.49	1.54	1.58	1.56	1.45	1.54	1.38	1.41
BIIB037	202	185	157	129	77	220	201	177	144	98
	1.45	1.43	1.21	1.14	1.08	1.07	1.02	0.97	0.89	0.87

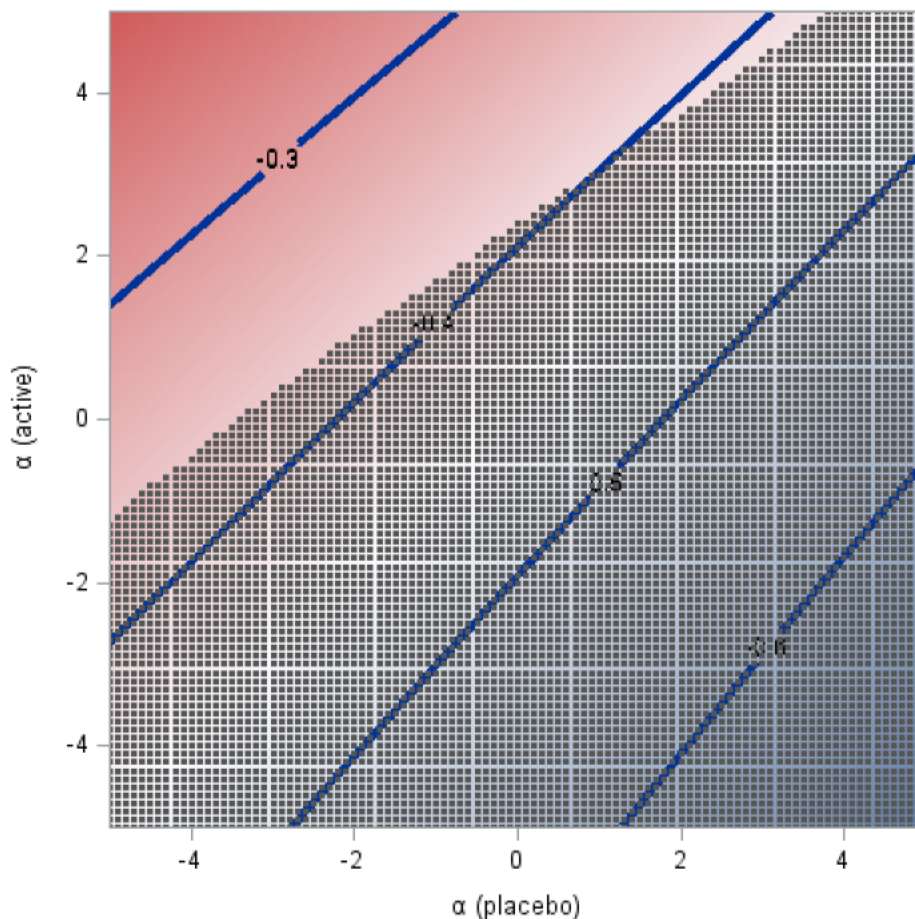
NOTE 1: Covariates in propensity score model include laboratory ApoE status, age, sex, baseline clinical stage, baseline scores of CDR-SB, MMSE, ADAS-Cog 13, ADCS-ADL-MCI, years of education, years since first AD symptom, AD symptomatic medication use at baseline, US/non-US and enrollment window of every 200 subjects. Placebo and treated subjects matched exactly on laboratory ApoE status. Subjects with undetermined laboratory ApoE status are grouped in the randomized ApoE subgroup.

NOTE 2: Results for each threshold were based on an MMRM (mixed model for repeated measures) model, with change from baseline in CDR-SB as dependent variable and with fixed effects of treatment group, categorical visit, treatment-by-visit interaction, baseline CDR-SB, baseline CDR-SB by visit interaction, baseline MMSE, AD symptomatic medication use at baseline, region, and laboratory ApoE status.

Note: The above figure was copied from page 80 of the sponsor's summary of clinical efficacy

Figure 14 shows a tipping point analysis of Study 302, combinations of informative missingness that could alter the significance (above the shaded grid region). Alpha (“active”) and Alpha (placebo) allow for adjustments to the mean response for dropouts independently for high dose and placebo. Combinations of the two above the dark shaded grid would suggest a loss of significance, e.g., if $\alpha(\text{“active”})$ is a little greater than 2 and $\alpha(\text{placebo})=0$ or anything combination of $\alpha(\text{“active”})$ and $\alpha(\text{placebo})$ above the line $2.2+1.9/3 * \alpha(\text{placebo})$. The same would be true for any combination of $\alpha(\text{“active”})$ and of $\alpha(\text{placebo})$ in the orange region at the upper left of the figure. Considering the overall outcome of Study 301 it is not unreasonable that high dose dropouts could be worse than placebo.

Figure 14. Study 302 CDRSB Tipping Point Analysis

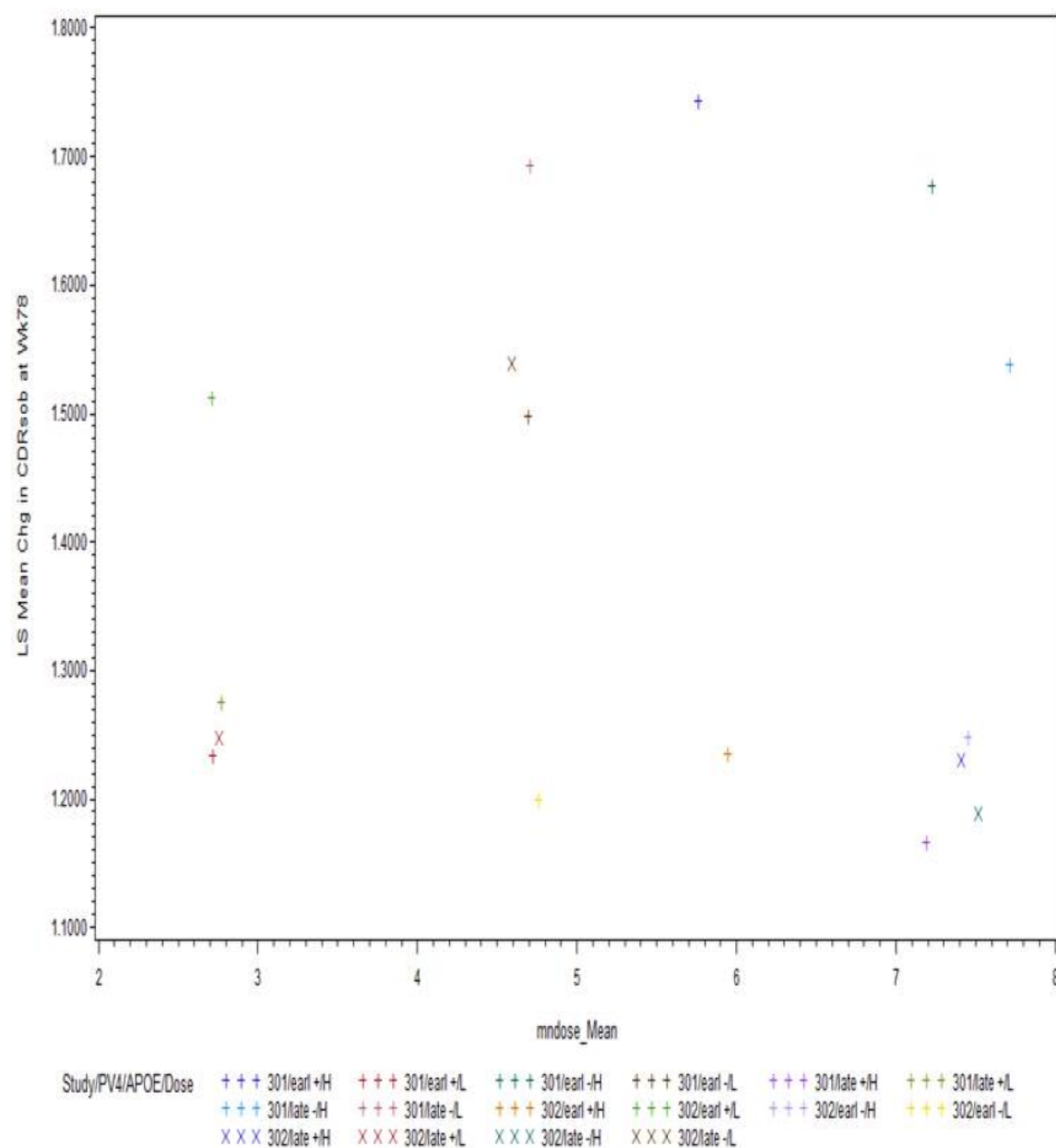


In Figure 15 since there is no correlation between achieved dose and outcome when low is plotted with high (not differencing from placebo) it suggests that the sponsor’s placebo differenced analysis of correlation is driven by changes in placebo outcomes between early and

late. Recall that the low dose was numerically better than high by 0.2 points in study 301 at Week 78 although the low dose had a much lower mean dose. This also seems to contradict their argument about intermediate dosing.

Notice also in the figure below that in 302 the late APOE+ low dose LS Mean which has the plotting symbol **X** is essentially the same as the late APOE + high dose with plotting symbol **X**.

Figure 15. Assessment of Correlation between Dose Achieved and Week 78 CDRSB across Low and High Doses (not placebo subtracted)



Biomarkers and Lack of Correlation between Clinical and Biomarker Changes within the High Dose

The SAP for biomarker analysis is dated 2020; it is not clear that it was prespecified.

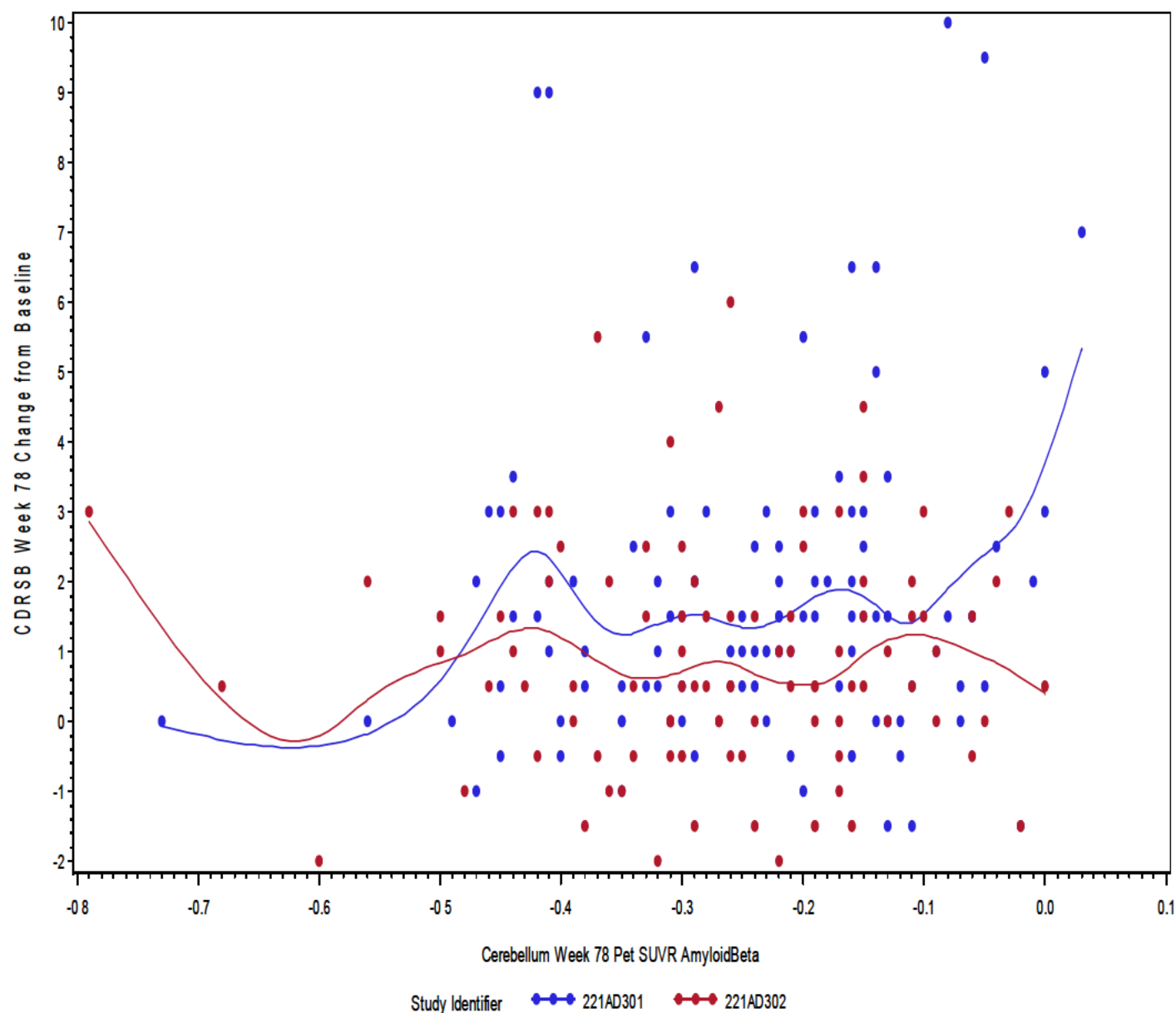
In the PET biomarker substudy (N=442 from 302) overall 29% placebo and 22% high dose were missing Week 78 and of those who had the opportunity to complete Week 78 (99/140 placebo) 12% of the placebo and 12% of the 113/(145) high dose were missing Week 78 assessment of cerebellum SUVR. The cerebellum SUVR was significant at Week 78 ($-.28 \pm .01$) as well as Week 26 ($-.09 \pm .01$). The low dose was also significant ($-.08$ at Week 26 and at Week 78 $-.18 \pm .01$). Placebo LS Mean cerebellum SUVR at Week 78 was $.0003 \pm .008$ in 301 and $+.0158 \pm .01$ in 302.

The PET substudy is not directly randomized or balanced. In PET substudy completers there were 12% more aged 71-80 in the high dose than placebo. 9-10% more age 71+ in the high dose and a 10% imbalance in APOE. There are 6% more aged 71-80 in the high dose than placebo in the PET population overall which was the group with the highest apparent estimated effect in study 302. Mean CDRSB at baseline in the PET subgroup was 2.39 in study 302 as compared to 2.51 in the high dose overall; in 301 it was also 2.39 in the PET subgroup but 2.40 in the high dose overall.

And why is the biomarker change positive in 301 (N=540 Hi vs Pbo $-.24 \pm .01$ at Week 78 and $-.07 \pm .008$ at Week 26) and the biomarker shows dose dependence (low dose $-.07 \pm .008$ at Week 26 and Week 78 $-.17 \pm .01$) but the clinical change in CDRSB is not significant and the low dose is numerically better than the high dose on CDRSB change from baseline at Week 78 in study 301? Correlations of Week 78 changes between the biomarker and CDRSB change within the high dose group are .135 for study 301 and $-.035$ for study 302 (see Figure 16 which includes a local regression curve to help the eye assimilate the data points). The correlation with the biomarker for the on face positive study 302 is in the wrong direction, i.e., the wrong side of no correlation! For example, the patient with the largest decrease in SUVR Amyloid beta in the Cerebellum $-.79$ had a Week 78 CDRSB that was a 3.0 point increase (worsening) from baseline. The patient with the second largest Week 78 decrease in Amyloid beta ($-.68$) in the Cerebellum via PET SUVR had a week 78 CDRSB, representing an increase of 0.5 from baseline. Pooling studies 301 and 302 the partial correlations adjusted for baseline CDRSB and baseline cerebellum SUVR was .137 ($p=.352$) and .182 ($p=.217$). For study 302 only, the baseline adjusted Pearson correlation was $-.031$ ($p=.889$) and the adjusted Spearman correlation was $+.116$ ($p=.589$). For 301 (N=24) the corresponding adjusted correlations were .219 (.327) and .278 (.210). In summary, for high dose patients the Week 78 CDRSB change from baseline and the Week 78 PET SUVR changes in A β in the Cerebellum (the prespecified primary brain

region) are uncorrelated which raises doubts about disease modification claims and substantial evidence of effectiveness in light of the mixed results of 301 and 302.

Figure 16. Assessing Correlation of Amyloid Pet and CDRSB in High Dose at Week 78



Secondary biomarkers also did not correlate well with primary efficacy outcome in the high dose as follows. Aducanumab appears to increase CSF A β 142. Thus, a negative correlation between CDRSB change and CSF A β change would support an association. There was a non-significant negative correlation within the high dose between CSF A β 142 and CDRSB change at Week

78. However, A β 140 was decreased by Aducanumab compared to placebo but not significantly and the correlation for A β 140 change and CDRSB change was positive and non-significant. With 301 and 302 pooled (N=25) CSF A β 1-42 Week 78 change correlations with CDRSB Week 78 change were -.07 (p=.74) and -.26 (p=.23). With 301 and 302 pooled (N=25) CSF A β 1-40 Week 78 change correlations with CDRSB Week 78 change were +.21 (p=.35) and +.17 (p=.43).

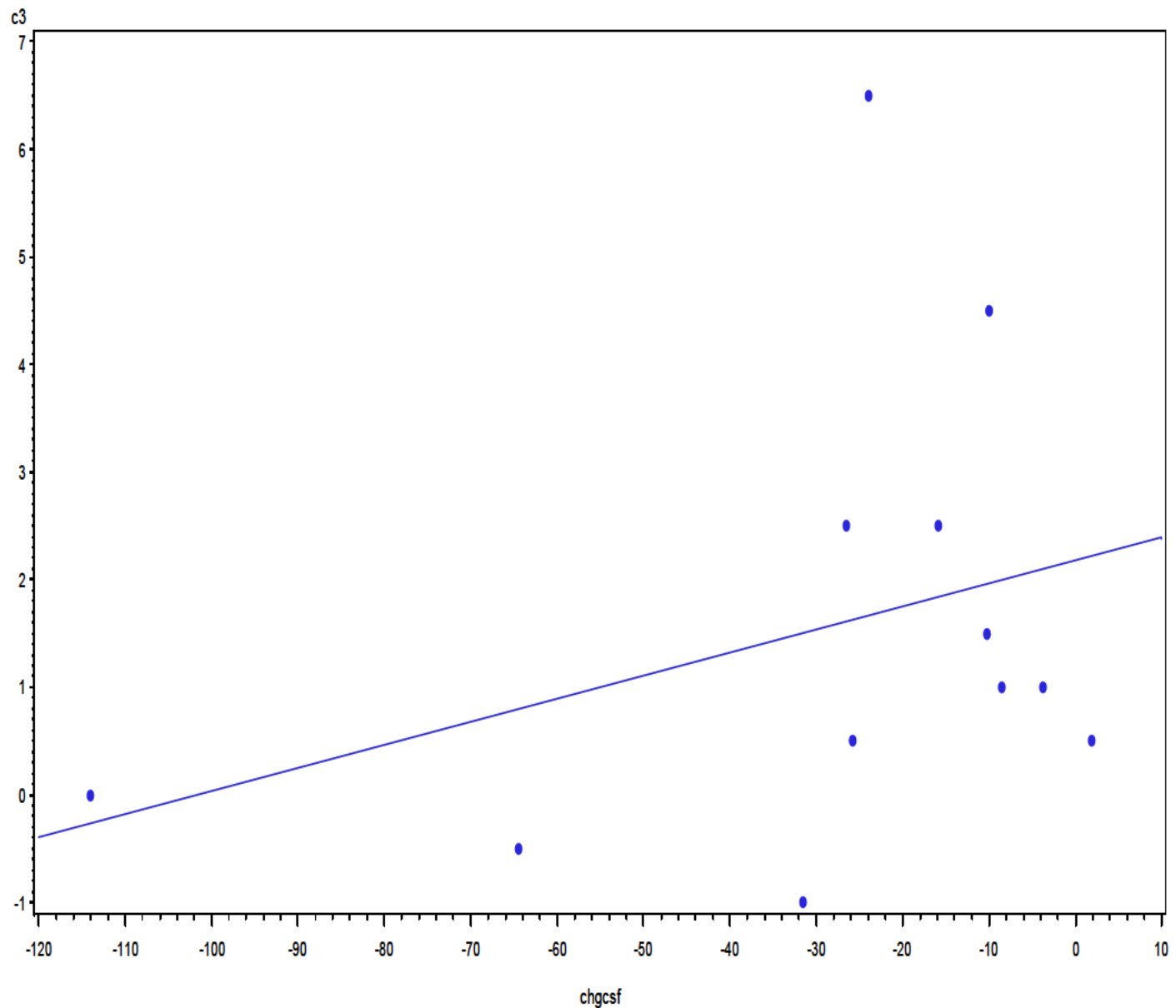
Aducanumab appears to decrease CSF ptau and tau protein. Thus a positive correlation between CDRSB change and CSF tau change would be required for a meaningful association between change on imaging and change on clinical endpoint. While the correlation with ptau is positive, the correlation with tau protein, another measured tau biomarker, is negative: -0.499.

Study 302 High Dose CDRSB Week 78 Change Correlations with Week 78 Biomarker Changes

	Pearson partial(p-value)	Spearman (P-value)
Correlation with CSF A β 140 n=13	.009 (.9787)	.334 (.315)
Correlation with CSF A β 142 n=13	-.236 (.4846)	-.453 (.1618)
Correlation with CSF ptau n=12	.634 (.0491)	.589 (.0731)
Correlation with CSF tauprot n=11	-.155 (.6942)	-.499 (.1719)

Figure 17 shows Correlation between Week 78 CSF Ptau and Week 78 CDRSB change within the High Dose group of Study 302 (with a fitted regression line). The Week 78 CDRSB change from baseline is identified by the y-axis and the Week 78 change in ptau is identified by the x-axis. The correlation is weak, e.g., the patient with the best CSF change did not improve on the CDRSB and was within 0.5 CDRSB points of the patient with the worst CSF ptau change at Week 78.

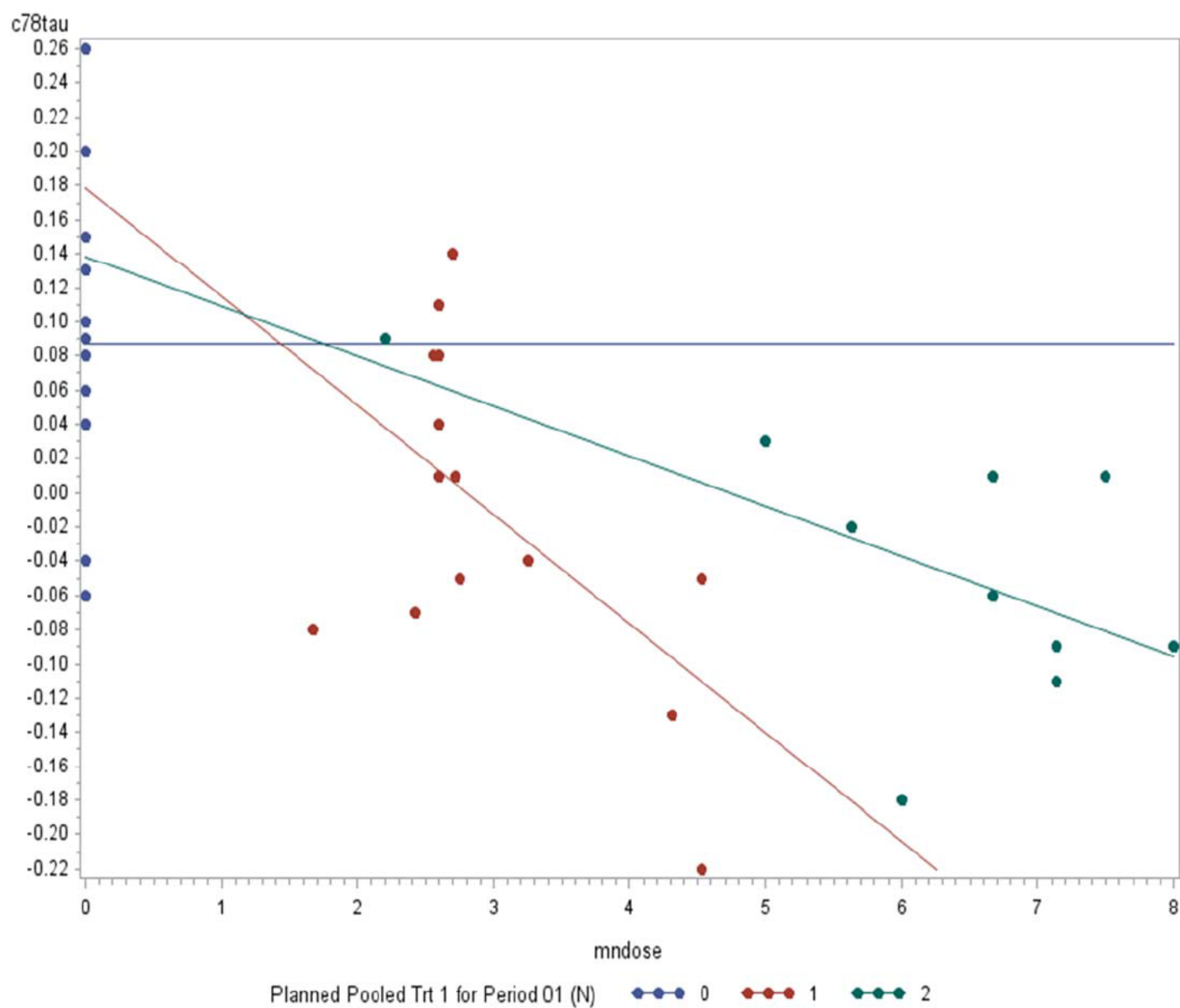
Figure 17. Correlation between Week 78 CSF PTau change and Week 78 CDRSB change within High Dose in 302



The sponsor also highlighted changes in the non-primary biomarker tau medial temporal SUVR at 78 weeks based on 12 placebo and 11 high dose patients. It is important to observe that in this small subgroup the low dose had a stronger slope/correlation between tau medial temporal SUVR and mean dose than did the high dose (Figure 18 which shows a fitted regression line for each dose 1 (red)=Low ; 2 (green)=High). This is only 23 patients total (excluding the low dose)

and tau is downstream in the disease process from the drug's target of Amyloid. Therefore, without a correlation between amyloid change and cdrsb change this possible relationship further downstream seems dubious. In fact, these two groups of 12 and 11 patients are not very well balanced. The baseline CDRSB was 2.75 for the 12 placebo and 2.27 for the 11 placebo, 8 vs. 18 % Mild disease, 33% vs. 55% used concomitant medications at baseline, average age 69 vs. 76 and 8% vs. 0% were from study 302. In short, this is a small sample and it's not clear that they were even comparable at baseline. There was no effect overall in study 301, the parent study from which most of these 23 patients came, so it is additionally hard to believe this tau finding in a tiny substudy. Furthermore, the correlation between tau and cdrsb change at week 78 in the high dose (all measured after the futility announcement) is not significant-- $-.35$ $p=.5272$ ($n=6$). This correlation should have been positive (>0) in order to support the high dose causality, i.e., reduction of tau translating to long term clinical change in CDRSB.

Figure 18. Week 78 Change in Medial Temporal Tau by Mean Dose



3.2.2 Study 301

3.2.2.1.1 Sponsor's Results

3.2.2.1.2 Reviewer's Results

Study 301 was negative overall with the high dose actually being numerically worse than placebo (.03 +/- .150 [-.262,.328], $p=0.8$). It can rule out a drug effect greater in magnitude than .262 with 95% confidence. The sponsor tried to find similar patients in 301 to 302 and highlighted an apparent high dose effect in those who were under the protocol version 4 amendment. However, there is a multiplicity problem with this (overall the trial was negative and looking at a subset that was not even prespecified there is only a partial randomization: it is only part of the process and there could be important exclusions). Since the sponsor made a lot of effort to find 301 subjects like 302 subjects this reviewer evaluated the consistency post-pV4 in study 301.

Some patients randomized as late as Sept 2017 by which one would have to be randomized in order to have the opportunity to complete by March 20, 2019) were not under protocol version 4 for whatever reason (patient consent/ IRB/site acceptance, etc.) , thus did not get 10 mg/kg dose or at least not for the maximum possible time if they were APOE+. The non-PV4 patients seem to have different baseline characteristics than those randomized under protocol version 4 in the same time frame. Once again, this means the effect of raising the dose in APOE+ is confounded with enrollment changes in baseline characteristics (e.g., country and baseline disease stage as shown in Figure 1).

In the APOE+ subgroup the high dose was in the wrong direction overall in study 301 $apoe+ 0.07014 \pm 0.1802 \quad p=0.6971$. (APOE 0.05766 ± 0.2716).

In the subset of data prior to protocol amendment 4 (N=844), overall, the high dose group was nominally significantly worse at Week 78 on CDRSB than the low dose ($p=0.0304$) [and worse than placebo at the .15 significance level].

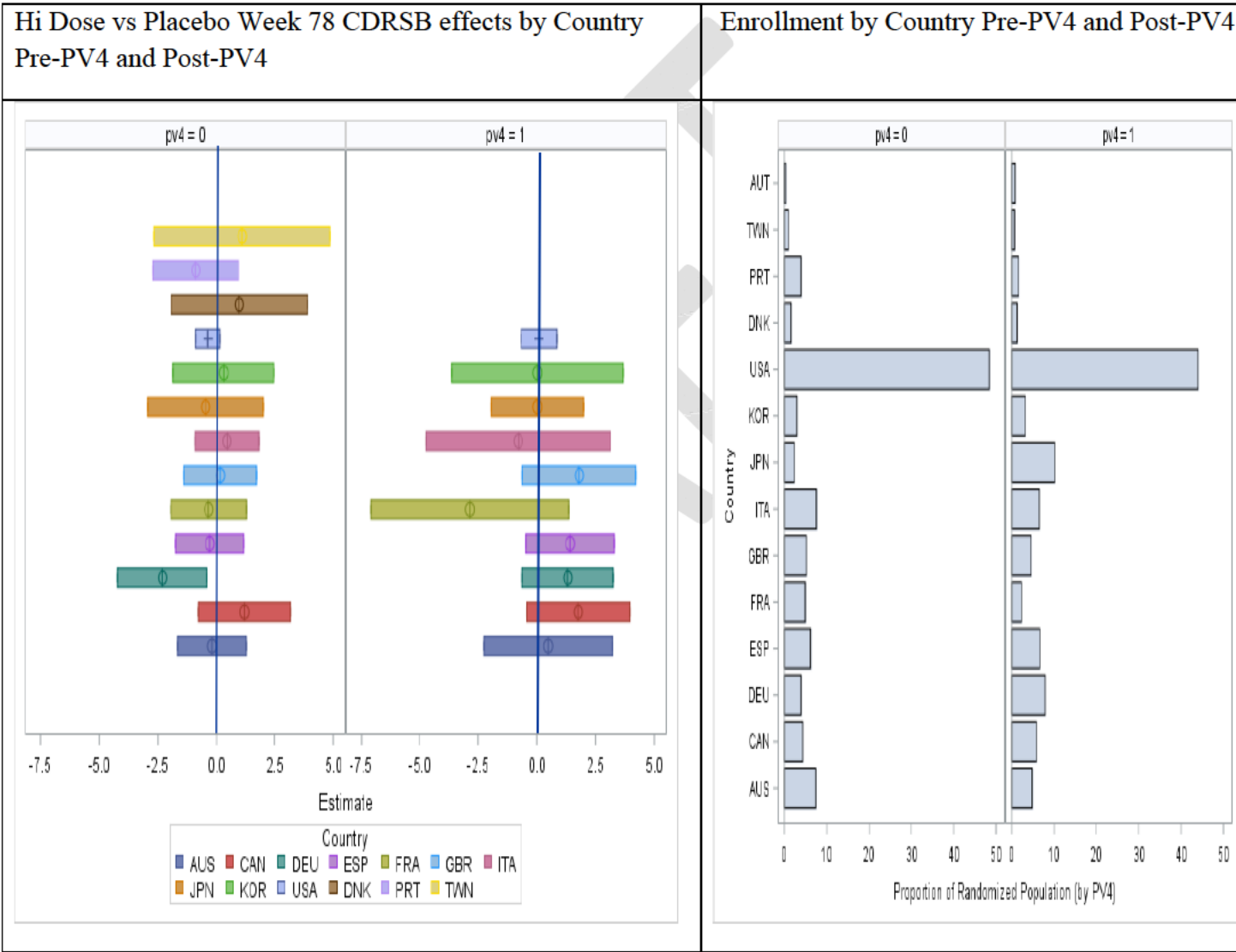
In the subset of data after protocol amendment 4 (and consenting to it by day 113) the estimated high dose treatment difference at Week 78 was in the correct direction: $N=787 \quad r=1589 \quad -.4868 \pm .2701 \quad p=.0724$.

However, in this post-PV4 subset there was a significant interaction with country*visit*and treatment (country main effect $p=.056$; 3 way Country by Visit by Treatment interaction: $p=.07$; all country terms $p=.0351$; all interaction effects beyond region main effect : $p=0.0625$ [101 numerator degrees of freedom]). Allowing for treatment differences over countries in the model and averaging over countries, the estimated average over countries high dose placebo subtracted effect at Week 78 is $.352 \pm .45 \quad p=.4308$, . If the coarser, prespecified Region effect is used instead of Country in the model the 3 way interaction Region*Visit*Treatment p-value is .0139. Thus, whether adjusting for prespecified region or Country the treatment differences appear to vary significantly by region or Country in the post-PV4 subset. The country interaction is even more compelling in the APOE+ subgroup (main $p=.0097$ 3 way .0548) and the overall high effect averaged over country non-significant $.657 \pm .537, p=.223$.

Figure 19 shows the hi vs. placebo estimated differences (± 1 StdError) by Country in pre-PV4 and post-PV4 subsets with positive differences favoring the high dose.

Even in the post-PV4 subset, the 301 high dose effect relative to placebo in the US the largest enroller is very small (and would not be nominally significant). Also, the post-PV4 high dose CDRSB effect in Japan is virtually zero although it was big in study 302.

Figure 19. Study 301 High Treatment Effect Estimates at Week 78 pre-PV4 and post-PV4



The estimated high dose effect in the post-PV4 subset also seems to vary significantly by age group. agegroup (agegroup main effect .0003; agegrp*visit*trt .0352; all terms involving agegroup.0031 18 df). Only the oldest group, >75 accounting for 34%, showed a nominal effect at Week 78 on CDRSB for the high dose and the <65 subgroup (accounting for 21%) was in the wrong direction (Table 9).

Table 9. Study 301 Post-PV4 subset High Dose vs Placebo CDRSB Treatment Effect Estimates by Age Group

Age Group	LSMean Difference at Week 78 +/- Std. Error
<65	-.16 +/- .6
65-74	.24 +/- .4
≥75	1.2 +/- .5

In study 301, the high dose effect in APOE- non-carriers is numerically better post-PV4 than in APOE+ (Table 10), but the dose increase for the high dose was in APOE+ only and pre-PV4 in APOE- there was essentially no effect (compare Figure 20 blue bars at Week 78).

Table 10. Study 301 Post-PV4 High -Placebo Estimated CDRSB Difference at Week 78 by APOE subgroups

APOE subgroup	Comparison	Estimated Difference (LSMean)	Std. Error	Nominal p-value
apoe+	Placebo - High Week 78	0.4689	0.3356	0.1633
apoe-	Placebo - High Week 78	0.5227	0.4922	0.2890

As can be seen in Figure 20, in the post-PV4 subset Placebo worsened at the same time as the high dose improved as compared to pre-PV4, but the high dose improved to a point little better than the low dose despite the higher and increased dose relative to pre-PV4 (LSMean 1.167 +/- .24 1.276 +/- .22).

Figure 20. Study 301 Placebo, Low Dose, and High Dose CDRSB Profiles in Study 301 APOE non-carriers Stratum

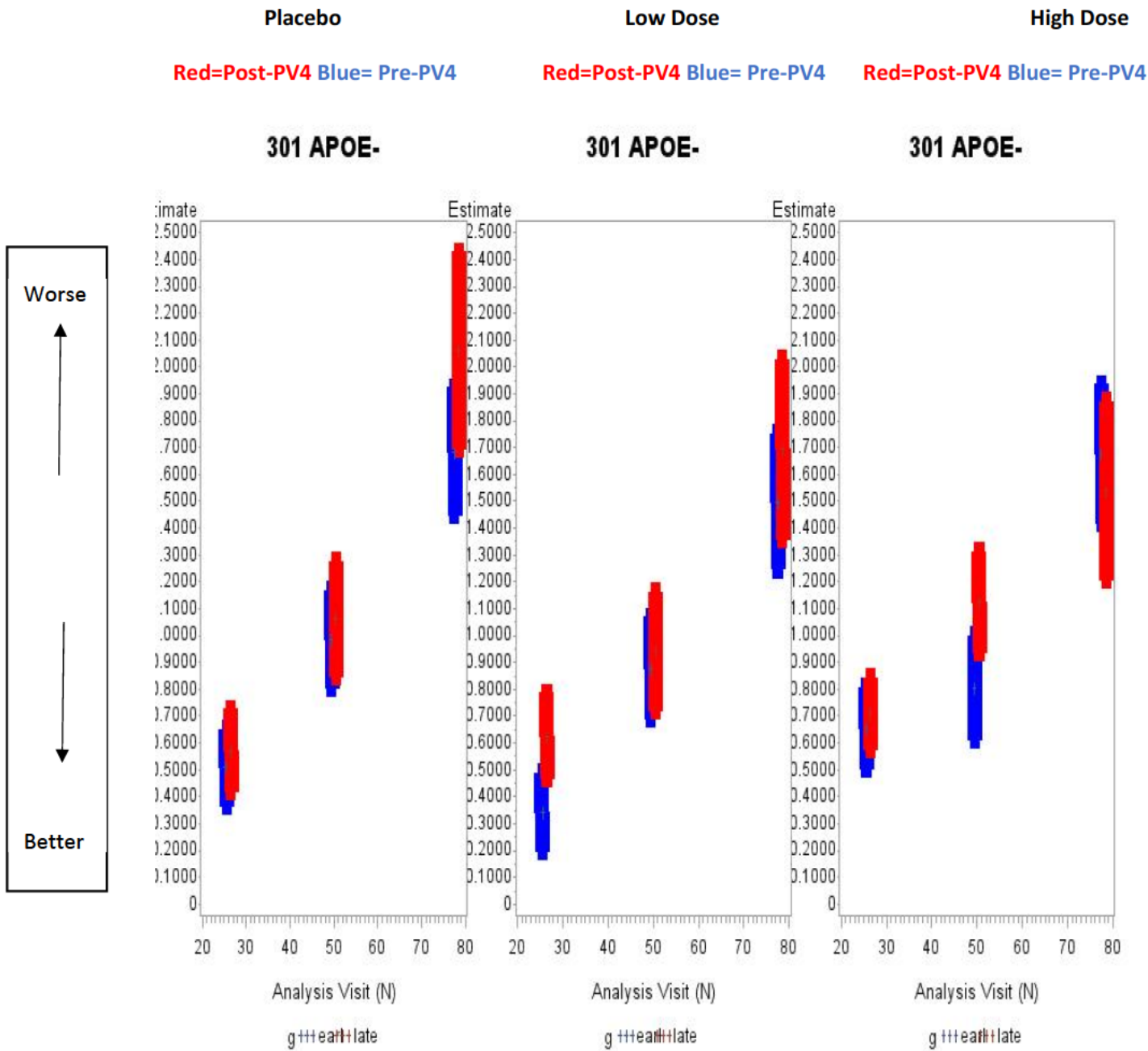
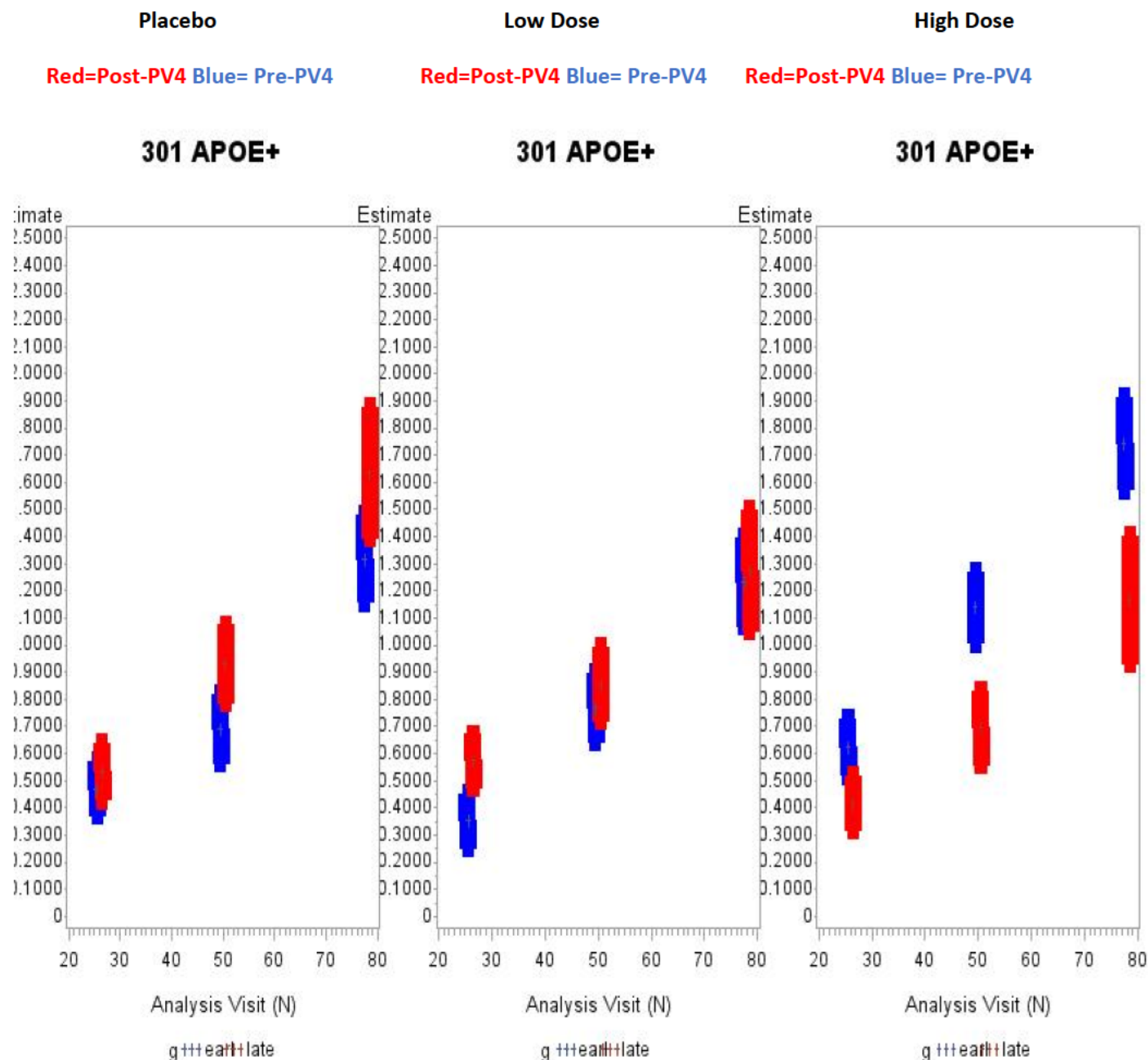


Figure 21 shows CDRSB profiles in Study 301 APOE + (carrier) stratum. While the high dose, on the right, improved post-PV4 relative to pre-PV4, placebo, on the left, also worsened from pre-PV4 to post-PV4. Also, comparing to the low dose, post-PV4 (red) high and low are almost the same, despite the high dose increase.

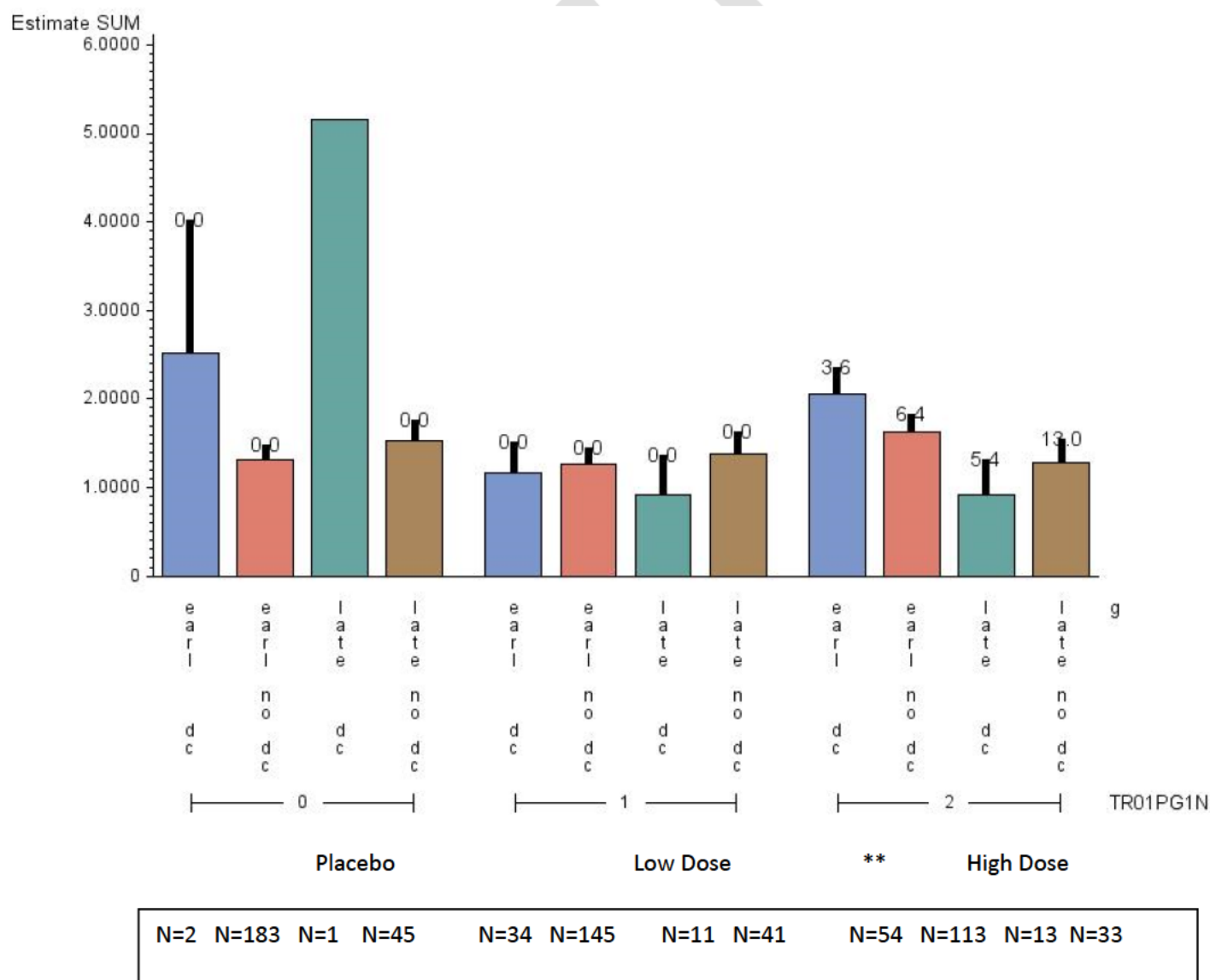
Figure 21 Study 301 Placebo, Low Dose, High Dose CDRSB profiles in Study 301 APOE carrier Stratum



The following graph shows the LS Mean Changes in CDR-sob at Week 78 for the APOE+ stratum by Early vs. Late and Dose Reduction due to ARIA status. The average number of 10 mg/kg doses completed by Week 78 is the value shown above the bars. It appears that the Late non-dose-reduced High Dose subgroup (brown) had significantly more 10 mg/kg doses, as

expected, but the LS Mean is numerically worse than for the Late dose-reduced High Dose subgroup (green). The sample sizes are admittedly small but it is not clear why in Figure 22, as in Figure 10 for study 302, the post-PV4 (late) subgroup not requiring dose modification or slowing, with more 10 mg/kg doses is numerically worse than the corresponding high dose subgroup that required dose modifications or slowing due to ARIA adverse events. There was some real and more potential unblinding in the latter subgroup. There also appears to be very little difference between the low (middle brown) and high dose (right brown) LS means post-PV4 after the high dose had been allowed to increase.

Figure 22. Study 301 LS Mean Change from Baseline in CDR-sob at Week 78 in APOE+ by PV4 and ARIA dose modification



Note: earl=Pre-PV4 late=Post-PV4; dc= dose titration slowing or reduction due to ARIA

Subjects with ARIA-E or H had additional MRIs every 4 or 2 weeks respectively including a biomarker and PK sample for ARIA-H until resolution, which being beyond the normal schedule could have resulted in unblinding and the sponsor “Prot1:Subjects should complete all scheduled clinic visits for assessments and, in addition, have an unscheduled visit for an MRI and MOCA approximately every 4 weeks until the ARIA-E has resolved per the centrally read MRI.

Investigators, study site staff (except for the designated unblinded pharmacist/technician), and study subjects will be blinded to the subjects’ randomized treatment assignment for the placebo-controlled period.

The central MRI reading center will report incident cases of ARIA-E and ARIA-H to both the Sponsor and the Principal Investigator within a specified time after observing the finding on MRI per the imaging manual procedures. All cases of ARIA will be reviewed by the Sponsor and the Principal Investigator; decisions on dosing continuation, interruption, or discontinuation will be based on clinical symptoms, and the MRI information provided by the central reader.

“Selective roles within study team at Biogen may have access to individual cases of ARIA in order to carry out the operational aspect of the study. They will not perform aggregate summaries or communicate ARIA related data to other team members who do not have access to ARIA data.”

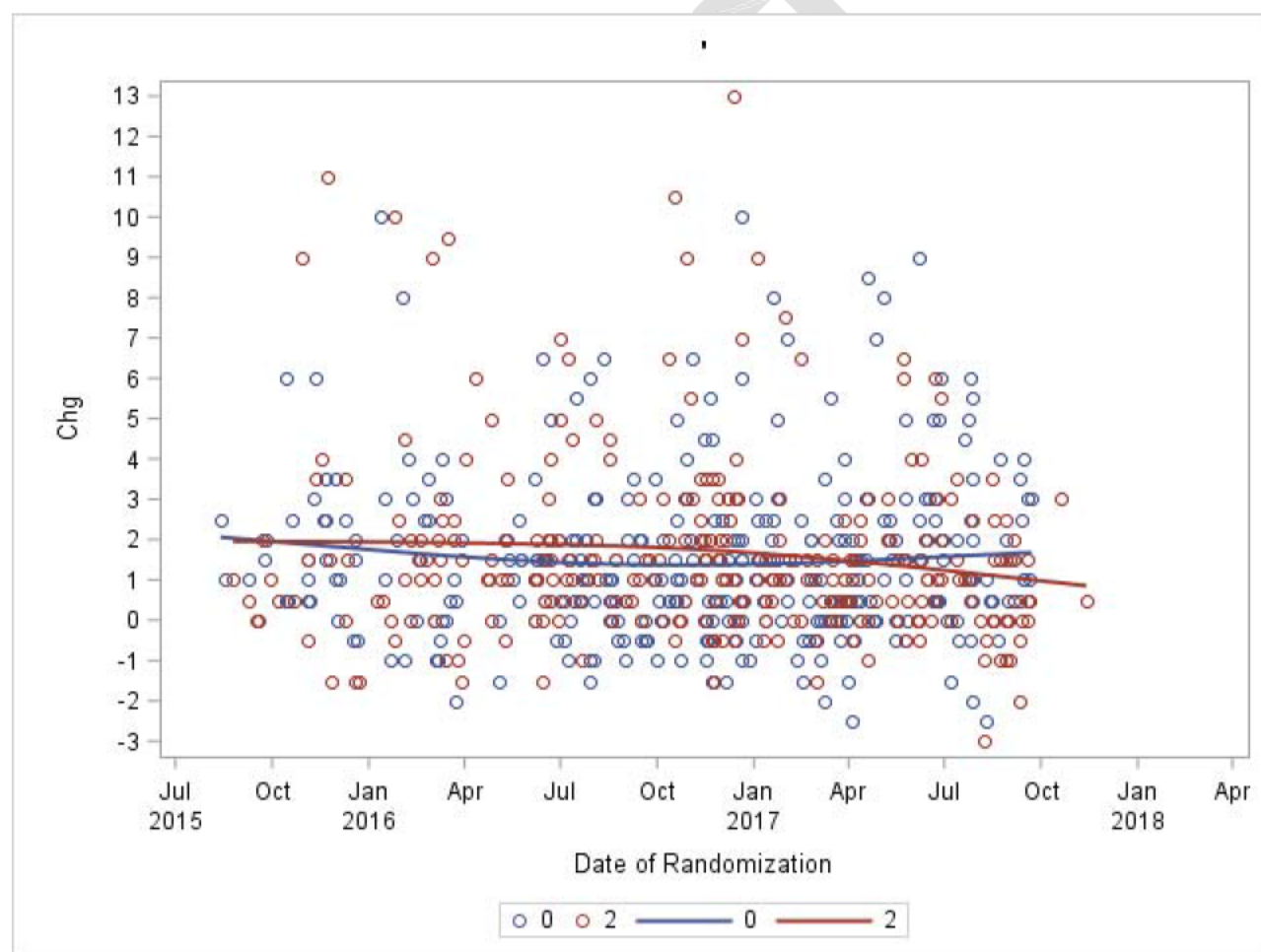
V4: Appropriate monitoring for ARIA has been implemented in this study including MRIs 2 weeks after the 4th dose (when subjects would have received 2 doses of 3 mg/kg and every 2 doses thereafter through the 8th dose (2 weeks after subjects in the 10 mg/kg group have received the 2nd administration of the target dose) and at least every 3 doses thereafter through Week 78”.

Protocol Version 4 was not implemented at the same time across sites and patients had to provide consent to the version, so there is not a unique implementation time across the study and which patients it applies to might have multiple possible definitions. The sponsor’s definition might not be the only possible definition, e.g., there are APOE + patients who had some 10 mg/kg doses but are classified as non-PV4 by the sponsor because they didn’t have the opportunity for all 14 10 mg/kg doses. There is a multiplicity issue if one evaluates the PV4 subgroup in addition to the overall study population. PV4 also added a MRI monitoring assessment for ARIA at Week 66. Since this drug related adverse event required some unblinding it could have created the potential for operational bias (even some sponsor personnel were involved in individual dose management decisions). Although they are subgroup estimates one would not expect the high

dose managed (decreased or slowed titration) subgroup due to ARIA to have a numerically better Week 78 outcome than the group that did not require dose management due to ARIA yet that is what Figure 21 shows. It seems to either suggest the possibility of operational bias due to ARIA related unblinding or at least conflicts with the claim that more 10 mg/kg doses lead to more effect vs. placebo.

Study 301 local trend in CDRSB change at Week 78 over randomization time for placebo (0 blue) and high dose (=2 red) is shown in Figure 23.

Figure 23 Study 301 local trend in CDRSB change at Week 78 over randomization time



3.2.3 Study 103

Protocol 221AD103 Version 11 (2018)

Primary Objective

The primary objective of the study is to evaluate the safety and tolerability of multiple doses of aducanumab in subjects with prodromal or mild AD.

This is a Phase 1b, multicenter, randomized, double-blinded, placebo-controlled, multiple dose study of aducanumab in subjects with prodromal or mild AD (Mini Mental State Examination [MMSE] scores from 20 to 30). The study will be conducted with a staggered, parallel group design, with the first 3 treatment arms (Arms 1-3) conducted in parallel, followed by Arms 4-5 beginning in parallel, followed by Arms 6-7 beginning in parallel, and finally followed by Arms 8-9 beginning in parallel. Arms 1-3 will comprise 2 aducanumab arms (1 and 3 mg/kg) and 1 placebo arm; Arms 4-5 will comprise 1 aducanumab arm (10 mg/kg) and 1 placebo arm; Arms 6-7 will comprise 1 aducanumab arm (6 mg/kg) and 1 placebo arm. Arms 8-9 will comprise 1 aducanumab titration arm (1 mg/kg for the first 2 doses, 3 mg/kg for the next 4 doses, 6 mg/kg for the next 5 doses, and 10 mg/kg thereafter) and 1 placebo arm. Titration up to 10 mg/kg in Arms 8-9 will only be implemented after the DMC's review of all the accumulated safety prior to the first subject receiving 10 mg/kg. Based on subsequent reviews, the DMC will determine whether titration up to 10 mg/kg in Arms 8-9 should continue. Approximately 188 subjects will be enrolled in total across approximately 35 centers. See Figure 24 for the study design. Subjects will receive 14 doses of aducanumab or placebo that will be administered by IV infusion once every 4 weeks.

Arms 1-3: approximately 80 subjects in total:

- ☐ Arm 1: 1 mg/kg aducanumab per dose (30 subjects)
- ☐ Arm 2: 3 mg/kg aducanumab per dose (30 subjects)
- ☐ Arm 3: Placebo (20 subjects)

Arms 4-5: approximately 40 subjects in total:

- ☐ Arm 4: 10 mg/kg aducanumab per dose (30 subjects),
- ☐ Arm 5: Placebo (10 subjects)

Arms 6-7: approximately 40 subjects in total:

- ☐ Arm 6: 6 mg/kg aducanumab per dose (30 subjects)
- ☐ Arm 7: Placebo (10 subjects)

Arms 8-9: approximately 28 subjects in total:

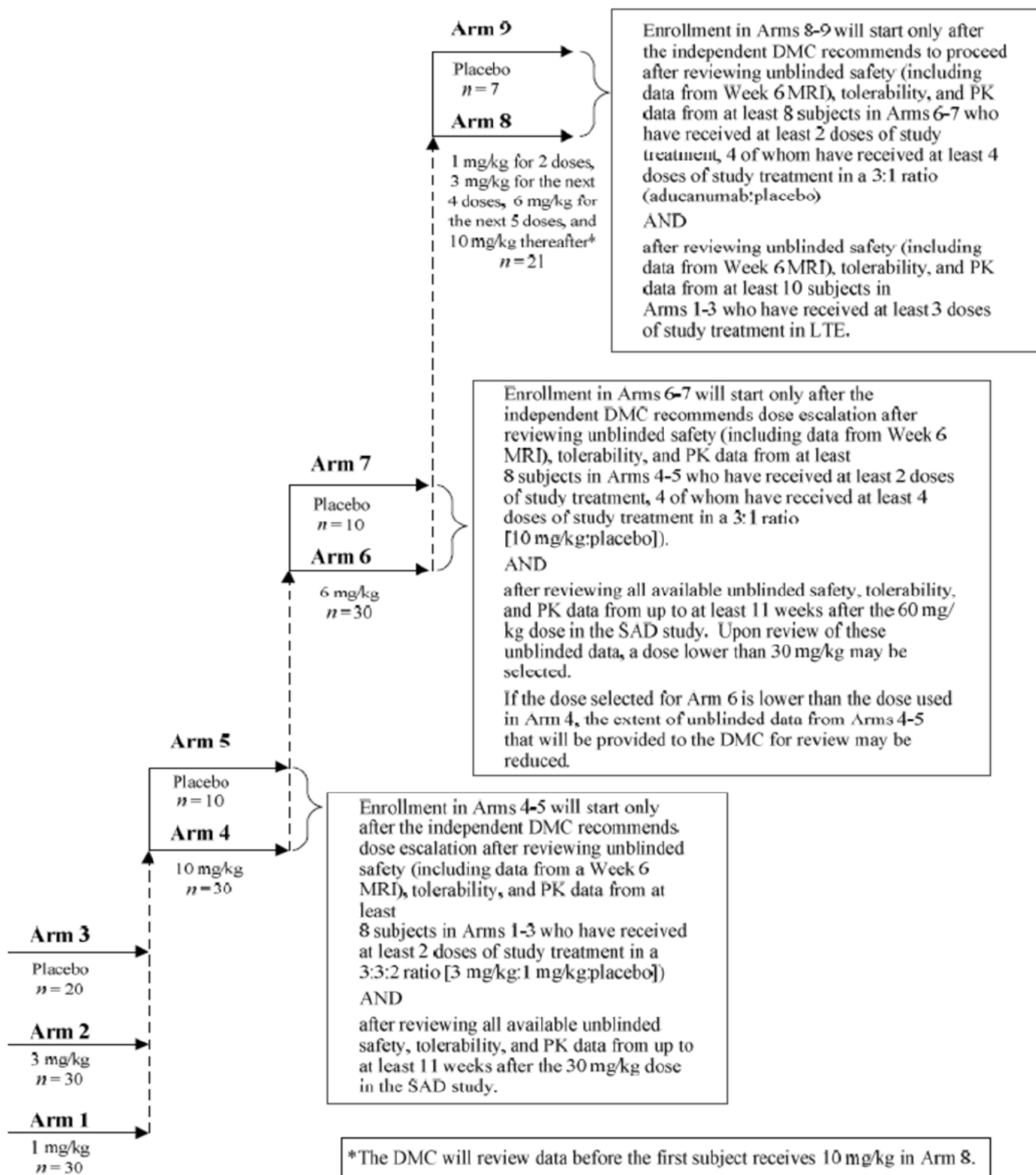
- ☐ Arm 8: aducanumab titration arm, 1 mg/kg for the first 2 doses, 3 mg/kg for the next 4 doses, 6 mg/kg for the next 5 doses, and 10 mg/kg thereafter (21 subjects)
- ☐ Arm 9: Placebo (7 subjects)

Subjects will be randomized into each treatment group within Arms 1-3, Arms 4-5, Arms 6-7, and Arms 8-9. The randomization will be stratified by the ApoE4 status (carrier or non-carrier), with the exception of Arms 8-9, which will contain ApoE4 carriers only. Enrollment in Arms 1, 2, and 3 will occur in parallel. Once enrollment in Arms 4 and 5 is open, the enrollment of Arms 1, 2, and 3 and Arms 4 and 5 will occur in parallel. Enrollment in Arms 6 and 7 will start once enrollment in Arms 1-5 is completed; enrollment in Arms 6 and 7 will occur in parallel.

Enrollment in Arms 8-9 will start once enrollment in Arms 6-7 is completed; enrollment in Arms 8-9 will occur in parallel. The assignment of subjects to Arms 1-3, Arms 4-5, and Arms 6-7 will then be made in a way to ensure the distribution of subjects is comparable between Arms 1-3, Arms 4-5, and Arms 6-7 with respect to the ApoE4 status; the ratio of ApoE4 carriers to non-carriers will be no more than 2:1 and no less than 1:2. All subjects in Arms 8-9 will be ApoE4 carriers.

Subjects who meet the LTE inclusion/exclusion criteria will be eligible to enter the LTE for an additional 42 intravenous doses of aducanumab during the first 3 years of the LTE, once every 4 weeks, with the first dose administered approximately 4 weeks after the final dose in the placebo-controlled portion of the study.

Figure 24 Study 103 Staggered Arm Design



Note: Figure copied from page 40 of sponsor's study report

Randomization will take place across all study sites using a centralized interactive voice/web response System (IXRS).

In Arms 1-3, approximately 80 subjects will be randomized 3:3:2 of aducanumab 1 mg/kg, 3 mg/kg, or placebo. In Arms 4-5, approximately 40 subjects will be randomized to a dose of 10 mg/kg aducanumab or placebo in a 3:1 ratio. In Arms 6-7, approximately 40 subjects will be randomized to a dose of 6 mg/kg aducanumab or placebo in a 3:1 ratio. In Arms 8-9, approximately 28 subjects will be randomized 3:1 of aducanumab (1 mg/kg for the first 2 doses, 3 mg/kg for the next 4 doses, 6 mg/kg for the next 5 doses, and 10 mg/kg thereafter) or placebo. The randomization will be stratified by the ApoE4 status (carrier or non-carrier), with the exception of Arms 8-9, which will contain ApoE4 carriers only. Enrollment in Arms 1, 2 and 3 will occur in parallel. Once enrollment into Arms 4 and 5 is open, the enrollment of Arms 1, 2, and 3 and Arms 4 and 5 will occur in parallel. Enrollment in Arms 6 and 7 will start once enrollment in Arms 1-5 is completed; enrollment in Arms 6 and 7 will occur in parallel.

-

Enrollment in Arms 8-9 will start once enrollment in Arms 6-7 is completed; enrollment in Arms 8-9 will occur in parallel. The assignment of subjects to Arms 1-3, Arms 4-5, and Arms 6-7 will then be made in a way to ensure the distribution of subjects is comparable between Arms 1-3, Arms 4-5, and Arms 6-7 with respect to the ApoE4 status; the ratio of ApoE4 carriers to non-carriers will be no more than 2:1 and no less than 1:2. All subjects in Arms 8-9 will be ApoE4 carriers.

Subjects who discontinue study treatment or withdraw prematurely from the placebo-controlled portion of the study may be replaced. Replacement subjects will be assigned to the same group (i.e., Arms 1-3, Arms 4-5, Arms 6-7, or Arms 8-9) as the subject(s) that withdrew and will be randomized into a treatment group within Arms 1-3, Arms 4-5, Arms 6-7, or Arms 8-9, as applicable.

Blinding Procedures

This study consists of a randomized, double-blinded, placebo-controlled portion, followed by a dose-blinded LTE with all subjects receiving aducanumab.

For the double-blinded placebo-controlled portion all study staff (except for a designated Pharmacist/Technician and the independent DMC) and subjects will be blinded to the subject treatment assignments. To maintain the study blind, it is imperative that subject treatment assignments are not shared with the subjects, their families, caregivers, legal representatives, or any member of the site study staff. The Biogen aducanumab study team will be blinded until the interim analysis on the Week 26 data; information from Arms 1-5 will be unblinded after all subjects in Arms 1-5 have completed tests and evaluations at the Week 26 Visit, information from Arms 6-7 will be unblinded after all subjects in Arms 6-7 have completed tests and evaluations at the Week 26 Visit, and information from Arms 8-9 will be unblinded after all subjects in Arms 8-9 have completed tests and

evaluations at the Week 26 Visit. Week 26 information will only be unblinded to a limited team within Biogen.

Analysis Population

The PD analysis population is defined as all subjects who were randomized, received at least one dose of study treatment, and have at least one post-baseline assessment of the parameter being analyzed.

Methods of Analysis

The change from Baseline to Week 26 and Week 54 in amyloid plaque burden as measured by PET will be summarized overall and by ApoE4 status for each treatment group and will be analyzed by analysis of covariance adjusting for baseline amyloid plaque burden and the ApoE4 status to assess the dose response and pairwise comparison with placebo. The analysis may be performed by the ApoE4 status. Due to the exploratory nature of this study, there will be no multiple comparison adjustment.

Change from baseline in CDR and other cognitive measures will be summarized by treatment group. Disease-related biomarkers, change from baseline in morphometric MRI measures, cerebral blood flow (by ASL-MRI), functional connectivity (by tf-fMRI), and glucose metabolism (by FDG PET) will be summarized by treatment group. Analysis of covariance or its non-parametric equivalent may be used to analyze these exploratory endpoints.

Interim Analyses

Interim analysis will be performed for the purpose of planning future studies of aducanumab, and no change will be made for this study based on the interim analysis results. A limited team from Biogen will be unblinded at the interim analysis. Up to 6 interim analyses for placebo-controlled period data may be performed.

- After all subjects in Arms 1-7 have completed tests and evaluations at the Week 26 Visit. Analysis of the primary PD endpoint (change from baseline to Week 26 in ^{18}F -AV-45 PET signal in certain brain areas) will be included in this interim analysis.
- After all subjects in Arms 1-5 have completed the Week 26 Visit
- After all subjects in Arms 8-9 have completed the Week 26 Visit.
- After all subjects in Arms 1-5 have completed the Week 54 Visit.
- After all subjects in Arms 6-7 have completed the Week 54 Visit.
- After all subjects in Arms 8-9 have completed the

Week 54 Visit. The Sponsor may perform additional interim analyses.

Sample Size Considerations

- In Arms 1-3, approximately 80 subjects will be randomized 3:3:2 to aducanumab 1 mg/kg, 3 mg/kg, or placebo. In Arms 4-5, approximately 40 subjects will be randomized 3:1 to 1

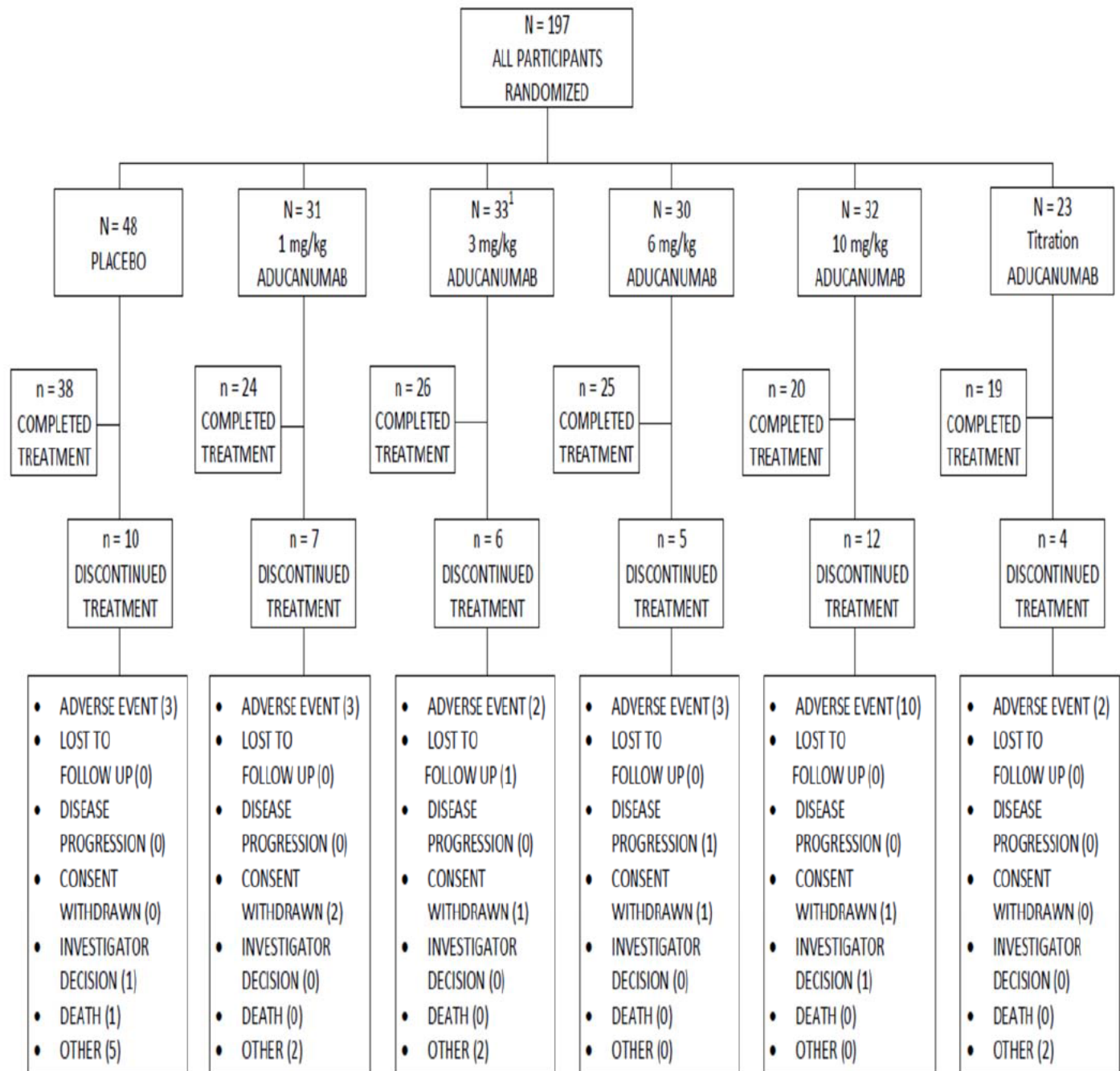
- aducanumab treatment arm (10 mg/kg) or placebo. In Arms 6-7, approximately 40 subjects will be randomized 3:1 to 1 aducanumab treatment arm (6 mg/kg) or placebo. In Arms 8-9, approximately 28 subjects will be randomized 3:1 to 1 aducanumab treatment arm (1 mg/kg for the first 2 doses, 3 mg/kg for the next 4 doses, 6 mg/kg for the next 5 doses, and 10 mg/kg thereafter) or placebo. Combining all 9 arms, there will be approximately 188 subjects of which approximately 30 subjects will be in each treatment group (1, 3, 6, and 10 mg/kg), approximately 21 subjects in the titration group (up to 10 mg/kg), and approximately 47 subjects in the placebo treatment groups.
- The primary PD endpoint is change from baseline to Week 26 in ^{18}F -AV-45 PET signal in certain brain areas. A sample size of 30 subjects per treatment group will provide over 90% power to detect a treatment difference of 1 standard deviation with respect to the reduction of amyloid from baseline, based on comparison of each aducanumab group with placebo,
- at a two-sided significance level of 0.05, and assuming a dropout rate of 20%. In addition, under the same assumptions, a sample of 21 subjects per treatment group will provide over 80% power. Due to the exploratory nature of this trial, no formal adjustment for multiplicity will be performed.
- Whole blood samples will be obtained for ApoE genotyping at Screening. Subject enrollment will be monitored (Arms 1-7) so that the ratio of ApoE4 carriers to non-carriers will be no more than 2:1 and no less than 1:2. If the treatment effect differs according to ApoE4 status, this sample size will provide at least 74% power within each ApoE4 stratum at a one-sided significance level of 0.1.

Reviewer Comment: only descriptive analysis methods were prespecified. Clearly, this is a learning trial.

3.2.3.1.1 Sponsor's Results

Figure 25 shows patient disposition in study 103 with all placebo arms grouped together.

Figure 25 Study 103 Patient Disposition



¹ 1 participant did not receive any doses of study treatment.

Note: figure copied from page 65 of sponsor's study report

The sponsor submitted the interim analysis 2 (IA-2) of Study 103 for breakthrough designation.

Table 11 shows the sponsor's final analysis of study 103, based on by-Visit ANCOVA model.

Reviewer's Comment: This completer's analysis would not be adequate for a primary analysis when dropouts are not unexpected.

Table 11 Sponsor's Completer's Analysis of Study 103

	Placebo	BIIB037 1 mg/kg	BIIB037 3 mg/kg	BIIB037 6 mg/kg	BIIB037 10 mg/kg	BIIB037 titration
Change from baseline at Week 54						
n	39	23	27	26	23	21
Adjusted mean	1.89	1.69	1.33	1.09	0.63	0.70
Standard error	0.350	0.441	0.413	0.423	0.446	0.499
95% CI	(1.198, 2.581)	(0.819, 2.564)	(0.517, 2.149)	(0.258, 1.930)	(-0.251, 1.511)	(-0.287, 1.686)
Difference with placebo		-0.20	-0.56	-0.80	-1.26	-1.19
95% CI for difference		(-1.308, 0.912)	(-1.612, 0.499)	(-1.855, 0.264)	(-2.356, -0.163)	(-2.343, -0.037)
p-value (comp. to placebo)		0.7249	0.2995	0.1398	0.0246	0.0432

NOTE: Adjusted mean for each treatment group, difference with placebo, 95% confidence interval and p-value were based on ANCOVA model at each timepoint. ANCOVA model was fitted with change from baseline as dependent variable, and with categorical treatment, baseline value and laboratory ApoE status (carrier and non-carrier) as independent variables.

- (a) The efficacy analysis population is defined as all subjects who were randomized, received at least one dose of study treatment, and have both baseline and at least one post-baseline questionnaire assessment.
- (b) Baseline population includes subjects in efficacy analysis population who have both baseline and at least one post-baseline assessment for CDR sum of boxes.

Note: this table was copied from page 110 of the sponsor's study report

3.2.3.1.2 Reviewer's Results

There was some potential for operational bias in this study due to having up to six interim analyses with some sponsor personnel unblinded while the study was ongoing. Cases of ARIA were managed by the Biogen and the investigator.

The staggered arm design means that there is no direct dedicated randomization to validate dose response analysis (comparison of doses) or ITT interpretation of the comparison of individual doses to pooled placebo (placebo arms 3,7, and 9 had no chance of being randomized to 10 mg/kg and placebo arms 3 and 9 were not concurrent with 10 mg/kg [arm 4]). None of the direct comparisons directly supported by the staggered randomization(s) are nominally significant. This is a serious limitation as compared to a typical confirmatory phase 3 design.

Only the Arm 5 placebo is concurrent with the 10 mg/kg group. The estimated difference based on only these two groups was -1.13, +/- .70, p=0.1178. For the titration to 10 mg/kg group the difference against the same cohort placebo was -.57 +/- 1.08 p=0.6046.

Table 12 compares the sponsor's non-randomization supported analyses and the fully supported randomization based analyses.

Table 12 Study 103 Randomization Supported and Sponsor's non fully randomization supported Analyses

Comparison	LS Mean Treatment Difference	Standard Error of Difference	Nominal p-value
10 mg/kg vs same cohort placebo (randomization supported)	1.13	0.70	.1178
10 mg/kg vs pooled placebo (not randomization supported)	1.08	0.54	.0462*
Titration to 10 mg/kg arm vs. same cohort placebo (randomization supported)	0.57	1.08	.6046
Titration to 10 mg/kg arm vs. pooled placebo (not randomization supported)	0.73	0.57	.2044

*loses significance when post randomization starting of approved AD medications data is censored: p=0.09

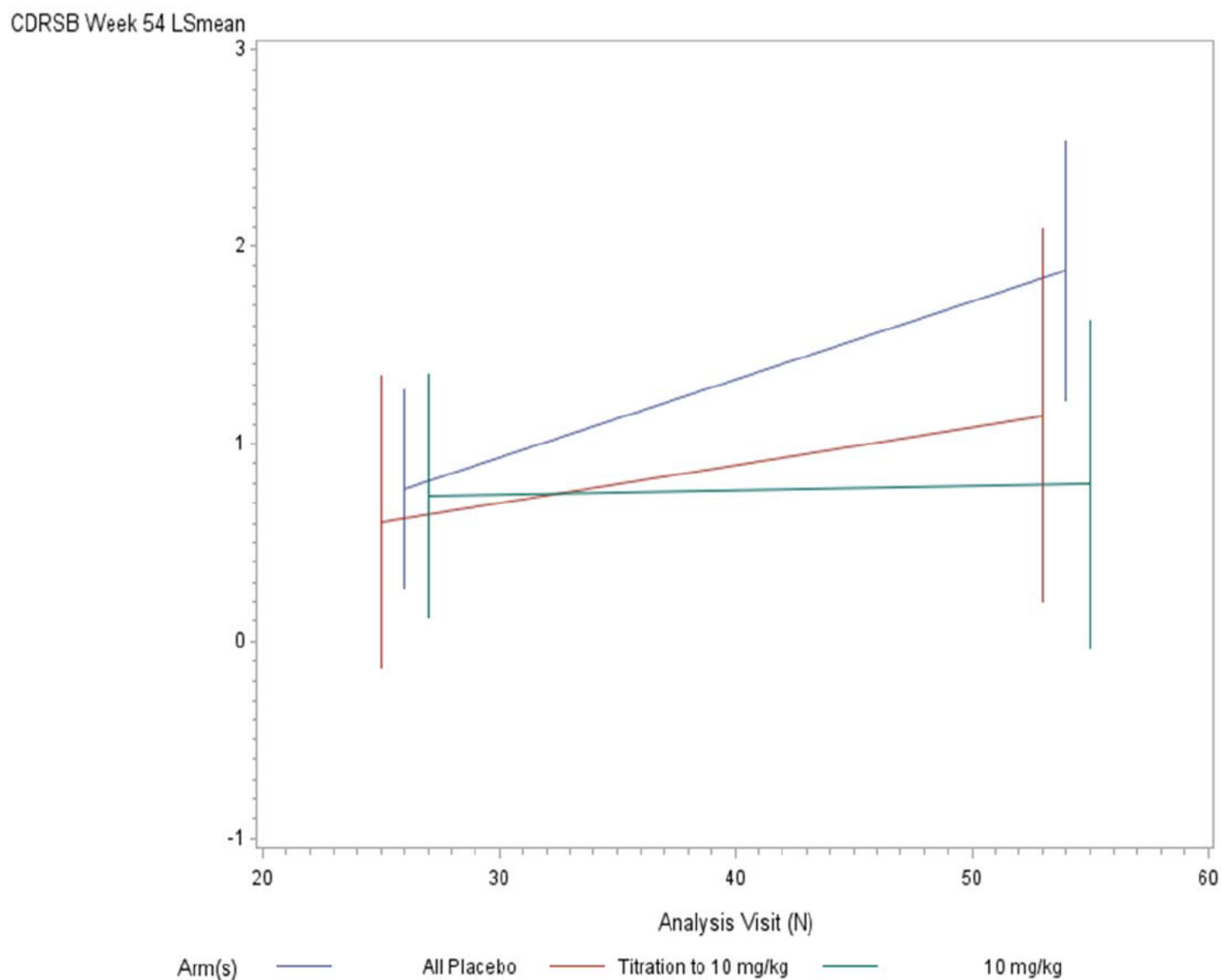
In the final study report the sponsor presents the analysis of Week 54 from an ANCOVA of Week 54 data only instead of the more appropriate MMRM (given the presence of dropouts 11% placebo [10% concurrent placebo] and 18% in 10 mg/kg). There were 20 patients excluded from the ANCOVA.

Their ANCOVA result for 10 mg vs. **pooled** placebo was $1.26 \pm .556$, $p=0.0246$. The MMRM has a less favorable p-value for 10 mg vs. pooled placebo: 1.08 ± 0.537 $p=0.0462$.

The MMRM estimate of the titration group vs pooled placebo comparison is not significant.: $.7317 \pm .5744$, $p=0.2044$ as compared to the ANCOVA $1.19 \pm .5838$, $p=.0432$. Dropouts were 5/44 for placebo overall and 1/22 for the titration group. The one titration group dropout had a 6.5 point increase from baseline to Week 26 and then dropped out. The ANCOVA seems to be biased by informative censoring of this Week 26 CDRSB value: the week 54 LS Mean for the titration group based on ANCOVA is .70 as compared to 1.14 for MMRM.

Figure 26 shows that neither the 10 mg/kg or the titration to 10 mg/kg had an effect at Week 24 on CDRSB.

Figure 26 Study 103 CDRSB profile for 10 mg/kg groups and Overall Pooled placebo



Three of the placebo (a fourth started right after) and three of the BIIB037 10 mg/kg had started AD medication after baseline and prior to the Week 54 assessment. If these subjects are excluded from the analysis the difference on CDRSB is (167 patients; 311 records) $.8957 \pm .5339$ $p = .0954$. For the titration group the result after excluding data from post-baseline started AD medications is $.442 \pm .561$, $p = .432$.

In the sponsor's 10 mg/kg vs. pooled placebo comparison, more of the effect was in the APOE- subgroup in study 103 which is the opposite of what was seen in study 302. There was no significant interaction between treatment and APOE but this would likely be underpowered (interaction test for hi dose vs. placebo at Week 54 by APOE, $p = 0.3746$).

apoe+ .883 +/- .642 .p=.1716
apoe- 1.63 +/- .956 p=.0906

There were also influential sites for study 103: w/out site 228 the 10 mg/kg estimated difference at Week 54 is 1.07 +/- .55, $p = .0561$ 162 df

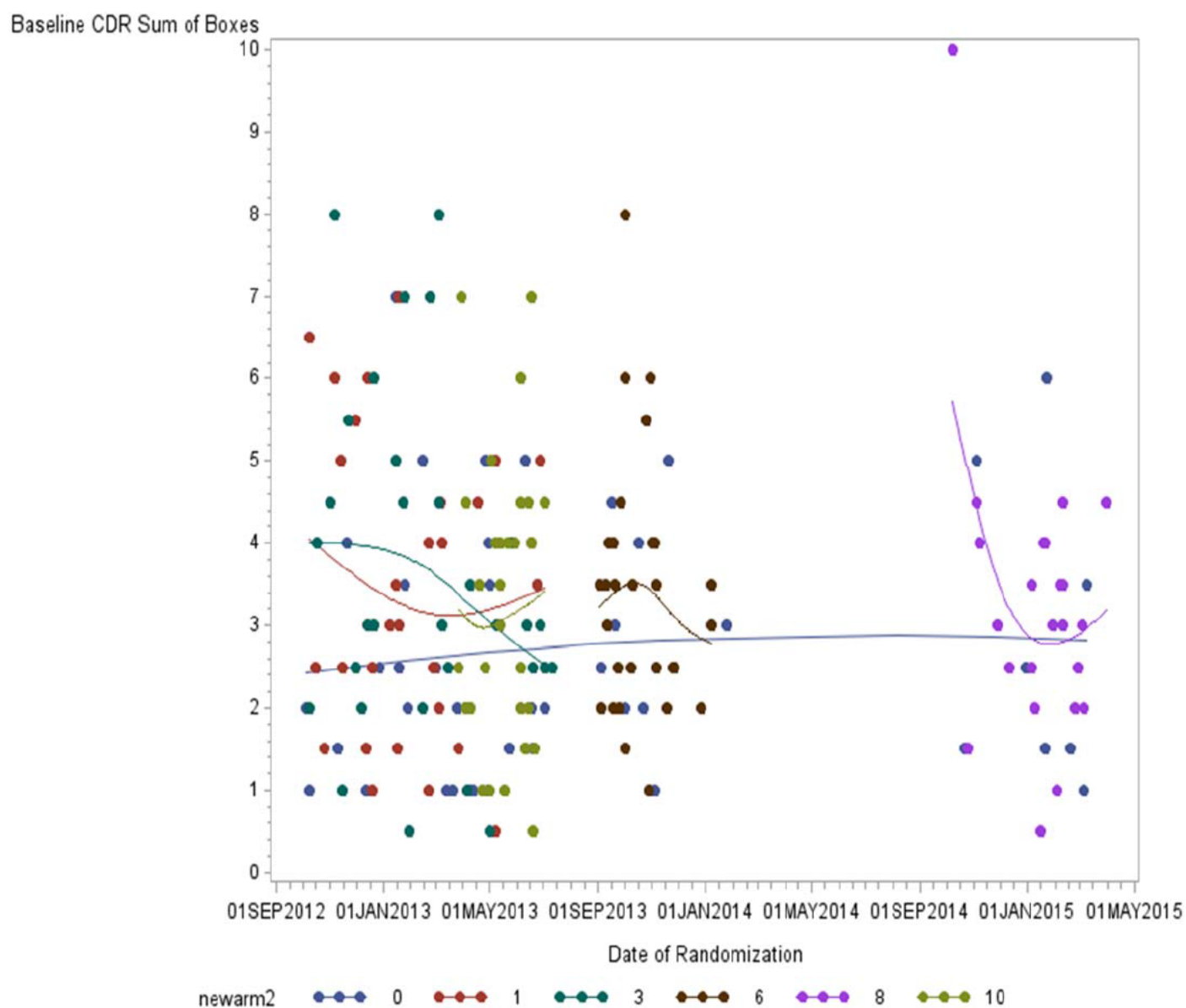
Without site 219 the estimated effect is .95 +/- .53, $p = .0789$.

Sites 229, 236, and 201 were the most influential on the 10 mg/kg results. Exploratory exclusion individually led to p -values $\geq .09$.

Figure 27 illustrates that all of the placebo was not randomized concurrently with the 10 mg/kg dose due to the sequential group randomization design. LOESS or local regression curves are drawn to aid the eye in the assimilation of the average CDRSB on a local [narrow window of time] level. Due to the staggering of the arms in time, the comparison of 10 mg/kg with all placebos may be biased in case of imbalances and, regardless, lacks the support of direct randomization between both complete groups. In fact, there are imbalances between 10 mg/kg and pooled placebo. For example, overall pooled placebo was 56% female but 10 mg/kg was only 41% female and overall placebo was 75.0% APOE carriers as compared to 65.6 % for 10 mg/kg. Also, some sites had no randomized patients from one group or the other, i.e., either placebo or 10 mg/kg. The overall comparison may be confounded with these differences. If the titration groups which add some extra carriers for placebo are excluded the 10 mg/kg vs all other placebo is not significant 1.05 +/- .57, $p = .0689$. And if only the concurrent placebo arm is used 10 mg/kg is not significant 1.12 +/- .71 $p = 0.122$.

Figure 27 illustrates the staggered arm design while showing the baseline CDRSB by arm. All placebo arms are grouped together. It is clear that there are placebos (dark blue) randomized both well before and well after the 10 mg/kg arm (olive). The number of the "newarm2" arm in the legend refers to the dose, except newarm2=8 refers to the titration to 10 mg/kg arm.

Figure 27 Baseline CDRSB by Arm showing Staggered Design



3.3 Evaluation of Safety

Incidence of Falling adverse events was 3.2 % higher in the high dose than placebo (Table 13 showing common adverse events).

Table 13 Adverse Events With at Least 5% Incidence in BIIB037 10 mg/kg and 2% Higher Incidence Than Placebo –Pool A1

	Placebo (N=1087) n (%)	BIIB037 3 mg/kg (N=760) n (%)	BIIB037 6 mg/kg (N=405) n (%)	BIIB037 10 mg/kg (N=1033) n (%)	BIIB037 total (N=2198) n (%)
Number of subjects with any event	945 (86.9)	700 (92.1)	347 (85.7)	946 (91.6)	1993 (90.7)
Amyloid related imaging abnormality-oedema/effusion	29 (2.7)	223 (29.3)	83 (20.5)	362 (35.0)	668 (30.4)
Headache	165 (15.2)	161 (21.2)	58 (14.3)	212 (20.5)	431 (19.6)
Amyloid related imaging abnormality-microhaemorrhages and hemosiderin deposits	71 (6.5)	141 (18.6)	50 (12.3)	197 (19.1)	388 (17.7)
Fall	128 (11.8)	105 (13.8)	50 (12.3)	155 (15.0)	310 (14.1)
Superficial siderosis of central nervous system	24 (2.2)	91 (12.0)	23 (5.7)	151 (14.6)	265 (12.1)
Diarrhoea	74 (6.8)	62 (8.2)	27 (6.7)	92 (8.9)	181 (8.2)

NOTE 1: A subject was counted only once within each preferred term (MedDRA version 22.0).

NOTE 2: Preferred terms are presented in decreasing frequency of the table's BIIB037 10 mg/kg column.

NOTE 3: Preferred terms are displayed if the incidence is at least 5% in the BIIB037 10 mg/kg and the incidence difference compared to placebo is at least 2%.

Note: Table copied from page 64 of Sponsor's Clinical overview document

This was also statistically significant when analyzed by time to first fall (HR= 1.33 [1.05, 1.68] .p=.0166) or number of falls adjusted by time at risk, risk ratio 1.30 (1.12,1.52)

When the CDRSB is transformed to a maximum of 1 as one may do to permit a quantitative risk benefit analysis, the transformed CDRSB difference at Wk 78 in 302 MMRM estimate is 0.02226 +/- 0.008644.

The recent International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) addendum R1 on statistical efficacy ICH E9 (R1) stressed the importance of accounting for potential biasing intercurrent events such as death in primary statistical analyses.

If a Joint rank of time to first fall and cognition/function is conducted as a benefit risk analysis high vs. placebo in OTC of Study 302 gives an estimated difference in ranks of: 36.6 +/- 28.81 p=0.2040. Note that this analysis method imposes a heavy penalty for falls, giving patients with them a rank worse than any observed Week 78 CDRSB in patients without any falls. Two patients who both had falls are ranked by the time to first fall and two patients without falls are ranked on the CDRSB changes at Week 78. The resulting (joint) rank sums are analyzed adjusting for the same covariates the sponsor specified in the primary analysis model except for those involving Visit because only the last common Visit between each pair of patients is used to determine the rank for the pair.

There were 6 high dose and 5 placebo deaths in the OTC or Died population (N=643 for placebo and high dose) of study 302. A joint rank analysis of CDRSB change and survival gave an estimated Week 78 difference in rank sums of 64.05 +/- 28.63, p=0.0256.

Please see the Clinical safety review for a review of general safety.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race, Age, and Geographic Region

4.1.1 Gender, Race, and Age

SEX

In Study 302, 51% and in Study 301 52% were Female. In study 302, both high and low dose differences from placebo are smaller for females and the SEX*VISIT interaction in 302 is marginally significant 2,1093 df F pr>F=0.0896 suggesting a potentially different CDRSB trend over time by Sex. Table 14 shows estimated treatment effects for the high dose by Sex and by Study as well as pooled studies. Pooling over studies (while allowing for a study effect) is often

recommended for subgroup estimation because subgroups, tend to be underpowered and have higher variability than the overall ITT population in individual studies. Females numerically favored placebo in Study 301 and in Study 302 had an estimated effect less than $\frac{1}{2}$ that of the estimated effect in Males. In the pooled analysis the high dose effect on CDRSB at Week 78 in Females was very close to 0.

Table 14 Differences on CDRSB by Sex Subgroup

Study	Group	Est. Diff.	Std. Error
301	Female h78	-0.2408	0.2092
301	Male h78	0.1976	0.2188
302	Female h78	0.2047	0.2153
301	Male h78	0.5862	0.2222
301/302 Pooled	Female h78	-0.01805	0.1500
301/302 Pooled	Male h78	0.3919	0.1559

AGE

The mean age in 302 is 70.7 (median 72) and 80% were ≥ 65 years of age. The SAP states that age subgroup categories would be Age < 65, 65-74 and >75. Proportions among all groups falling into these categories were 23, 47, and 30 in 301 and 20, 47, and 33 in study 302. Three factor interaction in 302 had a p-value of .0919 (8 num DF) overall age effects (18 num DF) the p-value is .1448.

When age is treated as continuous the p-value for the 3-way interaction between Age, Treatment Group, and Visit is .0058 (9 num DF). Testing only the three way interaction effects gives a p-value of .0051 (4 num DF). The implication of the interaction is that the estimated high dose effect was mainly in the highest age group.

Table 15 shows estimated treatment effects for the high dose by Age Group (<65,65-74,≥75) on CDRSB at Week 78 and by Study as well as pooled studies. The <65 group was in the wrong direction (placebo was numerically better) in 301 and in Study 302 had an estimated effect less than ¼ the size of the ≥ 75 group, as did the 65-74 age group.

Table 15 CDRSB Hi Placebo Differences on CDRSB at Week 78 by Age Group and Study

Study	Group	Est. Diff.	Std. Error
301	<65 h78	-0.3050	0.3039
301	65-74 h78	0.05618	0.2227
301	≥ 75 h78	0.07405	0.2793
302	<65 h78	0.1940	0.3409
302	65-74 h78	0.2088	0.2225
302	≥ 75 h78	0.8302	0.2756
301/302 Pooled	<65 h78	-0.05546	0.2284
301/302 Pooled	65-74 h78	0.1325	0.1574
301/302 Pooled	≥ 75 h78	0.4521	0.1962

When Age was treated as continuous instead of categorizing into 3 age ranges there were still interactions with treatment (AGE*AVISITN*TR01PG1N 9 1895 F=2.59 p=0.0058)

The estimated high dose effect at Week 78 as a function of age can be expressed by the equation $-2.2859*(P=0) + 0.03821*AGE*(P=0)$ note: this equals 0 at 59.8 0.19775 at 65 .38 at 69.8

Race

In Study 301 75.2% were white 10.6% were Asian, and 14.2 % were Other (including Black and Indian) and or not reported).

In Study 302, for Race 76% were white, 8% Asian, 16% were Other (including Black and Indian).

Race was actually not reported for 15.6 % in 302 and 13 % in 301.

Table 16 shows estimated high dose treatment effects at Week 78 by Race for individual studies, as well as for pooled studies.

There was some suggestion in study 302 of possible differential profiles and treatment effects by Race (Race Main effect p=0.0235; Race by Visit interaction p=0.0774 and Race*Visit*Treatment 3 way interaction p=0.1670). In Study 302 the estimated high dose treatment difference vs. Placebo on CDRSB at Week 78 in Asians was more than 2.5 times larger than the effect in Whites or Others.

Table 16 Pooled and By Study Analysis of Estimated High Dose Treatment Effects by Race Groups for CDRSB at Week 78

Study	Group	Est. Diff.	Std. Error
301	Asian h78	-0.06544	0.5140
301	White h78	0.1603	0.1720
301	Other h78	-1.0182	0.4005
302	Asian h78	1.0569	0.6827
302	White h78	0.3887	0.1741
302	Other h78	0.3000	0.3933
301/302 Pooled	Asian h78	0.4957	0.4273
301/302 Pooled	White h78	0.2745	0.1224
301/302 Pooled	Other h78	-0.3591	0.2807

4.1.2 Geographic Region

Enrollment by geographic region was not the same in studies 301 and 302 (Table 17). Several countries were only involved in one of the two studies (Australia, Austria, Great Britain, Denmark, Korea, Puerto Rico, Taiwan in 301; Belgium, Chechnya, Finland, Netherlands, Poland, Sweden in 302. The US proportion was slightly higher in 301: 46.4 vs. 39.6% in 302. There were 13 countries in study 302. The distribution of regional enrollment also changed over the course of the studies which may confound the impact of protocol amendment 4 (allowing the APOE+ high dose to reach 10 mg/kg instead of only 6 mg/kg). In 302 Japan increased from 1.7 to 11.9% and Germany increased from 2.8 to 11.1%, Poland decreased from 19.1 to 6% after PV4 and these countries were impactful on the high dose effect in the same direction as the enrollment change from before to after, thus making an assessment of the APOE+ dose increase from 6 mg/kg to 10 mg/kg in protocol amendment 4 confounded with Country enrollment changes from pre-PV4 to post-PV4.

Table 17 Enrollment by Country across Studies 301 and 302

Study Identifier		All
221AD301 221AD302		
Country		
AUS	N	99
	PctN	6.2
AUT	N	9
	PctN	0.6

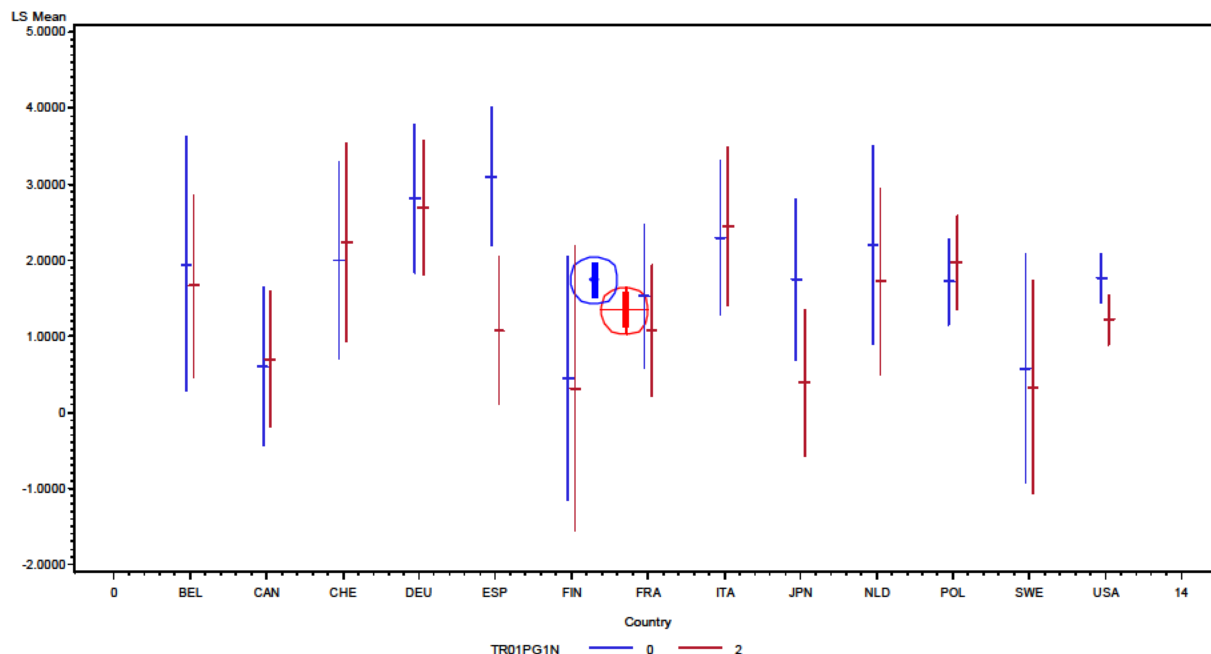
		Study Identifier		All
		221AD301	221AD302	
BEL	N		49	49
	PctN		3.1	1.5
CAN	N	78	95	173
	PctN	4.9	6.0	5.4
CHE	N		49	49
	PctN		3.1	1.5
DEU	N	96	110	206
	PctN	6.0	7.0	6.5
DNK	N	22		22
	PctN	1.4		0.7
ESP	N	103	77	180
	PctN	6.4	4.9	5.7
FIN	N		33	33
	PctN		2.1	1.0
FRA	N	59	76	135
	PctN	3.7	4.8	4.2
GBR	N	79		79
	PctN	4.9		2.5
ITA	N	113	80	193
	PctN	7.1	5.1	6.1
JPN	N	95	119	214
	PctN	5.9	7.5	6.7
KOR	N	47		47

		Study Identifier		All
		221AD301	221AD302	
NLD	PctN	2.9		1.5
	N		46	46
POL	PctN		2.9	1.4
	N		185	185
PRT	PctN		11.7	5.8
	N	46		46
SWE	PctN	2.9		1.4
	N		35	35
TWN	PctN		2.2	1.1
	N	13		13
USA	PctN	0.8		0.4
	N	743	626	1369
All	PctN	46.4	39.6	43.0
	N	1602	1580	3182

In study 302, three regions were specified for the Region effect in the primary analysis model: Europe/Canada, Asia, and the United States. There was a lot of variability within the Europe/Canada region. In four of the eleven Europe/Canada region countries, LS Mean Differences numerically favored Placebo compared to the High Dose. A test of whether including all of the countries fit the data better than just including the three regions was nominally significant, indicating the including the separate countries fit the data better (Country vs Region1 LR test ChiSq=32, on 10 and 1598 DF Europe/Canada/Australia Region 2 contrast 10DF has p=0.0004 for equality of constituent countries).

Figure 28 shows Placebo and High Dose LS means at Week 78 by Country. The big circles in the middle illustrate the overall LS means. The Europe/Canada/Australia region used in the analysis shows considerable variability (all bars except JPN 5th from the right and US 1st on the right).

Figure 28 Study 302 Placebo and High Dose LS means for CDRSB by Country



There was a significant country main effect ($p < 0.0001$) on the Change from baseline in CDRSB when country was added to the primary model to check for consistency across countries. A likelihood ratio test of the primary analysis model augmented with country effects, country by visit, country by treatment, and country by treatment by visit effects versus the primary model which did not adjust for any of these yielded a Chi square p-value of 0.0122, suggesting that there was statistically significant variation (lack of consistency) in the treatment effect across countries (Country Main effect $p < 0.0001$; COUNTRY*TR01PG1N $p = 0.7979$; COUNTRY*AVISIT $p = 0.0017$; COUNTRY*AVISIT*TR01PG $p = 0.0080$).

An F test of just the last 3 terms representing different profiles by Country and Treatment Group (96 num DF) has a p-value of 0.0015. (ADASCog .2439; ADCSADL .0774); Country*tr01pg1n and country*avisit*tr01pg terms only given country and country*avisit terms has a p-value of 0.0543.

An F test of the other parts of the COUNTRY*VISIT*TRT interaction than the COUNTRY*VISIT part but excluding VISIT*TRT which is already in the primary analysis model has a p-value of .0459. A likelihood ratio test of the same has a p-value of 0.0311. Therefore, to say that the three way interaction is driven by the COUNTRY*VISIT part is an oversimplification.

An exploratory analysis excluding the US (which accounts for 40% of the overall population), was not nominally significant (Non-US 302 N=954 Nobs=2159 HiPl Wk78 diff=.297 SE=.218 p=.1739).

For 301 the corresponding tests are as follows.

COUNTRY	13	1472	4.35	<.0001
COUNTRY*TR01PG1N	26	1477	1.19	0.2335
COUNTRY*AVISITN	26	1900	1.51	0.0467
COUNTR*AVISIT*TR01PG	51	2003	1.09	0.3020

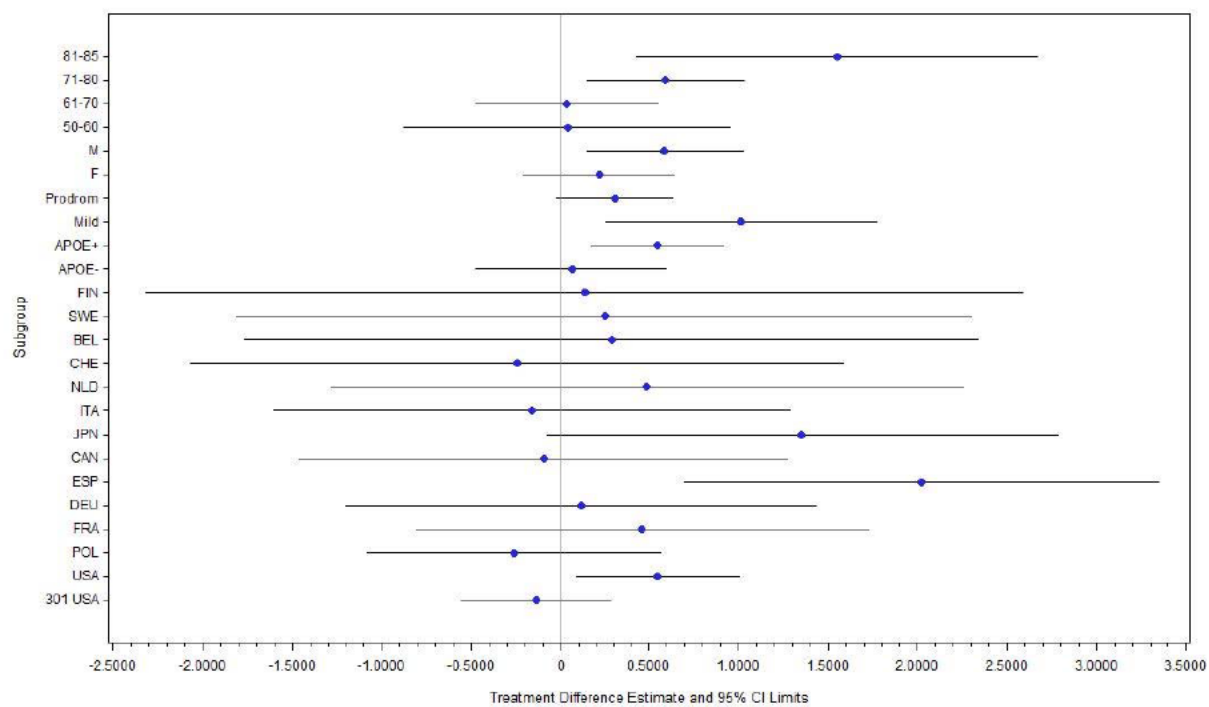
An F test of just the last 3 terms (103 num DF) has a p-value of 0.0678.(ADASCog .1000 ADCS .2178)

For Placebo alone country by Visit interaction has an type 3 F test p-value of 0.0059. For the high dose the corresponding test has a p-value of 0.2090 and for the low dose 0.0022.

A forest plot of treatment effects on CDRSB at Week 78 by Country follows. Lower CDRSB scores are better and the difference is presented as Placebo – Adu 10 mg/kg, so that positive differences favor Aducanumab.

The following forest plot shows inconsistencies of high dose treatment effect at Week 78 in study 302 for important subgroups such as Country, Mild AD vs. Prodromal AD (less effect in Prodromal [interaction p=0.09]), Age Group (more effect in older [interaction p=0.06]), APOE (more effect in + [interaction p=0.15]). These inconsistencies might be important if one was evaluating whether study 302 could stand on it's own, which however introduces selection bias given the 301 result.

Figure 29 Forest Plot of Change in CDRSB at Week 78 by Country and Other Subgroups



There were 7 countries included in both 301 and 302. Altogether they accounted for about 78% of the patients in both studies (80% in 301 and 75% in 302). Therefore, 25% of 302 patients were from countries not involved in 301. The pooled results by country, after allowing for study differences, for the high dose difference in CDRSB at Week 78 for the countries involved in both studies are as follows Table 18. In this table positive differences favor the high dose. As can be seen in the forest plot of Figure 29 the high dose was in the wrong direction in the US subgroup of Study 301. The high dose effect at Week 78 in Japan, the US, Canada, and Italy among others had different signs (favoring drug or favoring placebo numerically) between 301 and 302.

Table 18 CDRSB Hi vs. Placebo Results at Week 78 Averaged over 301 and 302 by Country

Country Sample Size at Week 26 and 78	Country/Visit	Estimated Diff	Std. Error
COUNTR*AVISIT*TR01PG 171/85	CAN 78	0.6279	0.4911
COUNTR*AVISIT*TR01PG 200/87	DEU 78	-0.5457	0.4648
COUNTR*AVISIT*TR01PG 177/106	ESP 78	1.0771	0.4410
COUNTR*AVISIT*TR01PG 133/89	FRA 78	-0.1951	0.4926
COUNTR*AVISIT*TR01PG 194/94	ITA 78	0.3288	0.4720
COUNTR*AVISIT*TR01PG 214/60	JPN 78	0.5403	0.5221
COUNTR*AVISIT*TR01PG N26=1348/N78=898	USA 78	0.1890	0.1569

4.1.2.1 Individual Sites

There were 142 sites among the 13 countries in study 302. Note that the randomization was stratified by site and APOE carrier status.

Figure 30 shows treatment effect at Week 78 on CDR-SB by Site. A few of the most influential sites were 872 (-8 pts 22 recs), 856, 849, and 664:

Figure 30 Study 302 CDRSB Change at Week 78 by Site

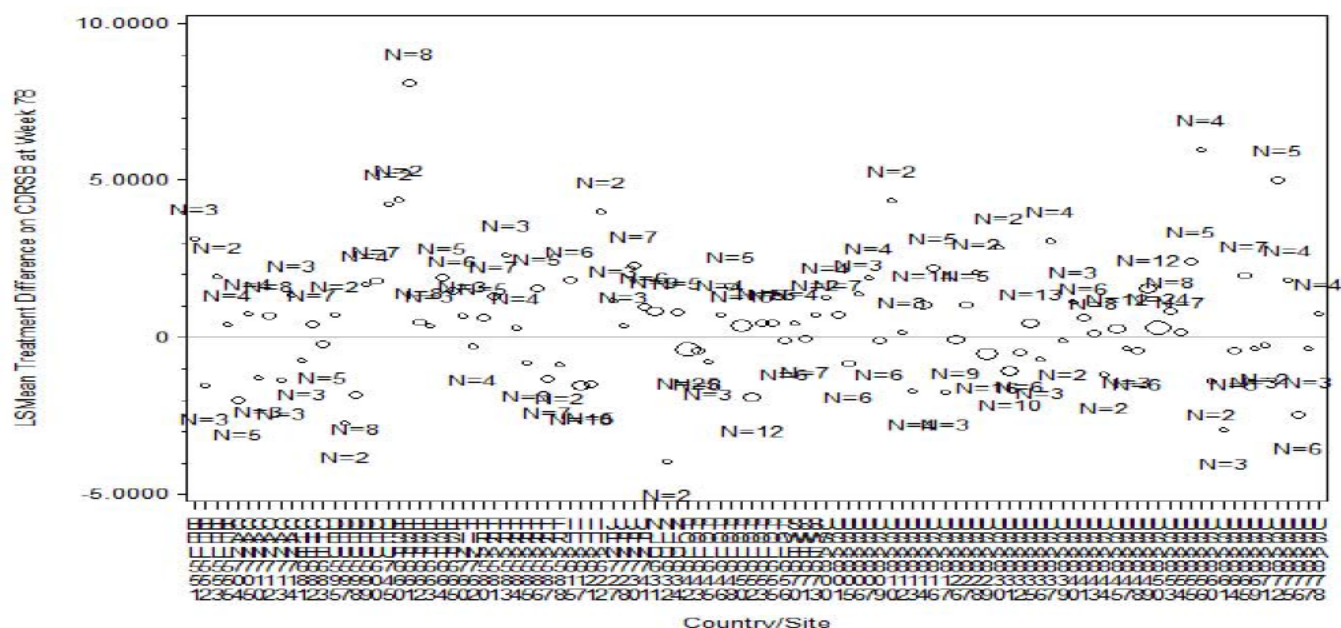


Table 19 shows the effect of exploratory exclusion of sites on the primary result.

Table 19 Study 302 Hi vs. Placebo LS Mean differences on CDRSB at Week 78 after excluding a site of interest

Site excluded	#Patients	#Records	Estimate	stderr	p
None	1580	3712	-0.4007	0.1556	0.0101
- 872 (USA)	1572	3690	-0.3593	0.1556	0.0211
- 856 (USA)	1572	3690	-0.3635	0.1553	0.0194
- 849 (USA)	1553	3647	-0.3785	0.1579	0.0167
- 664 (ESP)	1565	3680	-0.3842	0.1568	0.0144

If Site is added to the primary analysis model in place of Region the Site effect is highly significant and a simultaneous F test of all Site effects and interactions up to and including the

three way interaction with Visit and Treatment have $p < 0.0001$ (excluding the low dose data due to convergence issues related to small sites).

site	138	731	2.04	<.0001
tgrp1*site	125	719	1.12	0.1898
AVISITN*site	237	797	1.23	0.0201
tgrp1*AVISITN*site	223	789	1.13	0.1241

These Site interactions seem to further support the interactions discussed previously involving Country and suggesting inconsistency of the treatment effect by Site (and Country).

4.2 Other Special/Subgroup Populations

The dose was weight adjusted and yet, still baseline weight was found to be a nominally significant predictor of change from baseline in CDRSB ($p=0.0288$) which also varied significantly by Visit (WEIGHTBL*VISIT, $p=0.0204$). The sign of the estimated effect suggests that higher weights tended to have better CDRSB scores. For example, those above the median weight of 71 had an estimated high dose effect 0.5512 ± 0.2240 (S.E.), whereas those below had an estimated high dose effect of 0.2503 ± 0.2151 (S.E.).

Baseline disease stage (Prodromal or MCI) was adjusted for in the primary analysis model and, therefore, these baseline disease stage subgroups may be of importance and interest. There was a bigger effect in the smaller Mild subpopulation (18% of all 302 patients) in 302 (Figure 31). The difference between these subgroups for the high dose treatment effect at Week 78 would be significant at the .10 level (mild-v mci hi 78 0.7101 +/- 0.4219, $p=0.0926$), which may not be unreasonable for an underpowered interaction test. The baseline disease stage main effect and various interactions with it were somewhat compelling given the lower power for subgroups (ADBL

	1	1544	43.87	<.0001
ADBL*TR01PG1N	2	1522	2.28	0.1029
ADBL*AVISITN	2	1134	14.87	<.0001
ADBL*AVISITN*TR01PG1	4	1337	1.75	0.1358

. The difference was more compelling when High or Low differences were considered $p=0.0285$.

Figure 31 Study 302 Estimated High vs Placebo Differences at Week 78 on CDRSB by Baseline Disease Stage

Baseline Disease Stage	Comparator to Placebo and Timepoint	Estimated Difference from Placebo	Std. Error
Mci	high 78	0.3025	0.1672
Mild	high 78	1.0127	0.3872

Table 20 shows the results by baseline disease stage for each individual study and Pooled (interacting group trt and visit with study).

Table 20 Study 302 Estimated CDRSB Change Differences from Placebo at Week 78 by Baseline Stage Diagnosis

Study	Group	Est. Diff.	Std. Error
301	Mild h78	-0.2774	0.3592
301	MCI h78	0.02286	0.1649
302	Mild h78	1.0053	0.3885
302	MCI h78	0.2972	0.1670
301/302 pooled	Mild h78	0.3640	0.2645
301/302 pooled	MCI h78	0.1600	0.1173

APOE

The randomization was stratified by APOE and 67% were carriers (+). The treatment difference on CDRSB for the high dose in the APOE- (non-carrier) subgroup, by which the randomization was stratified, was very consistent across studies 301 and 302, very small and not nominally significant (Table 21). As discussed above in 3.2.1.4.2 the smaller estimated effect for APOE non-carriers than APOE carriers on CDRSB in Study 302 at Week 78 was also observed for all of the other key secondary endpoints (including a significant interaction for MMSE : APOE*TRT p=0.0065). APOE - vs+ hi diff p=.1511 at Week 78 for CDRSB, p=.07 for both MMSE and ADCS-ADL, also favoring carriers.

Table 21 CDRSB at Week 78 for High vs Placebo by APOE and Study and Pooled

Study	Group	Est. Diff.	Std. Error
301	apoegr3n - h78	0.06518	0.2732
301	apoegr3n + h78	-0.07219	0.1816
302	apoegr3n - h78	0.06720	0.2705
302	apoegr3n + h78	0.5357	0.1884
301/302 pooled	apoegr3n - h78	0.06619	0.1923
301/302 pooled	apoegr3n + h78	0.2317	0.1309

Concomitant Use of AD Medications at Baseline

Concomitant use of AD medications at baseline was also a model adjustment factor and thus defines subgroups of interest. Just over half used concomitant AD medications at baseline (51%). High dose vs. placebo Treatment group differences were relatively consistent across those using AD medications at baseline and those not using in study 302 (No 0.3971 +/-0.2255 , Yes 0.3979 +/- 0.2150) . However, the CDRSB profile across visits appeared to vary between them as indicated by an interaction test (ADCMBLFL*AVISITN 2 1095 7.79 p=0.0004).

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

Study 302 didn't complete according to the protocol because it was terminated for futility (<20% chance of success in either 302 or 301 study for either dose) based on the pooled analysis of it and study 301 when both were 50% complete. According to closed executive session minutes, because the analysis showed futility, the DMC was provided with the efficacy and safety data from both the ENGAGE and EMERGE studies [221AD301 and 221AD302]. Furthermore, the DMC discussed various aspects of the data and the analyses, agreeing that even with careful review of the data and analyses provided, that there was no evidence of movement that would have resulted in a non-futility conclusion. In the final open DMC session the DMC informed the Biogen attendees that there were no additional analyses, beyond the pre-specified analyses, requested to reach their recommendation regarding the trials. The DMC stated that it was their unanimous recommendation that the trials be halted. The DMC asked if the Biogen attendees would like to review with the DMC the prespecified outputs for the futility analysis. Biogen agreed, and the analyses were reviewed in a joint session between Biogen attendees and DMC members.

There was considerable unblinding to manage ARIA events during the course of the study (ARIA events :35% high dose [45% in APOE+] 26% and 2% for low dose and placebo in study 302, page 177 of June 14 2019 Type C briefing package). The absence of a significant impact of data collected after ARIA events does not necessarily imply that there was no bias due to unblinding due to ARIA; there is limited power for detecting such a difference. Moreover, one can't conclusively rule out an impact of those experiencing ARIA on the result because it requires making a comparison based on differential exclusions between the randomized groups (exclusion of drug patients and/or censoring of drug arm data) and the resultant groups without ARIA to be compared are no longer as randomized and/or have differential follow-up and selection bias due to conditioning on a post-randomization event.

There was a major protocol amendment in the middle of the study modifying dosing for the APOE+ stratum high dose. Changing enrollment by country over time and differential efficacy by country and variations in placebo response over time make it virtually impossible to elicit the effect of increasing the dose in protocol amendment 4. If the sponsor had not rushed from phase

1B to phase 3 phase 3 could have started the APOE+ group at 10 mg/kg and we would have a cleaner study and a clearer picture of whether or not more 10 mg/kg doses is more important.

We should keep in mind that prior to amendment 4 the high dose is not the same for APOE+ and APOE-, 6 mg/kg and 10 mg/kg respectively, and also the moderately common occurrence of ARIA limited the dosing (35% of high dose patients had dose modifications). This would seem to require drawing back from a blanket missing at random assumption for imputation of missing data for the high dose and call into question an imputation model not accounting for these differences. In 302 the occurrence of ARIA in the high dose was 1.4 times higher for APOE+ vs. APOE- prior to amendment 4 and 2.3 times higher post PV4. Limitation of dose titration in the high dose was 2.1 times higher for APOE+ prior to PV4 and 3.7 times higher post PV4. Thus, unblinding for dose managing would have been higher after PV4. The APOE- stratum had more 10 mg/kg doses but was worse on average than APOE+ in 4 out of the 4 primary and key secondary endpoints (and the low dose shows the same pattern).

Regional differences may offer an alternative explanation for the 301 and 302 differences (less US in 302: 39.6% vs. 46.4% in 301; Poland 11% in 302, 0% in 301; AUS 6% in 301, 0% in 302; several other countries only in one study). Please see Table 15 for enrollment by study details.

The effect of the protocol version 4 change of dose for APOE+ is confounded with other population enrollment changes:

8-13[H]% high dose drop in US region,

drop in Poland for Placebo from 19.7% pre-PV4 to 5.6% post-PV4 *Pol negative for high dose

Increase in JPN: Placebo 1.6% pre-PV4 to 12.9% post-PV4; *JPN big effect may drive result

DEU Placebo 3.3% pre-PV4 to Placebo 10.2% post-PV4

AgeGroup cat3 (71-80) P pre-PV4 40.2% post-PV4 31%

AgeGroupcat4 (>80) P pre-PV4 41.8% post-PV4 50%

These enrollment changes are confounded with pre-PV4 to post-PV4 dose changes and if responses differ by countries (as the data suggests) it would be almost impossible to balance dosing subgroups across countries using propensity scores. Poland which is in the wrong direction for the high dose effect in 302 at Week 78 had higher enrollment pre-PV4 as compared to post-PV4 and Japan which had a big positive effect for the high dose had enrollment increased post-PV4 as compared to pre-PV4.

There was also slightly more post randomization starting of AD symptomatic medications in 302 than 301 (9.0 P 10.4L 7.9H in 301 [6.9P 10.7L and 7.8H post PV4] vs. 302 10.4P 9.8 L 12.3H [10.2 P 9.2L 13.2H post-PV4]).

Consider the following region enrollment changes in study 302.

7-12% drops in US region from pre to post-PV4 across groups (Placebo: 43.6%→36.8% ; Low 45.0 →35.0; High: 46.0→33.6)

Decrease in Poland Placebo enrollment (19.7% pre-PV4 vs. 5.6% post-PV4) *Poland High dose effect is in wrong direction (-0.26 +/- .42 SE)

Increase in Japan Placebo enrollment (1.6% pre-PV4 vs. 12.9% post-PV4); *JPN may drive result -biggest country effect favoring High dose (+1.35 +/- .73 SE)

Increase in DEU Placebo 3.3% pre-PV4 vs. 10.2% post-PV4

There were also differences in CDRSB response by Age Group and changes in Age group distribution by pre-PV4 and post-PV4:

Agegr1n (Age group) Age 61-70 in Placebo group 40.2% Pre →31.0% Post; Age 71-80: Pre-PV4 41.8% →post-PV4 49.8%

Note that 302 has significant interactions with treatment*visit*Agegr1n and treatment*visit*Country, as well as visit*agegr1n and visit*country. Study 301 also has a significant country*visit interaction. These, among other interactions such as baseline disease stage group found by the reviewer, suggest that the model is not likely correct. Since demographic and disease characteristics changed somewhat after PV4 those missing Week 78 have different characteristics and if the model is not correct the model may be biased given the large amount of missing data for the ITT population. Different demographics of those without the opportunity complete such as more from the Asian region and DEU (Germany), more mild baseline disease stage, more symptomatic AD medications at baseline, and increased age may cause the model to not be valid under MAR since the model does not account for interactions with these variables and Visit but the data suggests they are significant (should be included). With more than 40% missing data and the uncertainty of the primary model's validity under MAR (it not including interactions that the data suggest are significant) it seems that the Opportunity to Complete Population result is more reliable and more relevant (also considering that Week 78 is the only significant timepoint in Study 302).

More than 40% missing Week 78, the only timepoint that showed nominal significance on the primary endpoint (45% for high dose and 48% for placebo, June 14 2019 Type C briefing package). Neither the primary nor any of the key secondary endpoints was significantly different from placebo at Week 50 in study 302. With only a single positive timepoint, no other positive timepoint to confirm it, it is not established that the progression was slowed. There are no compelling correlations within the high dose group between change in A β SUVR in the cerebellum, the primary biomarker, at Week 78 and Change in CDRSB at Week 78. In fact, in study 302 this unadjusted correlation was in the wrong direction. Furthermore, unlike the Solanezumab program, for Aducanumab there is no delayed start design to potentially support a slowing of progression. Even if there was it would likely be problematic due to the early futility stopping since there would be associated dropouts and disruptions.

There was a blinded sample size increase [Updated sample size (from 450 to 535 participants per treatment group) based on the BSSR conducted in November 2017] prior to the unblinded interim futility analysis but the conditional power for study 302 high dose was still only 58.63% at the interim for CDRSB based on the study 302 interim effect size only (N=803 the effect size was -.2818, SE=.1902, p=0.139). If the final 302 high dose effect on CDRSB was real, the chance of 301 succeeding given it's sample size would be 0.755 {the SE would be $\sqrt{3.547097*((333+293)/333/293)} = 0.1508578$, so the chance of success would be $1 - \text{pnorm}(1.96 - .4/.1508578) = .7553754$ }. The chance of observing a result as unimpressive as observed for 301 would be $\text{pnorm}(-.03/.1508578 - .4/.1508578) = 0.002183445$. The estimated difference in the patients not included in the interim N=835 is -0.5043 +/- .3279 (SE) for high dose and -.5654 +/- .3361 (SE) for the low dose (the low dose is numerically better after the interim!). This seems to conflict with the sponsor's assertion that protocol version 4's allowing the high dose to increase from 6 mg/kg to 10 mg/kg in APOE carriers optimized the dose response. Placebo showed a marked worsening after the interim 1.77 +/- .241 pbo; 1.26 +/- .228 hi; 1.20 +/- .241 lo) vs. (interim: 1.53 +/- .164 pbo; 1.25 +/- .160 hi; and 1.43 +/- .164 lo). The sponsor's explanation requires considering a post-randomization event defined subgroup which has no proper non-counterfactual control, ability to tolerate drug without ARIA, for statistical justification whereas the higher placebo response explanation for the better effect after the interim requires no breaking of the randomization.

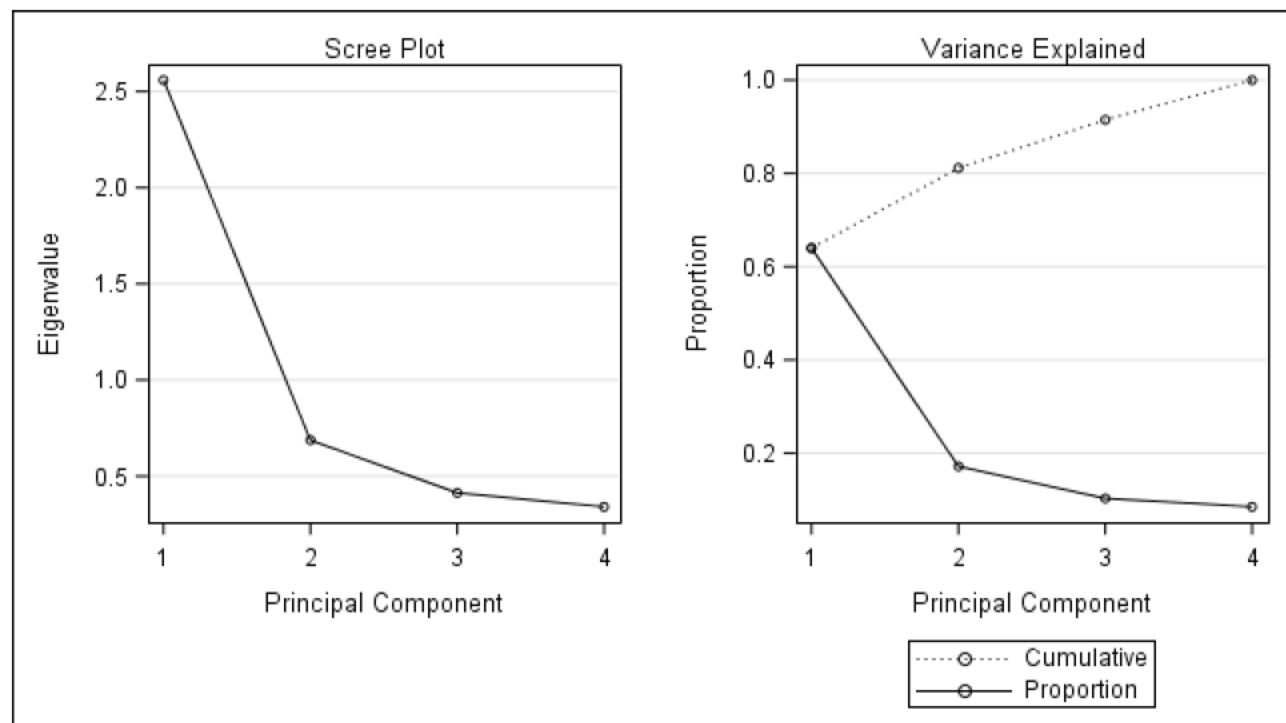
The correlations among the primary and key secondary endpoints at Week 78 are substantial and all easily nominally significant as shown below (Table 22). To suggest as the sponsor did that they measure relatively independent things seems highly questionable due to this and, e.g., all of the endpoints were positive in 302 and they all were negative in 301. In study 301 the correlations among the primary and key secondary endpoints are very similar to those in study 302 all > .40 in absolute value and highly significantly different from 0.

Table 22 Study 302 Correlations at Week 78 between changes from baseline on Primary and Key secondary endpoints

Prob > r under H0: Rho=0 Number of Observations				
	CDRSB	MMSE	ADAS-cog	ADCS-ADL-MCI
CDRSB	1.00000 876	-0.55312 <.0001 874	0.49447 <.0001 864	-0.64083 <.0001 859
MMSE	-0.55312 <.0001 874	1.00000 879	-0.58233 <.0001 868	0.44297 <.0001 862
ADAS-cog	0.49447 <.0001 864	-0.58233 <.0001 868	1.00000 869	-0.39773 <.0001 855
ADCS-ADL-MCI	-0.64083 <.0001 859	0.44297 <.0001 862	-0.39773 <.0001 855	1.00000 862

The scree plot in Figure 32 can be used to judge the number of important factors among the four key endpoints. Because it levels off very quickly after the first eigenvalue, it suggests that, rather than four independent factors, one or at most two are needed to explain the variation among the four key endpoints. Therefore, again, the four key endpoints do not measure very distinct efficacy information, i.e., one or at most 2 captures the key information.

Figure 32 Scree Plot of Principal Components of Primary and Key Secondary Endpoints



Why don't high dose non-carriers show a clinical effect in Study 302 if they got 10 mg/kg earlier, have less ARIA dose reductions and if they showed a bigger effect on cerebellum SUVR AB uptake? This seems to call into question whether AB PET SUVR in cerebellum is a surrogate. In fact, within the high dose group, there is actually no correlation between Week 78 change in cerebellum SUVR AB and Week 78 change in CDRSB. This seems to call into question a disease slowing claim.

Baseline disease severity group also showed this disconnect in the relationship between SUVR and CDRSB change, i.e., more on the one is less on the other.

APOE4+ effect is consistently larger in carriers on primary and key secondary endpoints in Study 302 which represents a disconnect with both the SUVR Cerebellum Amyloid Beta change data and study 103 trends. The high dose is in the wrong direction for non-carriers on the first key secondary MMSE in Study 302.

The current efficacy standard in AD is placebo controlled comparisons from clinical trials. An argument based on exposure response is circular. It presupposes that the drug is effective and/or

that the SUVR is correlated with clinical cognitive and functional change, neither of which has been consistently shown across the phase 3 clinical data. Patients who were most compliant or tolerated the drug best and achieved the highest exposure were not randomized to that outcome and so there is no corresponding control group that they can be compared to without possible confounding due to selection bias.

Rapid progressors are part of the reality of Alzheimer's and after the fact it is too late to address them in a completed randomized study. A highly effective drug would not be likely to fail because of rapid progressors. Study 302 could just as well be the outlier relative to the true proportion of outliers in the natural progression. In fact, the range of CDRSB changes in Study 301 at 18 months appears consistent with the ADNI study ([adnimerge_May15.2014 data](#)). There are slightly more outliers in the high dose in 301 but that is worrisome in itself since they are consistent with the ADNI data and so should again raise doubts about the representativeness of the 302 result. Furthermore, robust regression, techniques (M estimation, least trimmed squares, MM estimation, S estimation) designed to be resistant to and downweight outliers, applied to the 301 Week 78 data still suggested no effect of the high dose compared to placebo and that it was numerically worse than the low dose (vs. Pl +0.0024 +/- 0.1159 p=0.9836; vs. Low +0.0322 +/- 0.1154 p=0.7802). Without the worst change of +13 in the high dose group the 301 high dose vs. placebo result is +.0267 +/- .1495 p=0.8581 as compared to +.0316 +/- .1499 p=0.8330 including it. This shows that Study 301 is a big study and one outlier patient has limited influence. Totally excluding the patient instead of just the Week 78 observation the result is +.0072 +/- .1487 p=0.9615, still in the wrong direction. More than one outlier in the high dose is more of a systemic problem and should be more worrisome and harder to discount.

Study 103 is an outlier among the three available studies. Study 103 had a much larger effect (300%) by Week 52 than 302 had at Week 78 and 302 showed nothing significant at Week 52

(302 52 -0.1011 SE 0.1128 p=0.3702 CI=[-0.3224 0.1202];

301 52 +0.07479 0.1048 p=0.4757 CI=[-0.1309 0.2804]) despite the much larger sample size. (p 161 of June Type C Briefing Package)

More of the effect was in the APOE- subgroup in study 103 which is the opposite of what was seen in study 302. There was no significant interaction between treatment and APOE though(interaction test for hi dose vs. placebo at Week 54 by APOE, p=0.3746). There was a very small difference in the APOE- subgroup in study 302, .06 vs placebo on CDRSB at Week 78, and it was almost identical to the high dose effect in APOE- seen in study 301. Thus, pooled across 301 and 302 there is very little evidence of a high dose effect at Week 78 in the APOE- subgroup, although this group had 10 mg/kg dosing from the start of the study and less ARIA so fewer dose interruptions.

Study 103 had a staggered multi dose design, so the 10mg/kg group is not completely concurrent with the full placebo group and there could be a resulting bias. The analysis was acknowledged to be exploratory and the sponsor's result for CDRSB at Week 54 is sensitive to the handling of

starting of post randomization AD medications in 103. The 10 mg/kg is not significant if data after starting post randomization AD medications are censored ($p=0.09$). This study doesn't seem very supportive due to sensitivity to handling of starting of AD medications, the staggered design without direct support of a dedicated randomization between 10 mg/kg and pooled placebo, interim analyses with some Biogen personnel unblinded, no effect in phase 3 at the Visit (Week 50) closest to the 103 study duration time (Week 54), and APOE- had a numerically bigger effect in study 103 than APOE+ which is the opposite of study 302 APOE subgroup results.

A p -value $< .05$ doesn't necessarily reflect a clinically meaningful effect especially if there would be connotations of disease modification but the actual evidence supporting that is lacking and there is a failed study calling that p -value into question.

This final analysis is like a second interim analysis at 55% information since 45% of Week 78 is missing. A 3 stage O'Brien Fleming boundary with interim analyses at 45, 55, and 100% would spend .010 by the 2nd stage. The critical value for stage 2 would be $Z=2.68$. The observed t statistic is 2.51.

Reproducibility of 302 is in question since 301 did not replicate it (e.g., "reproducible research" issue with p -values).

5.2 Collective Evidence

The draft guidance on substantial evidence states that "poor execution can render a trial of any design to be not adequate or not well-controlled and, therefore, unable to provide substantial evidence of effectiveness. Examples of this include a randomized, double-blind, placebo-controlled trial in which unblinding is common due to an effect of the test drug, and where a modest treatment effect is found on a primary endpoint that is subject to bias when drug assignment is known (e.g., a physician global impression). In these cases, the trials might not be considered adequate and well-controlled." Both of these conditions are a concern in this application. Further it states "Findings from other trials that are not consistent with the findings of the single positive trial would need to be considered collectively, and could weaken the overall strength of evidence."

Conditional power based on all ITT data with censoring after March 21, 2019 of meta analysis being positive if non-OTC patients by the time of futility were somehow able to complete is estimated to be .61 as follows. This kind of pooled analysis was the basis for the interim treatment estimate on which the conditional power calculation was based. Perhaps, this calculation does not adequately take into account heterogeneity between studies, but it summarizes all available phase 3 evidence and follows the spirit of the sponsor's original plan to use a pooled treatment effect estimate at the interim.

Based on a Bayesian meta analysis of study 301 and 302 the posterior probability of the alternative hypothesis is 0.62 and a corresponding 0.38 probability for the null hypothesis.

5.3 Conclusions and Recommendations

The totality of the data does not seem to provide sufficient evidence to support the efficacy of the high dose. There is much inconsistency and no replication. There is only one positive study at best and a second study which directly conflicts with the positive study. Both studies were not fully completed as they were terminated early for futility and had sporadic unblinding for dose management of ARIA cases which was much higher in the drug group. The Amyloid PET substudy data suggested a larger effect in APOE- which is the opposite of what was observed for the clinical outcome data in phase 3 and also the phase 1B study 103. Within the high dose group there is no significant correlation between Week 78 cerebellum SUVR change and Week 78 CDRSB change in study 301 or 302. Therefore, there is no convincing evidence of delaying clinical progression cognitive or functional: only a single positive timepoint (unreplicated and conflicted by a second study) and no delayed start design (termination for futility does not help with completeness or interpretability of long term follow up). The increased placebo progression post PV4 and smaller effect (on all 4 key endpoints) in APOE non-carriers who all got 10 mg/kg from the start of the study rather than like APOE+ having to wait until PV4, call into question the sponsor's assertion about intermediate dosing early (less 10 mg/kg doses) being a challenge. In addition, the low dose in study 301 was numerically better than the high dose despite having no 10 mg/kg doses and this comparison is supported by randomization. For these reasons, a study fully completed according to protocol without major non-prespecified amendments while the study is ongoing is needed to confirm or deny the positive study or the negative phase 3 study.

I. Appendix

I.1 Additional Issues

The reviewer found many significant interactions of baseline demographic and disease characteristic variables with VISIT (APOE, ADBL, ADCMBLFL, WEIGHTBL, ADACOGBL), suggesting that the impact of these on CDRSB changed significantly across VISITs. The WAVE 1 (simulated trial completion) analysis model did not take this into account. The WAVE 1 study completion imputation model was close to the primary model in terms of covariates, but technically post-hoc, and this may have affected explanatory variable selection which was not exactly the same as for the primary analysis model (addition of Age, Baseline stage, Sex not included in the primary model). There were many simulations of trial completion which

appeared consistent with the ITT results of both trials. If the imputation model is not correct though, then the results could be biased. A trial completed through imputation will always be more uncertain than a truly completed trial and this completion required imputations of Week 78 for more than 40% of the patients. With such a high rate of missing data the chosen model could be biased if important changes in covariate effects across Visit are not accounted for.

DRAFT