

**CENTER FOR DRUG EVALUATION AND
RESEARCH**

APPLICATION NUMBER:

761248Orig1s000

STATISTICAL REVIEW(S)



Addendum

Date: April 26, 2024

To: Teresa Buracchio, MD, Director Office of Neurology Products
Sally Jo Yasuda, PharmD, Deputy Director for Safety, Division of Neurology 1 (DN1)
Natalie Branagan, MD, Clinical Reviewer (DN1)

From: Ping Li, MS, Statistical Analyst, Division of Analytics and Informatics (DAI)
Fengyu Zhao, PhD, Primary Statistical Reviewer, Division of Biometrics VII (DBVII)
Yueqin Zhao, PhD, Team Lead, DBVII
Mat Soukup, PhD, Deputy Director, DBVII

Application: BLA 761248/SDN 100, 161, 165 (resubmission)

Drug Name: Kisunla (donanemab)

Subject: Mortality Assessment

1 Introduction

From the clinical safety review, there was an observed imbalance in mortality from Phase 3 Trial AACI in BLA 761248. Given the complexities in the design and conduct of Trial AACI, DBVII was consulted to provide further analytical assessment of mortality in the donanemab development program. A statistical memorandum was signed into DARRTS on March 4, 2024. Subsequent to the memorandum, an information request was sent to the Sponsor to provide an updated analysis of mortality and obtain updated data on mortality and study discontinuation. This addendum is to update the statistical memorandum based on the responses and data submitted (SDN165, SDN161) to this information request.

Also of note is that the sponsor submitted additional vital status information (SDN 168, 169) for the cohorts who were loss to follow-up to respond to an informal request by the clinical team on February 27, 2024. Data from these participants is not included in the comparative analyses, but summarized in Section 4.3.

The tables and figures in this document are numbered to match with the original memorandum. For example, Table 1A is to update Table 1 in the original memorandum.

2 Approach to Analysis

The methods for analyzing mortality rely on an on-study approach to analysis that includes all deaths that occur by week 76 (or week 76 + 57 days) regardless of discontinuing randomized treatment – see original memorandum for detailed descriptions of the analysis approach. This addendum includes an additional sensitivity analysis to further assess the impact of missing vital status information on the assessment of mortality risk. The sensitivity analysis was conducted by assuming each of those participants who discontinued Trial AACI due to severe adverse events to be dead at random time point

after the discontinuation but within the 76 week visit. The sensitivity analysis is provided in Section 5 of this addendum.

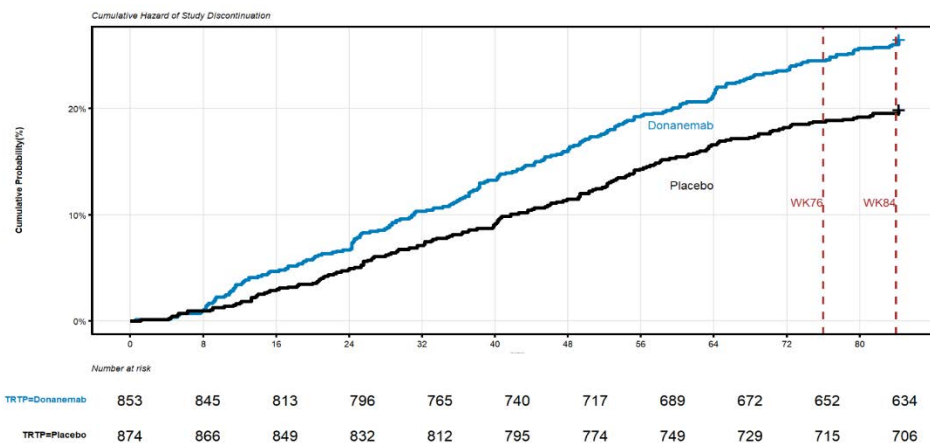
3. Assessment of Trial Disposition

The trial disposition remains the same.

3.1 Trial Discontinuation

Patterns of trial discontinuation have little changes from the original memo. Figure 1A depicts the time to trial discontinuation for Trial AACI. At 8 weeks, participants on donanemab start to discontinue at a higher rate than participants on placebo. As time progresses, the two curves continue to diverge slightly implying greater follow up for subjects on placebo than donanemab.

Figure 1A. Time to Trial Discontinuation, Trial AACI (Safety Population)



Source: Reviewer's analysis using SUBJINFO.XPT in SDN161

3.2 Baseline Characteristics – Overall Population and Participants Discontinuing Trial

Table 1A below depicts demographics and baseline characteristics for the overall population as well as for those participants that discontinued the study early. Interpretation of the findings from this descriptive summary is deferred to the clinical review team.

Table 1A. Baseline Demographics and Characteristics, Trial AACI (Safety Population)

		Overall Safety Population		Discontinued Trial Participants	
		Donanemab	Placebo	Donanemab	Placebo
		N=853	N=874	N=221	N=170
Sex, n (%)					
	Female	488 (57.2)	501 (57.3)	132 (59.7)	103 (60.6)
	Male	365 (42.8)	373 (42.7)	89 (40.3)	67 (39.4)
Age, years					
	Mean (SD)	72.9 (6.15)	73.1 (6.19)	73.7 (5.92)	74.7 (6.08)
	Median	73	74	74	76
	IQR	69.0, 78.0	69.0, 78.0	70.0, 78.0	71.0, 79.0
	Min, Max	59.0, 86.0	59.0, 85.0	59.0, 85.0	59.0, 85.0
Age categories, n (%)					
	<65	88 (10.3)	86 (9.8)	17 (7.7)	10 (5.9)
	>=65	765 (89.7)	788 (90.2)	204 (92.3)	160 (94.1)
RACE, n (%)					
NATIVE	AMERICAN INDIAN OR ALASKA	2 (<1)	0	1 (<1)	0
	ASIAN	57 (6.7)	47 (5.4)	12 (5.4)	8 (4.7)
AMERICAN	BLACK OR AFRICAN	19 (2.2)	20 (2.3)	4 (1.8)	3 (1.8)
	MULTIPLE	0	1 (<1)	0	0
	WHITE	774 (90.7)	806 (92.2)	204 (92.3)	159 (93.5)
	Missing	1 (<1)	0	0	0
Ethnicity, n (%)					
	HISPANIC OR LATINO	35 (4.1)	37 (4.2)	10 (4.5)	12 (7.1)
	NOT HISPANIC OR LATINO	806 (94.5)	823 (94.2)	208 (94.1)	157 (92.4)
	NOT REPORTED	12 (1.4)	14 (1.6)	3 (1.4)	1 (<1)
Pooled Site Group 1 (Countries Grp.), n (%)					
	Australia	13 (1.5)	4 (<1)	3 (1.4)	1 (<1)
	Europe	115 (13.5)	124 (14.2)	29 (13.1)	26 (15.3)
	Japan	45 (5.3)	43 (4.9)	7 (3.2)	7 (4.1)
Canada)	North American (US and	674 (79.0)	694 (79.4)	182 (82.4)	136 (80.0)
	Missing	6 (<1)	9 (1.0)	0	0
BMI Baseline Category, n (%)					
	Extreme obese (>=40 kg/m2)	11 (1.3)	7 (<1)	3 (1.4)	3 (1.8)

kg/m ²)	Normal (≥ 18.5 and < 25)	357 (41.9)	380 (43.5)	98 (44.3)	75 (44.1)
	Obese (≥ 30 and < 40 kg/m ²)	143 (16.8)	146 (16.7)	40 (18.1)	28 (16.5)
kg/m ²)	Overweight (≥ 25 and < 30)	310 (36.3)	307 (35.1)	75 (33.9)	56 (32.9)
	Underweight (< 18.5 kg/m ²)	32 (3.8)	34 (3.9)	5 (2.3)	8 (4.7)
BMI Tertiles, n (%)					
	< 23.36	268 (31.4)	298 (34.1)	78 (35.3)	62 (36.5)
	$23.36 - < 27.50$	293 (34.3)	286 (32.7)	71 (32.1)	53 (31.2)
	≥ 27.50	286 (33.5)	281 (32.2)	72 (32.6)	55 (32.4)
	Missing	6 (< 1)	9 (1.0)	0	0
Baseline for BMI					
	Mean (SD)	26.0 (4.83)	25.7 (4.76)	25.9 (5.00)	25.7 (5.18)
	Median	25.5	25.2	25.1	25.1
	IQR	22.6, 28.7	22.1, 28.7	22.3, 28.9	22.0, 28.8
	Min, Max	13.0, 47.4	15.9, 49.4	15.6, 44.2	16.4, 42.0
	Missing	6	9	0	0
ApoE status, n (%)					
	e2/e2	0	1 (< 1)	0	0
	e2/e3	18 (2.1)	20 (2.3)	6 (2.7)	5 (2.9)
	e2/e4	22 (2.6)	25 (2.9)	4 (1.8)	7 (4.1)
	e3/e3	237 (27.8)	229 (26.2)	53 (24.0)	48 (28.2)
	e3/e4	430 (50.4)	449 (51.4)	114 (51.6)	87 (51.2)
	e4/e4	143 (16.8)	146 (16.7)	41 (18.6)	19 (11.2)
	Missing	3 (< 1)	4 (< 1)	3 (1.4)	4 (2.4)
ApoE4 Status, n (%)					
	Carrier	595 (69.8)	620 (70.9)	159 (71.9)	113 (66.5)
	Non-Carrier	255 (29.9)	250 (28.6)	59 (26.7)	53 (31.2)
	Missing	3 (< 1)	4 (< 1)	3 (1.4)	4 (2.4)
ApoE Homozygous Status, n (%)					
	Heterozygote	452 (53.0)	474 (54.2)	118 (53.4)	94 (55.3)
	Homozygote	143 (16.8)	146 (16.7)	41 (18.6)	19 (11.2)
	Non-Carrier	255 (29.9)	250 (28.6)	59 (26.7)	53 (31.2)
	Unknown	3 (< 1)	4 (< 1)	0	0
	Missing	0	0	3 (1.4)	4 (2.4)
Disease Severity Status, n (%)					
	MCI (MMSE > 26)	144 (16.9)	136 (15.6)	19 (8.6)	16 (9.4)

Mild (20<=MMSE<=26)	702 (82.3)	729 (83.4)	201 (91.0)	154 (90.6)
Moderate (MMSE<20)	1 (<1)	0	1 (<1)	0
Missing	6 (<1)	9 (1.0)	0	0
Screening/Baseline MMSE Total Score				
Mean (SD)	23.7 (2.49)	23.6 (2.48)	23.1 (2.28)	22.9 (2.40)
Median	24	23	23	22
IQR	22.0, 26.0	21.0, 26.0	21.0, 25.0	21.0, 25.0
Min, Max	18.0, 28.0	20.0, 29.0	18.0, 28.0	20.0, 28.0
BL Antithrombotic use, n (%)	298 (34.9)	320 (36.6)	75 (33.9)	61 (35.9)
BL Anticoagulant use, n (%)	47 (5.5)	48 (5.5)	21 (9.5)	8 (4.7)
BL Aspirin use, n (%)	243 (28.5)	261 (29.9)	52 (23.5)	49 (28.8)
BL Aspirin >81 mg daily use, n (%)	12 (1.4)	18 (2.1)	4 (1.8)	2 (1.2)
BL Aspirin >81mg daily or antiplat usage, n (%)	49 (5.7)	47 (5.4)	17 (7.7)	9 (5.3)
BL Non-aspirin antiplatelet usage, n (%)	39 (4.6)	29 (3.3)	15 (6.8)	7 (4.1)
BL Hx of diabetes, n (%)	139 (16.3)	154 (17.6)	45 (20.4)	26 (15.3)
BL Hx of Hypertension, n (%)	469 (55.0)	480 (54.9)	118 (53.4)	106 (62.4)
BL HX of Cerebrovascular Disorder, n (%)	540 (63.3)	556 (63.6)	136 (61.5)	117 (68.8)
BL Prior cerebrovascular accident, n (%)	76 (8.9)	64 (7.3)	18 (8.1)	14 (8.2)
BL Prior cancer diagnosis w/in 5 years, n (%)	84 (9.8)	86 (9.8)	23 (10.4)	18 (10.6)
BL Number of microhemorrhages, n (%)				
0	742 (87.0)	736 (84.2)	196 (88.7)	142 (83.5)
1	80 (9.4)	88 (10.1)	20 (9.0)	13 (7.6)
2	19 (2.2)	31 (3.5)	3 (1.4)	11 (6.5)
3	7 (<1)	13 (1.5)	1 (<1)	3 (1.8)
4	5 (<1)	5 (<1)	1 (<1)	0
Missing	0	1 (<1)	0	1 (<1)
BL superficial siderosis only, n (%)	13 (1.5)	24 (2.7)	4 (1.8)	7 (4.1)
BL Number of superficial siderosis, n (%)				
1	16 (1.9)	29 (3.3)	5 (2.3)	7 (4.1)
2	0	1 (<1)	0	0
BL Severity of WMD, n (%)				
Severity 1	713 (83.6)	721 (82.5)	185 (83.7)	136 (80.0)
Severity 2	71 (8.3)	84 (9.6)	22 (10.0)	26 (15.3)
Severity 3 or more	0	1 (<1)	0	0

Abbreviations: IQR = interquartile range, ITT = Intention-to-treat population, SD = standard deviation BL = baseline, Hx = History, WMD = White Matter Disease.

Source: Reviewer's analysis using ADSL.XPT in SDN104 and SUBJINFO.XPT in SDN161.

3.3 Safety Profile – Overall Population and Participants Discontinuing Trial

The assessment of safety profiles among overall populations and participants discontinuing trial remain unchanged from the original memorandum.

4. Mortality Assessment

4.1. Trial AACI

Error! Reference source not found. A depicts analysis results for mortality at both 76 weeks and 76 weeks + 57 days when assuming that all censored subjects are non-informative (also see Figure 3A for the Kaplan-Meier curve for the cumulative probability of death). At 76 weeks, the cumulative incidence (i.e. cumulative percentage) of death for donanemab is 2.2% versus 1.2% for placebo with an estimated risk difference of 1.0% and 95% CI of (-0.3%, 2.3%). A risk difference of 1.0% for death at week 76, means that if 1000 patients from the population were treated with donanemab rather than placebo for 76 weeks, 10 additional patients would be expected to die. Estimates for mortality at 76 weeks + 57 days resulted in slightly increased estimates of the RD (1.1%) and HR (1.84). A listing of subjects who died is provided in the Appendix (Section A.2A).

Table 5A. Mortality Assessment of Trial AACI (Safety Population); On-study

n, (% ¹)	Donanemab N = 853	Placebo N = 874	RD (95% CI)	HR ² (95% CI)
Mortality at 76 weeks	17 (2.2%)	10 (1.2%)	1.0% (-0.3%, 2.3%)	1.80(0.83, 3.94)
Mortality at 76 weeks+57 days	19 (2.5%)	11 (1.4%)	1.1% (-0.2%, 2.5%)	1.84 (0.88, 3.88)

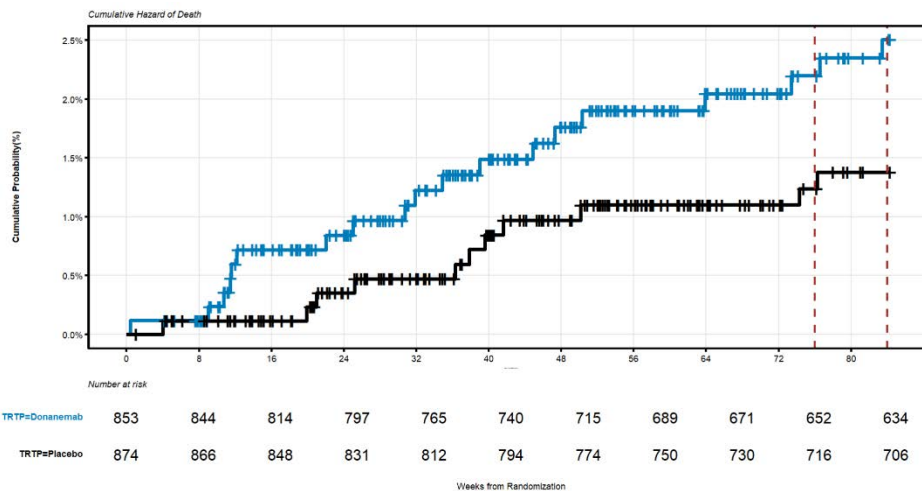
Source Reviewer's analysis using SUBJINFO.XPT in SDN161

¹ Kaplan-Meier estimates of cumulative incidence

² Hazard ratio estimate from Cox PH model without covariate adjustment

Abbreviations: HR = hazard ratio, RD = risk difference

Figure 3A. Kaplan-Meier Cumulative Probability for Death, Trial AACI (Safety Population); On-Study



Source: Reviewer's analysis using SUBJINFO.XPT in SDN161

4.2 Integrated Analysis – Trials AACG and AACI

Trial AACG contributed one death to each treatment arm. Integrating the deaths across both trials resulted in an estimated incidence rate difference of 0.61 deaths per 100 PY with 95% CI of (-0.21, 1.44). Roughly speaking this means that if 1000 patients were treated for 1 year, we would expect to see about 6 more deaths assuming the rate of death is constant over time.

Table 6A. Mortality Assessment of Trial AACI and AACG (Safety Population); On-study

	Donanemab N = 984	Placebo N = 999	IRD ¹ (95% CI)	HR (95% CI)
n (IR, per 100 PY)				
Mortality at 76 weeks	18 (1.44)	11 (0.84)	0.61 (-0.21, 1.44)	1.72(0.81, 3.65)
Trial AACI	17 (1.56)	10 (0.86)	0.70 (-0.21, 1.61)	
Trial AACG*	1 (0.62)	1 (0.63)	-0.005 (-1.73, 1.73)	

Source: Reviewer's analysis using SUBJINFO.XPT in SDN161

*In AACG study, Donanemab = LY3002813-M = 131

Abbreviations: IR, incidence rate; IRD= incidence rate difference, PY, person-years; CI, confidence intervals; HR, hazard ratio.

¹ Incidence rate difference for both trials is calculated using the MH method.

² HR is calculated using the stratified Cox model (stratified by trial)

4.3 Additional Observed Deaths in Trial AACI

The on-study analysis of mortality at 76 weeks, was based on including any death that occurred prior to or at week 76 regardless of discontinuation of randomized treatment. There were several other notable deaths observed in Trial AACI summarized below (see also Table A.1A in the Appendix). These deaths were not included in the 76 week mortality analysis results in Table 5A.

- Donanemab additional deaths
 - o Subject (b) (6) died 4 days after week 76 with reason listed as "Death."
 - o Subject (b) (6) died due to cerebrovascular accident 52 days after the week 76 visit with a fatal adverse event recorded prior to week 76.

- Placebo additional deaths
 - o Subject (b) (6) died 2 days after week 76 of a malignant lung neoplasm with a fatal adverse event recorded prior to week 76.

Responding to an informal discussion with the clinical review team on February 27, 2024, the sponsor established a process to retrieve vital status information from for the cohort of subjects who discontinued from the trial due to a reason of lost to follow-up. From the original memorandum, this represents 22 of 291 subjects that discontinued from the trial prior to week 76. As of April 23, 2024, the sponsor was able to retrieve vital status information from 18 of the 22 participants that were lost to follow up (SDN 168, 169). Of these 18 participants, four were identified to have died (2 on donanemab and 2 on placebo). Three of the deaths occurred after what would have been the week 76 visit, and one death on Placebo (ID (b) (6)) would have occurred prior to the week 76 visit. Causes of death and electronic data on these participants has not been submitted.

5. Sensitivity Analysis on Missing Mortality Data

Given that there is high percentage of missing vital status information at week 76 (26% for donanemab and 20% for placebo), a sensitivity analysis was conducted to assess the impact of missing data on the risk assessment. Among the participants who discontinued before week 76, there are 26 participants discontinued due to severe adverse events (16 in donanumab arm and 10 in placebo arm). Assuming each of these participants died within the 76 weeks at random time after the discontinuation, the risk differences and 95% confidence intervals can be re-calculated. Table A.2A in the Appendix presents the re-calculated risk differences and 95% confidence intervals, and it shows that the lower bound of 95% confidence intervals could go beyond zero only if there were two or three more additional death in the denonemab arm than the placebo arm among these 26 participants who discontinued due to severe adverse events.

6. Summary and Conclusions

The assessment of mortality, primarily based on the placebo-controlled portion of Trial AACI, resulted in an estimated risk difference of 1.0% for death at week 76 with 95% CI of (-0.3%, 2.3%). This means that if 1000 patients from the population were treated with donanemab rather than placebo for 76 weeks, 10 additional patients would be expected to die. Estimates for mortality at 76 weeks + 57 days resulted in slightly increased estimate of the RD (1.1%). An integrated analysis that incorporates 76 week data from the Phase 2 trial, Trial AACG, resulted in similar findings: IRD = 0.61 with 95% CI (-0.21, 1.44).

The main conclusions and recommendations remain the same. Overall, the mortality analyses identified a potential safety signal of an increased risk of death - the point estimate from the analysis of Trial AACI suggests an increase of 1.0% in the risk of death on donanemab vs. placebo with the upper 95% CI bound failing to rule out differences as large as 2.3%. In addition, the conduct of Trial AACI further adds uncertainty to the observed mortality effect sizes and there is some evidence to suggest that additional

vital status information from those that discontinued from the study may result in mortality effect sizes greater than those observed thereby suggesting increased levels of harm with donanemab relative to placebo. It's important that such uncertainty be incorporated into the benefit-risk assessment of donanemab along with the observed increased estimates of mortality risk of donanemab relative to placebo.

Based on information submitted in SDN 168 and 169, the sponsor was able to retrieve 18 out the 22 participants that discontinued from the trial early due to lost to follow-up within less than 2 months from the informal request. This informed that the sponsor has a process to retrieve vital status information and an information request was sent on April 24, 2024 to request the vital status for all the participants that discontinued Trial AACI early. This information once available would help reduce uncertainties around the mortality findings.

APPENDIX

A.1 Protocol Defined Information on Participant Discontinuation/Withdrawal from the Study

The information below is quoted from Section 7 of the protocol for Trial AACI. This suggests that by design, when a participant withdraws from the study no additional data (including vital status) is collected on that participant.

"A participant may withdraw from the study:

- at any time at his/her own request
- at the request of his/her designee (for example, parents or legal guardian)
- at the discretion of the investigator for safety, behavioral, compliance, or administrative reasons
- if the participant becomes pregnant during the study
- if the participant enrolls in any other clinical study involving an investigational medicinal product or enrollment in any other type of medical research judged not to be scientifically or medically compatible with this study
- if the participant, for any reason, requires treatment with another disease-modifying therapeutic agent that has been demonstrated to be effective for treatment of the study indication, discontinuation from the study occurs prior to introduction of the new agent
- participation in the study needs to be stopped for medical, safety, regulatory, or other reasons consistent with applicable laws, regulations, and good clinical practice (GCP).

At the time of discontinuing from the study, if possible, an early discontinuation (ED) visit should be conducted, as shown in the SoA. See SoA for data to be collected at the time of study discontinuation and follow-up and for any further evaluations that need to be completed. The participant will be permanently discontinued both from the IP and from the study at that time.

If the participant withdraws consent for disclosure of future information, the sponsor may retain and continue to use any data collected before such a withdrawal of consent. If a participant withdraws from

the study, he/she may request destruction of any samples taken and not tested, and the investigator must document this in the site study records.”

A.2A Listing of Participants that Died and Included in Mortality Analysis

Table A.1A depicts the 31 participants (19 donanemab, 12 placebo) that died during the placebo-controlled portion of Trial AACI. Among this, one participant (15T-MC-AACI- (b) (6)) is among the loss to follow-up cohort and the vital status was retrieved in a separate process (SDN169).

Table A.1A. Listing of Participants that Died by Week 76 + 57 Days, Trial AACI (Safety Population)

USUBJID	Treatment Arm	Age	Sex	Days to Death	Cause of Death
15T-MC-AACI- (b) (6)	Donanemab	69	F	3	Pulmonary Embolism
15T-MC-AACI- (b) (6)	Donanemab	83	F	63	Death
15T-MC-AACI- (b) (6)	Donanemab	72	M	75	ARIA-H
15T-MC-AACI- (b) (6)	Donanemab	72	F	80	ARIA-E
15T-MC-AACI- (b) (6)	Donanemab	68	M	81	Dehydration
15T-MC-AACI- (b) (6)	Donanemab	82	M	85	Coronavirus COVID-19 Pneumonia
15T-MC-AACI- (b) (6)	Donanemab	84	F	154	Death (Cause of death on Death Certificate is late onset Alzheimer's Disease)
15T-MC-AACI- (b) (6)	Donanemab	62	M	175	completed suicide
15T-MC-AACI- (b) (6)	Donanemab	79	F	215	PULMONARY EMBOLISM
15T-MC-AACI- (b) (6)	Donanemab	85	F	223	pneumonia
15T-MC-AACI- (b) (6)	Donanemab	83	M	244	SAH Subarachnoid Hemorrhage
15T-MC-AACI- (b) (6)	Donanemab	82	F	273	Retroperitoneal Bleeding
15T-MC-AACI- (b) (6)	Donanemab	75	M	314	Suicide
15T-MC-AACI- (b) (6)	Donanemab	72	M	331	Respiratory Arrest
15T-MC-AACI- (b) (6)	Donanemab	79	M	352	COVID-19
15T-MC-AACI- (b) (6)	Donanemab	75	M	447	Death
15T-MC-AACI- (b) (6)	Donanemab	75	F	514	The patient died of severe respiratory failure
15T-MC-AACI- (b) (6)	Donanemab	81	F	536*	Natural Cause Death
15T-MC-AACI- (b) (6)	Donanemab	75	F	584*	Cerebral Vascular Accident
15T-MC-AACI- (b) (6)	Placebo	85	M	28	RESPIRATORY FAILURE
15T-MC-AACI- (b) (6)	Placebo	81	M	139	lung infection
15T-MC-AACI- (b) (6)	Placebo	85	F	147	Community Acquired Pneumonia
15T-MC-AACI- (b) (6)	Placebo	83	F	176	Smoke Inhalation
15T-MC-AACI- (b) (6)	Placebo	74	F	254	Heart Attack
15T-MC-AACI- (b) (6)	Placebo	83	M	265	Death
15T-MC-AACI- (b) (6)	Placebo	77	M	277	Suicide by carbon monoxide poisoning
15T-MC-AACI- (b) (6)	Placebo	81	M	291	Sepsis



IST-MC-AACI-	(b) (6)	Placebo	65	F	351	Probable Atherosclerotic Cardiovascular Disease
IST-MC-AACI-		Placebo	78	M	520	Worsening of AD
IST-MC-AACI-		Placebo	76	M	534*	Lung neoplasm malignant
IST-MC-AACI-		Placebo	80	M	458	Unknown

Source: Reviewer's analysis using SUBJINFO.XPT in SDN161, 168 and 169.

* The participant is among the loss-to-follow up cohort and the death was later confirmed in SDN169. Death is not included in the analyses due to the lack of details on the vital status retrieval process.

* Death included only in the 76 weeks + 57 day analysis

A.3A Sensitivity Analysis on missing data

Table A.2A. Sensitivity Analysis: assuming each of the participants who discontinued due to severe adverse events to be dead within 76 weeks, Trial AACI (Safety Population)

	Number of Placebo Participants Imputed to be Dead											
RD (%)		0	1	2	3	4	5	6	7	8	9	10
Number of Donanemab Participants Imputed to be Dead	0	1.2 (-0.3,2.8)	1.1 (-0.4,2.7)	1 (-0.5,2.6)	0.9 (-0.7,2.4)	0.7 (-0.8,2.3)	0.6 (-1.2,2.2)	0.5 (-1.2,2.1)	0.3 (-1.3,2)	0.2 (-1.4,1.9)	0.1 (-1.6,1.8)	0 (-1.7,1.7)
	1	1.4 (-0.1,2.9)	1.3 (-0.3,2.8)	1.1 (-0.4,2.7)	1 (-0.6,2.6)	0.9 (-0.7,2.5)	0.7 (-0.9,2.4)	0.6 (-1.2,2.2)	0.5 (-1.2,2.1)	0.4 (-1.3,2)	0.2 (-1.5,1.9)	0.1 (-1.6,1.8)
	2	1.5 (0.3,3.1)	1.4 (-0.2,3)	1.3 (-0.3,2.8)	1.1 (-0.5,2.7)	1 (-0.6,2.6)	0.9 (-0.8,2.5)	0.7 (-0.9,2.4)	0.6 (-1.1,2.3)	0.5 (-1.2,2.2)	0.3 (-1.4,2.1)	0.2 (-1.5,2)
	3	1.7 (0.1,3.2)	1.5 (0.3,3.1)	1.4 (-0.2,3)	1.3 (-0.3,2.9)	1.1 (-0.5,2.8)	1 (-0.6,2.7)	0.9 (-0.8,2.6)	0.7 (-0.9,2.4)	0.6 (-1.1,2.3)	0.5 (-1.2,2.2)	0.4 (-1.4,2.1)
	4	1.8 (0.2,3.4)	1.7 (0.1,3.3)	1.5 (-0.1,3.1)	1.4 (-0.2,3)	1.3 (-0.4,2.9)	1.1 (-0.5,2.8)	1 (-0.7,2.7)	0.9 (-0.8,2.6)	0.8 (-1.2,5)	0.6 (-1.1,2.4)	0.5 (-1.3,2.3)
	5	1.9 (0.3,3.5)	1.8 (0.2,3.4)	1.7 (0,3.3)	1.5 (-0.1,3.2)	1.4 (-0.3,3.1)	1.3 (-0.4,2.9)	1.1 (-0.6,2.8)	1 (-0.7,2.7)	0.9 (-0.9,2.6)	0.7 (-1.2,5)	0.6 (-1.1,2.4)
	6	2 (0.4,3.6)	1.9 (0.3,3.5)	1.8 (0.1,3.4)	1.7 (0,3.3)	1.5 (-0.2,3.2)	1.4 (-0.3,3.1)	1.3 (-0.5,3)	1.1 (-0.6,2.9)	1 (-0.7,2.8)	0.9 (-0.9,2.6)	0.8 (-1.2,5)
	7	2.2 (0.5,3.8)	2 (0.4,3.7)	1.9 (0.3,3.6)	1.8 (0.1,3.5)	1.6 (-0.1,3.4)	1.5 (-0.2,3.2)	1.4 (-0.4,3.1)	1.3 (-0.5,3)	1.1 (-0.6,2.9)	1 (-0.8,2.8)	0.9 (-0.9,2.7)
	8	2.3 (0.6,4)	2.2 (0.5,3.9)	2.1 (0.4,3.7)	1.9 (0.2,3.6)	1.8 (0.1,3.5)	1.7 (-0.1,3.4)	1.5 (-0.2,3.3)	1.4 (-0.4,3.2)	1.3 (-0.5,3.1)	1.1 (-0.7,2.9)	1 (-0.8,2.8)
	9	2.4 (0.8,4.1)	2.3 (0.6,4)	2.2 (0.5,3.9)	2 (0.3,3.8)	1.9 (0.2,3.6)	1.8 (0.3,5)	1.6 (-0.1,3.4)	1.5 (-0.3,3.3)	1.4 (-0.4,3.2)	1.3 (-0.6,3.1)	1.1 (-0.7,3)
	10	2.6 (0.9,4.3)	2.5 (0.8,4.2)	2.3 (0.6,4.1)	2.2 (0.5,3.9)	2.1 (0.3,3.8)	1.9 (0.2,3.7)	1.8 (0.3,6)	1.7 (-0.1,3.5)	1.6 (-0.3,3.4)	1.4 (-0.4,3.3)	1.3 (-0.6,3.2)
	11	2.7 (1.4,4)	2.6 (0.9,4.3)	2.5 (0.7,4.2)	2.3 (0.6,4.1)	2.2 (0.4,4)	2.1 (0.3,3.9)	1.9 (0.1,3.8)	1.8 (0.3,7)	1.7 (-0.1,3.5)	1.6 (-0.3,3.4)	1.4 (-0.4,3.3)
	12	2.8 (1.1,4.6)	2.7 (1.4,5)	2.6 (0.9,4.4)	2.5 (0.7,4.2)	2.3 (0.5,4.1)	2.2 (0.4,4)	2.1 (0.2,3.9)	1.9 (0.1,3.8)	1.8 (0.3,7)	1.7 (-0.2,3.6)	1.6 (-0.3,3.5)
	13	3 (1.2,4.7)	2.8 (1.1,4.6)	2.7 (1.4,5)	2.6 (0.8,4.4)	2.5 (0.7,4.3)	2.3 (0.5,4.1)	2.2 (0.3,4)	2.1 (0.2,3.9)	1.9 (0.1,3.8)	1.8 (-0.1,3.7)	1.7 (-0.2,3.6)
	14	3.1 (1.4,4.9)	3 (1.2,4.8)	2.9 (1.1,4.7)	2.7 (0.9,4.5)	2.6 (0.8,4.4)	2.5 (0.6,4.3)	2.3 (0.5,4.2)	2.2 (0.3,4.1)	2.1 (0.2,4)	2 (0.1,3.9)	1.8 (-0.1,3.8)
	15	3.3 (1.5,5)	3.1 (1.3,4.9)	3 (1.2,4.8)	2.9 (1.1,4.7)	2.7 (0.9,4.6)	2.6 (0.8,4.5)	2.5 (0.6,4.4)	2.4 (0.5,4.2)	2.2 (0.3,4.1)	2.1 (0.2,4)	2 (0.3,9)
	16	3.4 (1.6,5.2)	3.3 (1.5,5.1)	3.1 (1.3,5)	3 (1.2,4.8)	2.9 (1.4,7)	2.7 (0.9,4.6)	2.6 (0.7,4.5)	2.5 (0.6,4.4)	2.4 (0.4,4.3)	2.2 (0.3,4.2)	2.1 (0.2,4)

Abbreviation: RD: risk difference. The numbers in the parentheses are the 95% confidence intervals of risk differences.

Each of the 26 participants who discontinued due to severe adverse events were assumed died at random points within 76 weeks after the discontinuation.

Bolded cells indicate the lower bound of 95% confidence intervals exceeds zero.

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

/s/

FENGYU ZHAO
05/14/2024 02:39:51 PM

PING LI
05/14/2024 05:52:50 PM

YUEQIN ZHAO
05/14/2024 05:58:04 PM

MATTHEW J SOUKUP
05/15/2024 07:51:50 AM



Memorandum

Date: February 29, 2024

To: Teresa Buracchio, MD, Director Office of Neurology Products
Sally Jo Yasuda, PharmD, Deputy Director for Safety, Division of Neurology 1 (DN1)
Natalie Branagan, MD, Clinical Reviewer (DN1)

From: Ping Li, MS, Statistical Analyst, Division of Analytics and Informatics (DAI)
Fengyu Zhao, PhD, Primary Statistical Reviewer, Division of Biometrics VII (DBVII)
Yueqin Zhao, PhD, Team Lead, DBVII
Mat Soukup, PhD, Deputy Director, DBVII

Application: BLA 761248/SDN 100 (resubmission)

Drug Name: Kisunla (donanemab)

Subject: Mortality Assessment

1 Introduction

From the clinical safety review, there was an observed imbalance in mortality from Phase 3 Trial AACI in BLA 761248. Given the complexities in the design and conduct of Trial AACI, DBVII was consulted to provide further analytical assessment of mortality in the donanemab development program.

This memorandum is focused solely on the assessment of mortality in BLA 761248. It describes the approach taken to analyze mortality, provides findings from the analysis and provides some concluding comments on the observed findings.

2 Approach to Analysis

The following details document the approach used in the assessment of mortality. Note that additional descriptive analyses are also provided in Section 3 to assess the extent of follow-up in Trial AACI and provide descriptive analyses to assess any potential differences in participants that completed the trial vs those that discontinued early¹.

The comparison of interest is donanemab to placebo, based on the treatment actually received in the clinical trial. Analyses will be based on the safety analysis population, defined as all subjects randomized to investigational product (IP) that received at least one dose of IP.

¹ Descriptive analyses are based on data submitted in the original resubmission (SN104).

The mortality assessment will be based on:

- The placebo-controlled period of Trial AACI using data generated during the following two time points²:
 - o 76 weeks which is the end of the double-blind period of Trial AACI
 - o 76 weeks + 57 days which includes 57 days of follow-up after the double-blind extension period³.
- An integrated analysis of Trial AACI and Phase 2 Trial AACG using the 76 week double-blind treatment period.

To address the intercurrent event of treatment discontinuation, analyses will be based on an on-study approach. Such an approach considers participants at risk of death while exposed to randomized IP or not exposed to randomized IP but still being followed for vital status.

The within arm summary measure will be based on the cumulative incidence of death at 76 weeks⁴. Cumulative incidence at week 76 will be estimated using Kaplan-Meier methods. Following the on-study approach to analysis the participant is censored at the time of study completion or last contact date⁵. Study-size adjusted estimates will be provided for integrated analyses.

Comparisons of donanemab to placebo will be based on the absolute risk difference (RD) at 76 weeks as well as using a hazard ratio based on a Cox PH model without covariate adjustment. Comparisons will include an estimate of the treatment effect along with a 95% confidence interval. Integrated analyses will be based on incidence rates at 76 weeks and comparisons will be based on incidence rate differences (IRD). Pooled IRD will be calculated using Mantel-Haenszel methods.

3. Assessment of Trial Disposition

This section of the review is intended to assess the extent of follow-up in Trial AACI and provide descriptive analyses to assess any potential differences in the overall population vs those that discontinued study early. Per Trial AACI protocol, participants were allowed to discontinue from study for a number of reasons (see Appendix A.1) and additional data was not planned to be collected after study discontinuation. To further inquire about this, prior information requests from the review team requested additional information about vital status follow-up from the sponsor. In response⁶, the sponsor stated, “vital status information is not available for the vast majority of participants who discontinued or were lost to follow-up in Study AACI.”

² The mortality assessment will be based on data submitted in the 90 day safety update (SN120).

³ The analysis using the follow-up extension is based on the randomized treatment actually received from trial start to the conclusion of the follow-up extension. By design, after 76 weeks, participants randomized to placebo are assigned to receive placebo. As such, the analysis conducted attributes any deaths observed in the 57 day window to the placebo arm in subjects initially randomized to placebo.

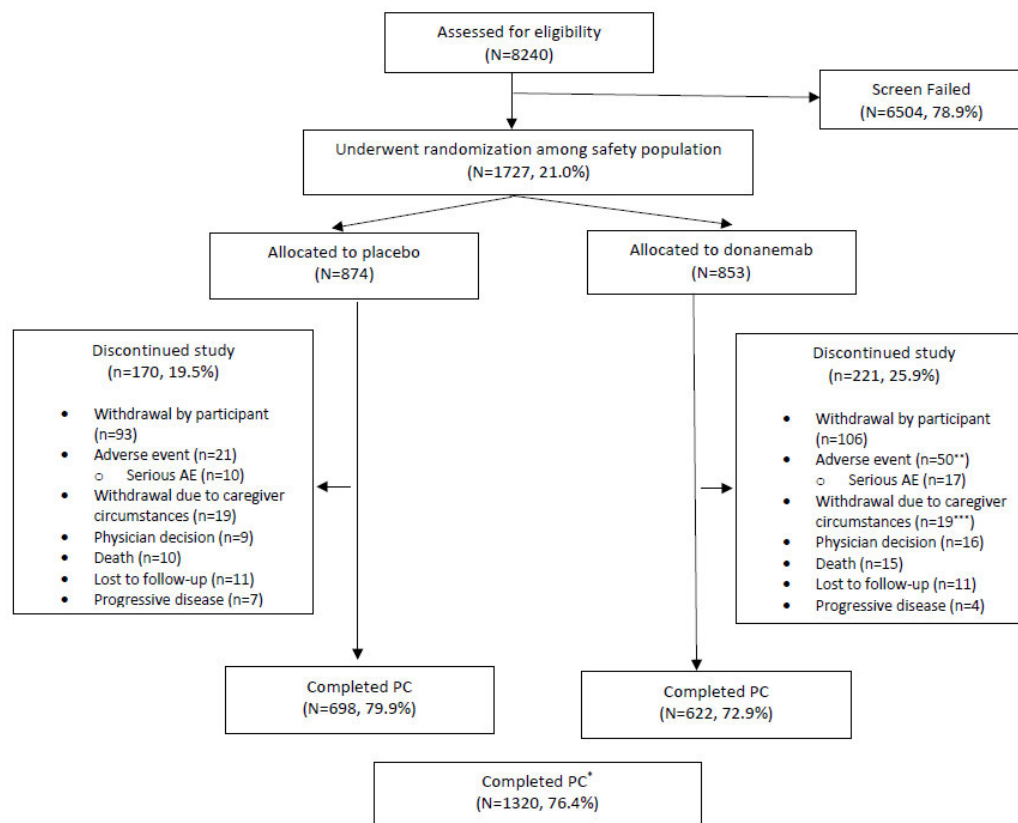
⁴ Assessment that includes the 57 days of follow-up will be based on cumulative incidence at 76 weeks + 57 days.

⁵ Last contact date is defined as the trial discontinuation date.

⁶ See response to information request submitted on 09/11/2023 (SDN: 123)

Figure 1 depicts the trial disposition for Trial AACI for the safety population which forms the basis of the mortality assessment in this memo. Of the 853 participants randomized to donanemab, 221 (25.9%) discontinued the trial early. Of the 874 participants randomized to placebo, 170 (19.5%) discontinued the trial early. Notably, 50 participants on donanemab discontinued due to an adverse event, 17 of which were considered serious, whereas 21 participants on placebo discontinued due to an adverse event, 10 of which were considered serious (see Table 3 with listing of subjects that discontinued due to SAE).

Figure 1. Participant Disposition, Trial AACI (Safety Population)



Source: Reviewer's analysis based on data in SN104

N = number of participants in the population; n = number of participants in the specified category; PC = placebo controlled: 76 weeks+57 days

*12 participants did not complete final visit prior to database lock and are neither considered completers or discontinuers. 4 participants were reported discontinued in the ADSL dataset but in the ADDS dataset the discontinuation dates are beyond placebo controlled period.

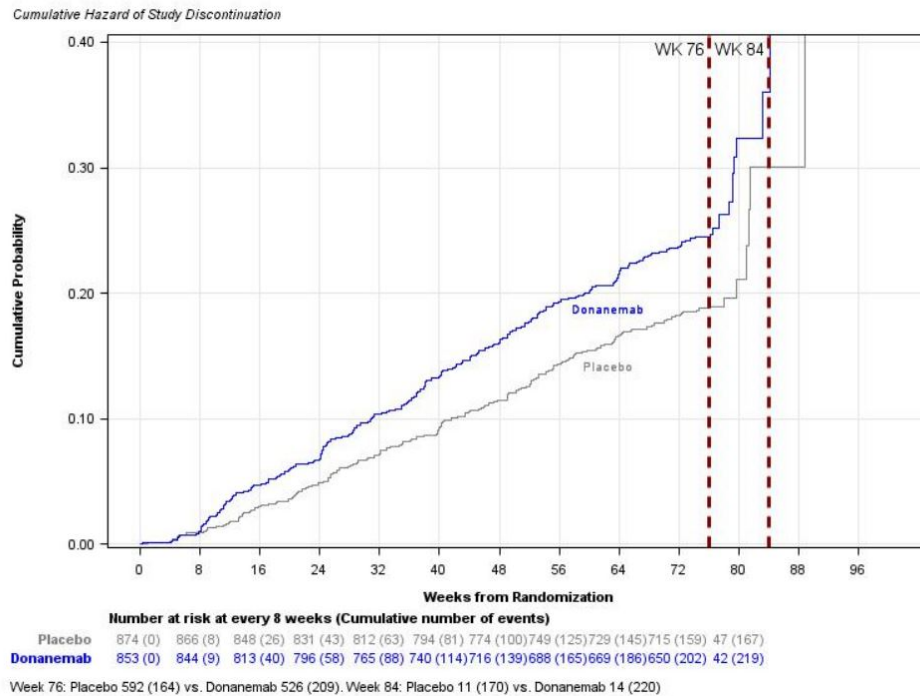
** 1 patient discontinued due to AE and then died within the PC period.

***1 patient discontinued due to caregiver circumstances and then died within the PC period.

3.1 Time to Trial Discontinuation

Figure 2 depicts the time to trial discontinuation for Trial AACI. At 8 weeks, participants on donanemab start to discontinue at a higher rate than participants on placebo. As time progresses, the two curves continue to diverge slightly implying greater follow up for subjects on placebo than donanemab.

Figure 2. Time to Trial Discontinuation, Trial AACI (Safety Population)



Source: Reviewer's analysis using ADLX.XPT in SN104

3.2 Baseline Characteristics – Overall Population and Participants Discontinuing Trial

Table 1 below depicts demographics and baseline characteristics for the overall population as well as for those participants that discontinued the study early. Interpretation of the findings from this descriptive summary is deferred to the clinical review team.

Table 1. Baseline Demographics and Characteristics, Trial AACI (Safety Population)

	Overall Safety Population		Discontinued Trial Participants	
	Donanemab	Placebo	Donanemab	Placebo
	N=853	N=874	N=221	N=170
Sex, n (%)				
Female	488 (57.2)	501 (57.3)	132 (59.7)	103 (60.6)
Male	365 (42.8)	373 (42.7)	89 (40.3)	67 (39.4)
Age, years				
N	853	874	221	170
Mean (SD)	72.9 (6.15)	73.1 (6.19)	73.7 (5.92)	74.7 (6.08)
Median	73.0	74.0	74.0	76.0
IQR	69.0, 78.0	69.0, 78.0	70.0, 78.0	71.0, 79.0
Min, Max	59.0, 86.0	59.0, 85.0	59.0, 85.0	59.0, 85.0
Age categories, n (%)				

<65	88 (10.3)	86 (9.8)	17 (7.7)	10 (5.9)
>=65	765 (89.7)	788 (90.2)	204 (92.3)	160 (94.1)
RACE, n (%)				
AMERICAN INDIAN OR ALASKA NATIVE	2 (<1)	0	1 (<1)	0
ASIAN	57 (6.7)	47 (5.4)	12 (5.4)	8 (4.7)
BLACK OR AFRICAN AMERICAN	19 (2.2)	20 (2.3)	4 (1.8)	3 (1.8)
MULTIPLE	0	1 (<1)	0	0
WHITE	774 (90.7)	806 (92.2)	204 (92.3)	159 (93.5)
Missing	1 (<1)	0	0	0
Ethnicity, n (%)				
HISPANIC OR LATINO	35 (4.1)	37 (4.2)	10 (4.5)	12 (7.1)
NOT HISPANIC OR LATINO	806 (94.5)	823 (94.2)	208 (94.1)	157 (92.4)
NOT REPORTED	12 (1.4)	14 (1.6)	3 (1.4)	1 (<1)
Pooled Site Group 1 (Countries Grp.), n (%)				
Australia	13 (1.5)	4 (<1)	3 (1.4)	1 (<1)
Europe	115 (13.5)	124 (14.2)	29 (13.1)	26 (15.3)
Japan	45 (5.3)	43 (4.9)	7 (3.2)	7 (4.1)
North American (US and Canada)	674 (79.0)	694 (79.4)	182 (82.4)	136 (80.0)
Missing	6 (<1)	9 (1.0)	0	0
BMI Baseline Category, n (%)				
Extreme obese (>=40 kg/m2)	11 (1.3)	7 (<1)	3 (1.4)	3 (1.8)
Normal (>=18.5 and <25 kg/m2)	357 (41.9)	380 (43.5)	98 (44.3)	75 (44.1)
Obese (>=30 and <40 kg/m2)	143 (16.8)	146 (16.7)	40 (18.1)	28 (16.5)
Overweight (>=25 and <30 kg/m2)	310 (36.3)	307 (35.1)	75 (33.9)	56 (32.9)
Underweight (<18.5 kg/m2)	32 (3.8)	34 (3.9)	5 (2.3)	8 (4.7)
BMI Tertiles, n (%)				
<23.36	268 (31.4)	298 (34.1)	78 (35.3)	62 (36.5)
23.36 - < 27.50	293 (34.3)	286 (32.7)	71 (32.1)	53 (31.2)
>= 27.50	286 (33.5)	281 (32.2)	72 (32.6)	55 (32.4)
Missing	6 (<1)	9 (1.0)	0	0
Baseline for BMI				
N	847	865	221	170
Mean (SD)	26.0 (4.83)	25.7 (4.76)	25.9 (5.00)	25.7 (5.18)
Median	25.5	25.2	25.1	25.1
IQR	22.6, 28.7	22.1, 28.7	22.3, 28.9	22.0, 28.8
Min, Max	13.0, 47.4	15.9, 49.4	15.6, 44.2	16.4, 42.0
Missing	6	9	0	0

Not on Antithrombotic, n (%)	501 (58.7)	507 (58.0)	131 (59.3)	103 (60.6)
On Antithrombotic, n (%)	346 (40.6)	358 (41.0)	90 (40.7)	67 (39.4)
Anticoagulation (alone or combined with other antithrombotic), n (%)	82 (9.6)	86 (9.8)	30 (13.6)	17 (10.0)
Aspirin ≤ 81 mg alone, n (%)	204 (23.9)	219 (25.1)	41 (18.6)	42 (24.7)
Aspirin > 81 mg, other antiplatelet or combination of aspirin and other antiplatelet, n (%)	60 (7.0)	53 (6.1)	19 (8.6)	8 (4.7)
ApoE status, n (%)				
e2/e2	0	1 (<1)	0	0
e2/e3	18 (2.1)	20 (2.3)	6 (2.7)	5 (2.9)
e2/e4	22 (2.6)	25 (2.9)	4 (1.8)	7 (4.1)
e3/e3	237 (27.8)	229 (26.2)	53 (24.0)	48 (28.2)
e3/e4	430 (50.4)	449 (51.4)	114 (51.6)	87 (51.2)
e4/e4	143 (16.8)	146 (16.7)	41 (18.6)	19 (11.2)
Missing	3 (<1)	4 (<1)	3 (1.4)	4 (2.4)
ApoE4 Status, n (%)				
Carrier	595 (69.8)	620 (70.9)	159 (71.9)	113 (66.5)
Non-Carrier	255 (29.9)	250 (28.6)	59 (26.7)	53 (31.2)
Missing	3 (<1)	4 (<1)	3 (1.4)	4 (2.4)
ApoE Homozygous Status, n (%)				
Heterozygote	452 (53.0)	474 (54.2)	118 (53.4)	94 (55.3)
Homozygote	143 (16.8)	146 (16.7)	41 (18.6)	19 (11.2)
Non-Carrier	255 (29.9)	250 (28.6)	59 (26.7)	53 (31.2)
Unknown	3 (<1)	4 (<1)	0	0
Missing	0	0	3 (1.4)	4 (2.4)
Disease Severity Status, n (%)				
MCI (MMSE>26)	144 (16.9)	136 (15.6)	19 (8.6)	16 (9.4)
Mild (20≤MMSE≤26)	702 (82.3)	729 (83.4)	201 (91.0)	154 (90.6)
Moderate (MMSE<20)	1 (<1)	0	1 (<1)	0
Missing	6 (<1)	9 (1.0)	0	0

Abbreviations: IQR = interquartile range, ITT = Intention-to-treat population, SD = standard deviation

Source: Reviewer's analysis using ADSL.XPT in SN104

3.3 Safety Profile – Overall Population and Participants Discontinuing Trial

Table 2 depicts a high-level summary of the AE and SAE incidence at 76 weeks + 57 days for each treatment arm both overall and in participants that discontinue trial. In this table, AE and SAE incidence refer to the participant experiencing an AE or SAE at any time during the trial. Note it does not depict that AEs or SAEs that lead to trial discontinuation (refer to Figure 1 above) nor does it capture any AE occurrence after a participant discontinued from study. In addition, estimates of incidence for

discontinued trial participants would be expected to be even higher when accounting for time in the estimation procedure. Focusing on the last two columns of Table 2, there is a higher incidence of subjects discontinuing the trial and experiencing an AE or SAE on donanemab relative to placebo.

Table 2 Tabular summary of AE and SAE Incidence at Week 76 + 57 Days, Trial AACI (Safety Population)

	Overall Population		Discontinued Trial Participants	
n, (%)	Donanemab N = 853	Placebo N = 874	Donanemab (n = 221)	Placebo (n = 170)
AE Incidence	768 (90.0%)	729 (83.4%)	202 (23.7%) 91.4% ¹	133 (15.2%) 78.2% ¹
SAE Incidence	148 (17.4%)	138 (15.8%)	67 (7.9%) 30.3% ¹	52 (5.9%) 30.6% ¹

Source: Reviewer's analysis using ADAE.XPT in SN104

Percentages for the overall population are based on crude percentages (i.e. $100 \times n/N$); percentages for the discontinued trial participants is based on the crude proportion that experience an AE/SAE at any time for those that discontinue trial.

¹ Percentage is based on the number of participants that discontinue trial (i.e. 221 for donanemab and 170 for placebo); such percentages should only be used to qualitatively compare against the overall population.

Table 3 lists the 27 subjects that discontinued Trial AACI by 76 weeks + 57 days due to an SAE. We defer to clinical review on the interpretation of the findings from this table.

Table 3. Listing of Participants that Discontinued Study Due to SAE, Trial AACI (Safety Population)

SUBJID	ACTARM	AGE	SEX	Days Until Trial Discontinuation	Verbatim Term for SAE	MedDRA SAE PT
(b) (6)	Donanemab	75	M	29	ARIA-E	Amyloid related imaging abnormality-oedema/effusion
	Donanemab	75	M	61	Hospitalization for psychotic agitation	Agitation
	Donanemab	68	M	63	Paranoid delirium with psychotic elements	Delirium
	Donanemab	74	M	89	Delirium	Delirium
	Donanemab	82	M	119	RIGHT HOLOHEMISPHERIC SUBDURAL HEMATOMA	Subdural haematoma
	Donanemab	80	M	126	Symptomatic ARIA-E	Amyloid related imaging abnormality-oedema/effusion
	Donanemab	71	M	156	multiple myeloma	Plasma cell myeloma



(b) (6)	Donanemab	77	F	193	Symptomatic ARIA-E	Amyloid related imaging abnormality-oedema/effusion
	Donanemab	64	F	244	stage 4 adenocarcinoma of the lung	Lung adenocarcinoma stage IV
	Donanemab	79	M	249	ARIA-E	Amyloid related imaging abnormality-oedema/effusion
	Donanemab	80	M	261	Multiple Myeloma	Plasma cell myeloma
	Donanemab	67	M	280	ARIA E	Amyloid related imaging abnormality-oedema/effusion
	Donanemab	68	F	309	Severe Symptomatic ARIA-E	Amyloid related imaging abnormality-oedema/effusion
	Donanemab	61	F	316	pleomorphic invasive lobular breast carcinoma	Invasive lobular breast carcinoma
	Donanemab	75	F	385	Pneumonia	Pneumonia
	Donanemab	62	F	469	Acute agitated state	Agitation
	Donanemab	76	M	491	Symptomatic Sinus Bradycardia	Sinus bradycardia
	Placebo	67	F	61	Non-Hodgkin lymphoma	Non-Hodgkin's lymphoma
	Placebo	77	F	105	Stroke (Primary Haemorrhagic)	Haemorrhagic stroke
	Placebo	74	M	165	Recurrent Symptomatic Bradycardia	Bradycardia
	Placebo	80	M	177	Ischemic stroke of mid basilar artery	Ischaemic stroke
	Placebo	62	M	218	Sectioned under the Mental Health Act 1983	Mental disorder
	Placebo	73	F	364	Electrocardiogram QT corrected interval prolonged	Electrocardiogram QT prolonged
	Placebo	76	M	381	Lung Cancer	Lung neoplasm malignant
	Placebo	68	M	394	Stage IV Adenocarcinoma of Lung	Lung adenocarcinoma stage IV



(b) (6)	Placebo	76	F	496	Squamous cell carcinoma of rectum-Biopsy performed (b) (6)	Rectal cancer
---------	---------	----	---	-----	--	---------------

Source: Reviewer's analysis; listing is based on participants that discontinued by week 76 + 57 days using ADAE.XPT in SN104
Abbreviations: ID, identification; MedDRA, Medical Dictionary for Regulatory Activities; PT, preferred term; VT

4. Mortality Assessment

4.1. Trial AACI

Error! Reference source not found. depicts analysis results for mortality at both 76 weeks and 76 weeks + 57 days when assuming that all censored subjects are non-informative (also see Figure 3 for the Kaplan-Meier curve for the cumulative probability of death). At 76 weeks, the cumulative incidence (i.e. cumulative percentage) of death for donanemab is 2.2% versus 1.2% for placebo with an estimated risk difference of 1.0% and 95% CI of (-0.3%, 2.3%). A risk difference of 1.0% for death at week 76, means that if 1000 patients from the population were treated with donanemab rather than placebo for 76 weeks, 10 additional patients would be expected to die. Estimates for mortality at 76 weeks + 57 days resulted in slightly increased estimates of the RD (1.1%) and HR (1.84). A listing of subjects who died is provided in the Appendix (Section A.2).

Table 5. Mortality Assessment of Trial AACI (Safety Population); On-study

n, (%)	Donanemab N = 853	Placebo N = 874	RD (95% CI)	HR ² (95% CI)
Mortality at 76 weeks	17 (2.2%)	10 (1.2%)	1.0% (-0.3%, 2.3%)	1.79 (0.82, 3.93)
Mortality at 76 weeks+57 days	19 (2.5%)	11 (1.4%)	1.1% (-0.2%, 2.5%)	1.84 (0.87, 3.86)

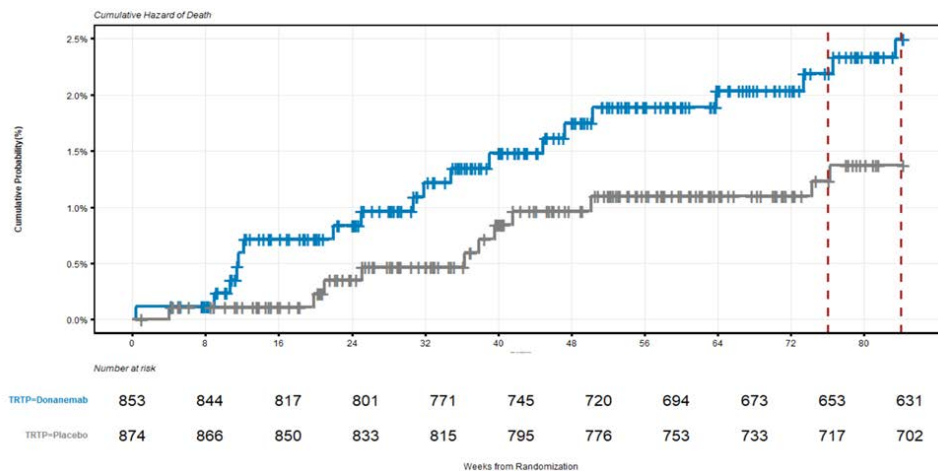
Source Reviewer's analysis using ADSL.XPT in SN120

¹ Kaplan-Meier estimates of cumulative incidence

² Hazard ratio estimate from Cox PH model without covariate adjustment

Abbreviations: HR = hazard ratio, RD = risk difference

Figure 3. Kaplan-Meier Cumulative Probability for Death, Trial AACI (Safety Population); On-Study



Source: Reviewer's analysis using ADSL in SN120

4.2 Integrated Analysis – Trials AACG and AACI

Trial AACG contributed one death to each treatment arm. Integrating the deaths across both trials resulted in an estimated incidence rate difference of 0.59 deaths per 100 PY with 95% CI of (-0.22, 1.41). Roughly speaking this means that if 1000 patients were treated for 1 year, we would expect to see about 6 more deaths assuming the rate of death is constant over time.

Table 6. Mortality Assessment of Trial AACI and AACG (Safety Population); On-study

	Donanemab N = 999	Placebo N = 999	IRD ¹ (95% CI)	HR (95% CI)
n (IR, per 100 PY)				
Mortality at 76 weeks	18 (1.41)	11 (0.83)	0.59 (-0.22, 1.41)	1.71(0.81, 3.61)
Trial AACI	17 (1.55)	10 (0.86)	0.69 (-0.21,1.60)	
Trial AACG*	1 (0.55)	1 (0.63)	-0.07 (-1.71, 1.56)	

Source: Reviewer's analysis using ADSL.XPT in SN120

*In AACG study, Donanemab = LY3002813-M + LY3002813-C = 131 + 15

Abbreviations: IR, incidence rate; IRD= incidence rate difference, PY, person-years; CI, confidence intervals; HR, hazard ratio.

¹ Incidence rate difference for both trials is calculated using the MH method.

² HR is calculated using the stratified Cox model (stratified by trial)

5. Summary and Conclusions

The assessment of mortality, primarily based on the placebo-controlled portion of Trial AACI, resulted in an estimated risk difference of 1.0% for death at week 76 with 95% CI of (-0.3%, 2.3%). This means that if 1000 patients from the population were treated with donanemab rather than placebo for 76 weeks, 10 additional patients would be expected to die. Estimates for mortality at 76 weeks + 57 days resulted in slightly increased estimate of the RD (1.1%). An integrated analysis that incorporates 76 week data from the Phase 2 trial, Trial AACG, resulted in similar findings: IRD = 0.59 with 95% CI (-0.22, 1.41).

The mortality assessment utilized an on-study approach to address the intercurrent event of treatment discontinuation. Such an approach considers participants at risk of death while exposed to randomized IP or not exposed to randomized IP but still being followed for vital status through trial end (i.e. 76 weeks or 76 weeks + 57 days). However, the conduct of Trial AACI, resulted in high trial discontinuation rates of 26% for donanemab and 20% for placebo and the sponsor did not make efforts to obtain vital status for subjects discontinuing from the trial.

Given the high percentage of subjects missing vital status information, the estimated effects from the mortality analysis would require the assumption that missing (i.e. censoring) is non-informative meaning that the reason a participant discontinues from the trial is unrelated to death. Based on the descriptive analyses shown in Section 3, such an assumption is likely implausible, especially with the high rate of study discontinuation and the reason for discontinuation being documented as due to an adverse event; both of which occur more frequently on donanemab than placebo.

Overall, the mortality analyses identified a potential safety signal of an increased risk of death - the point estimate from the analysis of Trial AACI suggests an increase of 1.0% in the risk of death on donanemab vs. placebo with the upper 95% CI bound failing to rule out differences as large as 2.3%. Note the trial was not powered to assess all-cause mortality and importantly, not powered to assess any cause-specific death. In addition, the conduct of Trial AACI further adds uncertainty to the observed mortality effect sizes and there is some evidence to suggest that additional vital status information from those that discontinued from the study may result in mortality effect sizes greater than those observed thereby suggesting increased levels of harm with donanemab relative to placebo. It's important that such uncertainty be incorporated into the benefit-risk assessment of donanemab along with the observed increased estimates of mortality risk of donanemab relative to placebo.

The following are some considerations on additional work beyond what is provided in this memo to further support the benefit-risk assessment of donanemab.

- Additional analyses are unlikely to further elucidate the impact of missing vital status information on the observed mortality findings. A concerted effort on behalf of the sponsor to recover vital status information, assuming such an effort is ethically/legally viable and can retrieve a high percentage of vital status information, would result in a more reliable estimate of the effect of donanemab on mortality.
- As the benefit-risk assessment of donanemab involves consideration of the therapeutic context, including the benefit-risk profile of available therapies, it is of interest to understand the safety findings, including mortality results, for other approved drugs (e.g., lecanemab) to understand if the observed mortality effect is unique to donanemab.

APPENDIX

A.1 Protocol Defined Information on Participant Discontinuation/Withdrawal from the Study

The information below is quoted from Section 7 of the protocol for Trial AACI. This suggests that by design, when a participant withdraws from the study no additional data (including vital status) is collected on that participant.

“A participant may withdraw from the study:

- at any time at his/her own request
- at the request of his/her designee (for example, parents or legal guardian)
- at the discretion of the investigator for safety, behavioral, compliance, or administrative reasons
- if the participant becomes pregnant during the study
- if the participant enrolls in any other clinical study involving an investigational medicinal product or enrollment in any other type of medical research judged not to be scientifically or medically compatible with this study
- if the participant, for any reason, requires treatment with another disease-modifying therapeutic agent that has been demonstrated to be effective for treatment of the study indication, discontinuation from the study occurs prior to introduction of the new agent
- participation in the study needs to be stopped for medical, safety, regulatory, or other reasons consistent with applicable laws, regulations, and good clinical practice (GCP).

At the time of discontinuing from the study, if possible, an early discontinuation (ED) visit should be conducted, as shown in the SoA. See SoA for data to be collected at the time of study discontinuation and follow-up and for any further evaluations that need to be completed. The participant will be permanently discontinued both from the IP and from the study at that time.

If the participant withdraws consent for disclosure of future information, the sponsor may retain and continue to use any data collected before such a withdrawal of consent. If a participant withdraws from the study, he/she may request destruction of any samples taken and not tested, and the investigator must document this in the site study records.”

A.2 Listing of Participants that Died and Included in Mortality Analysis

Table A.1 depicts the 29 participants (19 donanemab, 10 placebo) that died during the placebo-controlled portion of Trial AACI.



Table A.1. Listing of Participants that Died by Week 76 + 57 Days, Trial AACI (Safety Population)

USUBJID	Treatment Arm	Age	Sex	Days to Death	Cause of Death
15T-MC-AACI (b) (6)	Donanemab	69	F	3	Pulmonary Embolism
15T-MC-AACI	Donanemab	83	F	63	Death
15T-MC-AACI	Donanemab	72	M	75	ARIA-H
15T-MC-AACI	Donanemab	72	F	80	ARIA-E
15T-MC-AACI	Donanemab	68	M	81	Dehydration
15T-MC-AACI	Donanemab	82	M	85	Coronavirus COVID-19 Pneumonia
15T-MC-AACI	Donanemab	84	F	154	Death (Cause of death on Death Certificate is late onset Alzheimer's Disease)
15T-MC-AACI	Donanemab	62	M	175	completed suicide
15T-MC-AACI	Donanemab	79	F	215	PULMONARY EMBOLISM
15T-MC-AACI	Donanemab	85	F	223	pneumonia
15T-MC-AACI	Donanemab	83	M	244	SAH Subarachnoid Hemorrhage
15T-MC-AACI	Donanemab	82	F	273	Retroperitoneal Bleeding
15T-MC-AACI	Donanemab	75	M	314	Suicide
15T-MC-AACI	Donanemab	72	M	331	Respiratory Arrest
15T-MC-AACI	Donanemab	79	M	352	COVID-19
15T-MC-AACI	Donanemab	75	M	447	Death
15T-MC-AACI	Donanemab	75	F	514	The patient died of severe respiratory failure
15T-MC-AACI	Donanemab	81	F	536*	Natural Cause Death
15T-MC-AACI	Donanemab	75	F	584*	Cerebral Vascular Accident
15T-MC-AACI	Placebo	85	M	28	RESPIRATORY FAILURE
15T-MC-AACI	Placebo	81	M	139	lung infection
15T-MC-AACI	Placebo	85	F	147	Community Acquired Pneumonia
15T-MC-AACI	Placebo	83	F	176	Smoke Inhalation
15T-MC-AACI	Placebo	74	F	254	Heart Attack
15T-MC-AACI	Placebo	83	M	265	Death
15T-MC-AACI	Placebo	77	M	277	Suicide by carbon monoxide poisoning
15T-MC-AACI	Placebo	81	M	291	Sepsis
15T-MC-AACI	Placebo	65	F	351	Probable Atherosclerotic Cardiovascular Disease
15T-MC-AACI	Placebo	78	M	520	Worsening of AD
15T-MC-AACI	Placebo	76	M	534*	Lung neoplasm malignant

Source: Reviewer's analysis using ADSL.XPT in SN120

* The participant was not recorded as a death in the ADSL.XPT data set but recorded as a death in ADAE.XPT along with narrative confirming the death of the participant.

† Death included only in the 76 weeks + 57 day analysis

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

/s/

FENGYU ZHAO
03/04/2024 12:26:41 PM

PING LI
03/04/2024 12:55:26 PM

YUEQIN ZHAO
03/04/2024 01:27:12 PM

MATTHEW J SOUKUP
03/04/2024 01:38:20 PM



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

BLA #: BLA 761248/SDN 100 (resubmission)
Drug Name: Kisunla (donanemab-azbt)
Indication(s): Alzheimer's disease
Applicant: Eli Lilly and Company
Date(s): Resubmission Date: 6/12/2023
Primary Review Due Date: 11/19/2023
PDUFA Date: March/12/2024
Review Priority: Resubmission – Class 2

Biometrics Division: DBI
Statistical Reviewer: Van Tran PhD
Statistical Analyst: Xiaofeng (Tina) Wang MPH
Concurring Reviewers: John P. Lawrence PhD (secondary), James Hung PhD (tertiary)

Medical Division: Division of Neurology I
Clinical Team: Efficacy: Kevin Krudys PhD, Ranjit B. Mani MD
Safety: Natalie Branagan MD, Sally Yasuda PharmD
Project Manager: Justine Kankam

Keywords:

Link to keywords: Early Alzheimer's Disease, MMRM, natural cubic spline, iADRS, CDR-SB, amyloid plaque, Time-PMRM, tipping point analysis, missing data

Table of Contents

1	EXECUTIVE SUMMARY	5
2	INTRODUCTION	6
2.1	OVERVIEW.....	7
2.2	DATA SOURCES	8
3	STATISTICAL EVALUATION	8
3.1	DATA AND ANALYSIS QUALITY	8
3.2	EVALUATION OF EFFICACY	9
3.2.1	<i>Study Design and Endpoints</i>	9
3.2.2	<i>Statistical Methodologies</i>	14
3.2.3	<i>Patient Disposition, Demographic and Baseline Characteristics</i>	21
3.2.4	<i>Results and Conclusions</i>	26
3.3	EVALUATION OF SAFETY	37
4	FINDINGS IN SPECIAL/SUBGROUP POPULATIONS	37
5	SUMMARY AND CONCLUSIONS	39
5.1	STATISTICAL ISSUES	39
5.2	COLLECTIVE EVIDENCE	39
5.3	CONCLUSIONS AND RECOMMENDATIONS	40
5.4	LABELING RECOMMENDATIONS	40
6	APPENDICES.....	41
6.1	NATURAL CUBIC SPLINE	41
6.2	RANK ANCOVA.....	42
6.3	RESULTS.....	43
6.4	MODEL DIAGNOSTICS.....	48
6.5	NCS2 MODEL FIT ESTIMATES.....	51

LIST OF TABLES

Table 1 List of all studies included in analysis	8
Table 2 Demographics/Baseline Characteristics	24
Table 3 Baseline Disease Scales and Biomarkers.....	25
Table 4 Primary Efficacy Analysis (NCS2) of iADRS Change from Baseline at Week 76 in Evaluable Efficacy Set by Study Population.....	26
Table 5 Tipping Point Analysis of CDR-SB Change from Baseline at Week 76 in Evaluable Efficacy Set, Intermediate Tau Population.....	36
Table 6 Tipping Point Analysis of CDR-SB Change from Baseline at Week 76 in Evaluable Efficacy Set, Overall Population.....	36
Table 7 Analysis of Select Endpoints Change from Baseline at Week 76 in Evaluable Efficacy Set by Study Population.....	44
Table 8 Analysis (Death Imputed as the Worst Score) of Clinical Endpoints at Week 76 in Evaluable Efficacy Set by Study Population	45
Table 9 Tipping Point Analysis of iADRS Change from Baseline at Week 76 in Evaluable Efficacy Set, Intermediate Tau Population.....	46
Table 10 Tipping Point Analysis of iADRS Change from Baseline at Week 76 in Evaluable Efficacy Set, Overall Population.....	46
Table 11 Covariate Estimates - Primary Efficacy Analysis (NCS2) of iADRS Change from Baseline at Week 76 in Evaluable Efficacy Set, Overall Population.....	52
Table 12 Covariance Matrix Estimate - Primary Efficacy Analysis (NCS2) of iADRS Change from Baseline at Week 76 in Evaluable Efficacy Set for the Overall Population.....	56
Table 13 Covariance Matrix Estimate - NCS2 Analysis of iADRS Change from Baseline at Week 76 Including Baseline Score Covariates in Evaluable Efficacy Set, Overall Population.....	57

LIST OF FIGURES

Figure 1 Study Schema.....	10
Figure 2 Hypothesis Testing Scheme to Control Type I Error Rate at Two-sided Alpha Level of 0.05	19
Figure 3 Subject Disposition	22
Figure 4 Time to Study Discontinuation.....	23
Figure 5 Time to Treatment Discontinuation.....	23
Figure 6 Primary Efficacy Analysis (NCS2) of iADRS Change from Baseline with Observed Mean Change from Baseline in Evaluable Efficacy Set by Study Population	27
Figure 7 Analysis of Select Clinical Endpoints Change from Baseline at Week 76 in Evaluable Efficacy Set by Study Population	29
Figure 8 Kaplan-Meier Plot: Cumulative Incidence of Time to Substantial Decline in CDR-G by Week 76 in Evaluable Efficacy Set, Intermediate Tau Population	30
Figure 9 Kaplan-Meier Plot: Cumulative Incidence of Time to Substantial Decline in CDR-G by Week 76 in Evaluable Efficacy Set, Overall Population.....	31
Figure 10 Primary (NCS2) and Sensitivity Efficacy Analysis of iADRS Change from Baseline at Week 76 in Evaluable Efficacy Set by Study Population	32
Figure 11 Analysis of Amyloid Change from Baseline at Week 76 in Evaluable Efficacy Set by Study Population.....	37
Figure 12 Subgroup Analysis (NCS2) of iADRS Change from Baseline at Week 76 in Evaluable Efficacy Set, Overall Population.....	38
Figure 13 Natural Cubic Spline vs. Mixed Model for Repeated Measures Analysis of iADRS Change from Baseline in Evaluable Efficacy Set by Study Population	43
Figure 14 Subgroup Analysis (MMRM) of CDR-SB Change from Baseline at Week 76 in Evaluable Efficacy Set, Overall Population.....	47
Figure 15 Subjects' Score Change and Estimated Mean Change from the Primary Efficacy Analysis (NCS2) of iADRS Change from Baseline at Week 76 in Evaluable Efficacy Set, Overall Population	49
Figure 16 Residual Quantile-Quantile Plot from Primary Efficacy Analysis (NCS2) of iADRS Change from Baseline at Week 76 in Evaluable Efficacy Set, Overall Population.....	50
Figure 17 Residuals versus Fitted Values by Scheduled Visit from the Primary Efficacy Analysis (NCS2) of iADRS Change from Baseline at Week 76 in Evaluable Efficacy Set, Overall Population.....	51

1 EXECUTIVE SUMMARY

Kisunla (donanemab-azbt) is an antibody directed at amyloid beta present only in brain amyloid plaques and intended for the treatment of Alzheimer's disease (AD). On June 12, 2023, Sponsor resubmitted the application to contain safety and efficacy data from the Phase 3 study AACI to address a complete response letter and apply for traditional approval.

Study AACI was a multicenter, randomized, double-blind, placebo-controlled study of participants with early symptomatic AD. Participants who met entry criteria were randomized 1:1 to donanemab or a matching placebo. The primary objective was to test the hypothesis that donanemab slowed the cognitive or functional decline of AD as measured by iADRS (integrated Alzheimer's Disease Rating Scale) compared to placebo in the population of participants with intermediate tau pathology at baseline and the overall population. The Division had not agreed to iADRS as the primary endpoint and had advised instead CDR-SB (Clinical Dementia Rating Scale – Sum of Boxes) as one of the alternatives. A natural cubic spline analysis with 2 degrees of freedom was pre-specified to assess the difference between treatment groups.

Study AACI showed that the donanemab arm compared to the placebo arm had statistically significantly less decline at Week 76 as measured by iADRS: intermediate tau population (3.3; 95% CI 1.9, 4.6; $p < 0.001$); overall population (2.9; 95% CI 1.5, 4.3; $p < 0.001$). Compared to the range of the scale (0 to 141), the treatment difference of approximately 3 points was small but statistically significant. The donanemab arm compared to the placebo arm had statistically significantly less decline at Week 76 as measured by the secondary endpoint CDR-SB: intermediate tau population (-0.7; 95% CI -1.0, -0.4; $p < 0.001$); overall population (-0.7; 95% CI -1.0, -0.5; $p < 0.001$). The treatment effect was also small (CDR-SB ranged 0 to 18) but statistically significant.

Sensitivity analyses of iADRS using mixed model with repeated measures, a model with less assumptions and more commonly used as the primary analysis of longitudinal AD data, produced a statistically significant treatment effect. A non-parametric rank analysis of covariance without imposition of parametric model assumption also produced a statistically significant treatment effect.

Analysis of iADRS and CDR-SB with the unfavorable outcome of death imputed as the worst score produced a similar and statistically significant treatment effect compared to its primary analysis. However, the analysis with death imputed as worst score for ADAS-Cog13 (Alzheimer's Disease Assessment Scale – Cognitive subscale) change from baseline at Week 76 and time to substantial decline in CDR-G (CDR global score) did not achieve statistical significance when compared to p-value cutoffs that were multiplicity controlled.

Study discontinuation was high with the donanemab arm having a higher rate compared to the placebo arm (27% vs. 20%). The tipping point analysis to assess sensitivity of the CDR-SB primary analysis to missing data assumptions showed that the results in both the intermediate tau and overall population were not statistically significant near the delta pair (0, 0.5), where 0.5 is the minimum within-subject scale increment change.

Despite the potential for the primary analysis of CDR-SB to be sensitive to missing data assumptions, study AACI demonstrated evidence of effectiveness. The Phase 2 Study AACG provided confirmatory evidence of effectiveness with the primary endpoint of iADRS (3.2; 95% CI 0.1, 6.3; p-value = 0.04). Overall, we recommend approval of donanemab.

For future studies of AD, we recommend that the Division provide specific protocol recommendations for sponsors to proactively follow up study participants to the end of the trial to minimize missing data. High missing data rates can result in efficacy and safety estimates that are unreliable for decision making. Additionally, pre-specifying a clinically relevant cutoff for a tipping point analysis would allow for a structured assessment of plausible missing data assumptions.

2 INTRODUCTION

Kisunla (donanemab-azbt) is an antibody directed at the pyroglutamate modification of the third amino acid of amyloid beta epitope that is present only in brain amyloid plaques. Donanemab binds to deposited amyloid plaque and aids its removal through microglial-mediated phagocytosis. Accumulation of amyloid plaque has been proposed to be the primary driver of the Alzheimer's disease (AD) process. Other anti-amyloid beta-directed antibodies included aducanumab and lecanemab that received accelerated approval for the treatment of AD (specifically for patients with mild cognitive impairment or mild dementia stage of disease) based on the observed reduction of amyloid beta, a surrogate endpoint. Lecanemab received full approval on July 6, 2023, based on significantly less decline as measured by Clinical Dementia Rating Scale – Sum of Boxes (CDR-SB).

Sponsor applied for accelerated approval for donanemab on May 18, 2022, based on evidence of amyloid plaque reduction from the Phase 2 study AACG. The Division issued a complete response letter (CRL), dated January 18, 2023, because the safety database was insufficient to adequately characterize the long-term safety of donanemab in the treatment of AD. On June 12, 2023, Sponsor resubmitted the application to contain safety and efficacy data from the Phase 3 study AACI to address the CRL and apply for traditional approval.

While study AACI was being conducted, the Division disagreed with Sponsor's proposal to change the primary endpoint from CDR-SB to the integrated Alzheimer's Disease Rating Scale (iADRS), because there was uncertainty that iADRS could assess the effects of treatment on both cognition and function, independent of each other¹. The Division advised the use of CDR-SB as the primary efficacy endpoint or the 13-item Alzheimer's Disease Assessment Scale – Cognitive (ADAS-Cog13) and the Alzheimer's Disease Cooperative Study – Instrumental Activities of Daily Living Scale (ADCS-iADL) as co-primary efficacy measures. However, Sponsor proceeded with iADRS as the primary endpoint based on the assertions that iADRS showed more consistent effects across trials, more reliable within-subject change, a smaller study sample size, shorter trial duration, and lower risk of a failed study.

¹ March 2, 2021, Type C Meeting Minutes. Available in DAARTS.

Additionally, there were multiple exchanges between Sponsor and the Division on the choice of primary statistical analysis model for study AACI. MMRM had typically been used in AD studies, including study AACG, as the primary analysis model. However, for study AACI, Sponsor proposed natural cubic splines with 2 degrees of freedom (NCS2) based on analyses that showed NCS2 attaining higher statistical power compared to MMRM. After Sponsor responded to the Division's concerns about bias and Type I error inflation², the Division agreed to the NCS2 model as the primary analysis model.

2.1 Overview

There were two studies, Phase 2 study I5T-MC-AACG (referred to as AACG) and Phase 3 study I5T-MC-AACI (referred to as AACI), relevant to Sponsor's application for full approval. Study AACG was previously reviewed by Massie 2023³ in Sponsor's application for accelerated approval. The focus of this review is on the efficacy results for the double-blind period of study AACI, referred to as I5T-MC-AACI-PC. See study summary in Table 1. The first subject was screened on June 19, 2020, and the last subject was enrolled on November 17, 2021. The data cut-off date was April 14, 2023, and the database lock date was April 28, 2023. Safety results for both studies are covered by the Clinical Team's safety review.

Note that study AACI has an ongoing long-term double-blind extension to further evaluate donanemab efficacy and safety. The analyses and results of this extension period will be submitted by Sponsor in a future report.

² December 22, 2022, Type B meeting minutes and briefing document. Written response only, dated February 22, 2023. Available in DARRTS.

³ Massie, T. Statistical Review and Evaluation of Kisunla (Donanemab) - Biomarker Accelerated Approval. January 17, 2023. Available in DARRTS.

Table 1 List of all studies included in analysis

Applicant defined study number	Phase and Design	Treatment Period	Follow-up Period	# of Subjects per Arm	Study Population
I5T-MC-AACG	Phase 2, MC, R, PG, DB, PC Donanemab: 700 mg IV Q4W for the first 3 doses and then 1400 mg IV Q4W Placebo: placebo IV Q4W ^a	72 weeks	4 weeks (after treatment period)	Placebo: 126 Donanemab:131 Donanemab + BACE inhibitor: 15 (arm discontinued)	Early symptomatic AD with elevated amyloid and intermediate tau burden
I5T-MC-AACI-PC	Phase 3, MC, R, PG, DB, PC Donanemab: 700 mg IV Q4W for the first 3 doses and then 1400 mg IV Q4W Placebo: placebo IV Q4W	72 weeks	4 weeks (after treatment period)	Placebo:876 Donanemab:860	Early symptomatic AD with the presence of brain amyloid and tau pathology (intermediate and high)

AD: Alzheimer's Disease, DB: double-blind, IV: intravenous, MC: multi-center, PC: placebo controlled, PG: parallel group, Q4W: every 4 weeks, R: randomized

^a Fifteen patients were randomized to a donanemab/BACE inhibitor, but the combination therapy was discontinued from the study.

2.2 Data Sources

The BLA resubmission is located in <\\CDSESUB1\evsprod\BLA761248\0104>. The submission contained the clinical study report, STDM data, ADAM data, and programming code for study I5T-MC-AACI-PC. The dataset containing the primary endpoint is located in <\\CDSESUB1\evsprod\BLA761248\0104\m5\datasets\i5t-mc-aaci\analysis\adam\datasets\adcogadl.xpt>.

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

The quality of the submitted data and analyses was adequate. The statistical analyst was able to verify with independent coding in SAS and reference to Sponsor's SAS programs pre-specified analyses that included:

- Primary analysis, sensitivity analyses, and subgroup analyses of the primary endpoint of iADRS
- Analyses of other endpoints (CDR-SB, ADCS-iADL, ADAS-Cog13, CDR-G, amyloid centiloid) (b) (4).

When using R software, the NCS2 model of iADRS did not converge with an unstructured variance-covariance structure matrix. As stated in the protocol, if the unstructured covariance matrix did not converge, a heterogeneous Toeplitz covariance structure would be fitted, but heterogeneous Toeplitz was not an option in R software. NCS2 did converge with heterogeneous autoregressive order 1 covariance structure in R software, however, an unstructured covariance was preferred because it had less assumptions. The NCS2 fit with a different covariance structure using R versus SAS software was considered a minor concern.

3.2 Evaluation of Efficacy

3.2.1 Study Design and Endpoints

3.2.1.1 Design

Overall Design

Study AACI is a multicenter, randomized, double-blind, placebo-controlled, Phase 3 study of participants with early symptomatic AD. Participants who met entry criteria were randomized 1:1 to the following treatment groups:

- Donanemab: 700 mg administered by intravenous (IV) infusion once every four weeks (Q4W) for the first 3 doses, then 1400 mg IV Q4W
- Placebo: Q4W

Randomization was stratified by intermediate or high tau level and by study sites. Participants received up to 72 weeks of treatment. Participants had planned assessments every 4 weeks after randomization with final endpoint measures and safety assessments performed 4 weeks following the last dose (Figure 1). Scales measuring efficacy endpoints were collected at Baseline, Week 12, Week 24, Week 36, Week 52, Week 64, and Week 76. After Week 76, participants were assigned to donanemab or placebo in the long-term extension (LTE) phase. This efficacy review covers up to 76 weeks (double blind period) of study AACI.

Participants whose amyloid plaque reduction (measured by florbetapir F18 positron emission tomography (PET) scans) met dose reduction criteria had a double-blind dose reduction of donanemab to IV placebo for the remaining duration of the study. These dose reduction rules were defined as: amyloid plaque level was <11 centiloid (CL) at any single amyloid PET scan or $11 \leq \text{CL} < 25$ from two consecutive amyloid PET scans.

	Screening Period	Double-Blind Period																				Extension Period		Follow-up Period		
Visit	601	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	V22 through V41		801-804	
Week		-7	0	4	8	12	16	20	24	28	32	36	40	44	48	52	56	60	64	68	72	76	W78 through W154		c	
		Donanemab 700 mg IV Q4W x 3 doses		Donanemab 1400 mg IV Q4W																				Dose reduction criteria not met Donanemab ^{a,b}		
				Placebo ^a																				All Placebo		
																								All		

Source: Study Protocol Amendment e, Study Schema

The primary estimand was the treatment policy strategy⁴, i.e., the intervention difference in slowing of progression of AD relative to placebo across 76 weeks of intervention in participants with early symptomatic AD, regardless of initiation or change to standard of care medications and regardless of whether a participant stopped taking study drug.

- 60 to 85 years of age, inclusive
- Gradual and progressive change in memory function reported by the participant or informant for ≥ 6 months
- A Mini-Mental State Examination (MMSE) score of 20 to 28 (inclusive) at Visit 601 or 1
- Patient met flortaucipir F18 PET scan criteria for evidence of pathologic tau deposition
- Patient met florbetapir F18 PET or florbetaben F18 PET scan criteria

Reference ID: 5275619

The main exclusion criteria were:

- Patient had any contraindications for MRI or PET
- Patient had centrally read MRI demonstrating presence of ARIA-E, greater than 4 cerebral microhemorrhages, more than 1 area of superficial siderosis, any macrohemorrhage or severe white matter disease at screening
- Patient was sensitive to florbetapir F18, florbetaben F18, flortaucipir F18

There were two study populations of interest:

- Intermediate tau pathology population: participants with baseline
 - Standardized Uptake Value ratio (SUVr) ≤ 1.46 and a topographic deposition pattern consistent with advanced AD (AD++) or
 - $1.10 \leq \text{SUVr} \leq 1.46$ and a topographic deposition pattern consistent with moderate AD (AD+)
- The overall population comprised all randomized participants, i.e., both intermediate and high tau⁵ population.

Treatment Discontinuation

The investigational product (IP), i.e., control or donanemab, could be discontinued for a participant for the following reasons:

- By request of the participant or participant's designee (for example, legal guardian)
- Clinical judgment
- A hepatic event or liver test abnormality
- Columbia Suicide Severity Rating Scale (C-SSRS) results
- Electrocardiogram (ECG) finding
- Amyloid-related imaging abnormalities (ARIA)
- Systemic hypersensitivity reaction
- Severe noncompliance
- The participant requires an excluded therapeutic agent

If IP were permanently discontinued, the participant would remain in the study to be evaluated for safety and efficacy according to the schedule of activities. Temporary discontinuation of IP was allowed for reasons such as development of ARIA, short-term treatment with an excluded medication, hospitalization, personal or exceptional circumstances, or to evaluate the IP impact on an uncertain adverse event.

Study Discontinuation

A participant could withdraw from the study:

- at any time at his/her own request
- at the request of his/her designee (for example, parents or legal guardian)
- at the discretion of the investigator for safety, behavioral, compliance, or administrative reasons

⁵ High tau pathology population: participants with baseline SUVr > 1.46 and a topographic deposition pattern consistent with either moderate (AD+) or advanced AD (AD++).

- if the participant became pregnant during the study
- if the participant enrolled in any other clinical study involving an investigational medicinal product or enrollment in any other type of medical research judged not to be scientifically or medically compatible with this study
- if the participant, for any reason, required treatment with another disease-modifying therapeutic agent that had been demonstrated to be effective for treatment of the study indication; discontinuation from the study occurred prior to introduction of the new agent
- participation in the study needed to be stopped for medical, safety, regulatory, or other reasons consistent with applicable laws, regulations, and good clinical practice

At the time of discontinuing from the study, if possible, an early discontinuation visit would be conducted. The participant would be permanently discontinued both from the IP and from the study at that time.

Reviewer Comment: Although the pre-specified estimand was the treatment policy strategy, the study protocol (Amendment e) contained conflicting language that allowed study subjects to discontinue the study for some reasons that appeared to be intercurrent events. This study conduct resulted in a high rate of study discontinuation (23%) and missing data (see Section 3.2.3 Patient Disposition, Demographic and Baseline Characteristics, Section 3.2.4.1 Primary Analysis of iADRS).

3.2.1.2 Endpoints

The Primary Objective: To assess the effect of donanemab versus placebo on clinical progression in participants with early symptomatic AD

Primary Efficacy Endpoint: iADRS (integrated Alzheimer’s Disease Rating Scale) change from baseline at Week 76. The iADRS score was calculated from the Alzheimer’s Disease Cooperative Study – instrumental Activities of Daily Living (ADCS-iADL) and Alzheimer’s Disease Assessment Scale – Cognitive subscale (ADAS-Cog13) scores. iADRS assesses the impact of cognitive loss on the ability to conduct everyday activities and provides a measure of global AD severity as a single summary score. The total score ranged from 0 to 141, with lower scores reflecting worse performance.

Reviewer Comment: The Division had disagreed with Sponsor’s use of iADRS as the primary endpoint. See Section 2 INTRODUCTION for the regulatory background.

Secondary endpoints:

- CDR-SB (Clinical Dementia Rating Scale – Sum of Boxes) change from baseline at Week 76. The range for CDR-SB score was 0 to 18 with higher scores indicating a greater level of impairment.
- ADCS-iADL change from baseline at Week 76. The range for the iADL score was 0 to 59 with lower scores indicating greater disease severity.
- ADAS-Cog13 change from baseline at Week 76. The range for ADAS-Cog13 score was 0 to 85 with higher scores indicating greater disease severity.

- MMSE (Mini Mental State Examination) change from baseline at Week 76. The range for MMSE score was 0 to 30 with lower scores indicating a greater level of impairment.

Endpoints tested with multiplicity control:

- CDR-SB change from baseline at Week 76
- ADCS-iADL change from baseline at Week 76
- ADAS-Cog13 change from baseline at Week 76
- Time to substantial decline in CDR global score (CDR-G) by Week 76. Any increase in CDR-G from baseline for two consecutive visits met the criterion of substantial decline. The range of CDR-G score was 0 to 3 with higher scores indicating greater level of impairment.
- CDR-SB non-progressing status (binary) at Week 52. Participants' status was classified as non-progressing if their CDR-SB change from baseline was less than or equal to 0, which was calculated as a binary outcome at each of the scheduled visits.
- Amyloid centiloid (CL) change from baseline at Week 76
- Amyloid complete clearance (CCL) (centiloid < 24.1) status (binary) at Week 24
- Amyloid CCL status (binary) at Week 76
- Brain tau deposition (P-tau217) change from baseline at Week 24
- Brain tau deposition (P-tau217) change from baseline at Week 76
- Tau position emission tomography (PET) frontal standard uptake value ratio (SUVR) change from baseline at Week 76

Safety Endpoints:

- Spontaneously reported adverse events (AEs)
- Clinical laboratory tests
- Vital sign and body weight measurements
- ECGs
- Physical and neurological examinations
- Magnetic resonance imaging (MRI) ARIA and emergent radiological findings
- Infusion related reactions
- C-SSRS

3.2.1.3 Analysis Sets

For efficacy analyses, the primary analysis set was the Evaluable Efficacy analysis set (EES), defined as all randomized participants with a baseline and at least one post-baseline efficacy scale. Participants were summarized according to the treatment group to which they were randomized.

For safety analyses, the Safety analysis set was defined as all randomized participants who were exposed to the study drug. Participants were summarized according to the treatment group to which they were randomized.

3.2.2 Statistical Methodologies

3.2.2.1 Primary Analysis of iADRS

The primary objective was to test the hypothesis that donanemab slowed the cognitive or functional decline of AD as measured by the iADRS score compared with placebo in the population of participants with intermediate tau pathology at baseline and the overall population.

A natural cubic spline analysis (see Appendix 6.1 Natural Cubic Spline) with 2 degrees of freedom (NCS2) was pre-specified to assess the difference between treatment groups in iADRS change from baseline at Week 76. For the NCS2 model, 3 knots over the observation time were placed: 2 at the boundaries (minimum and maximum observation time) and 1 internal knot at the median observation time. The baseline estimates were restricted to be the same for treatment and placebo groups. The model was estimated using restricted maximum likelihood method. The iADRS score at baseline and at each of the scheduled post-baseline visits were included in model as a dependent variable. Study visit was assumed continuous with values equal to weeks between baseline and post-baseline exam dates. NCS basis functions were constructed using visits in weeks. The model included the fixed effects: NCS basis expansion terms (two terms), NCS basis expansion term-by-treatment interaction (two terms), baseline age, concomitant acetylcholinesterase inhibitor (AChEI) and/or memantine use at baseline (yes/no), and pooled investigator⁶. Baseline tau category was included as a covariate in the model for the overall population. An unstructured variance-covariance structure matrix was assumed for the within-subject variance-covariance errors. If the unstructured variance-covariance structure matrix resulted in a lack of convergence, the following structures were used in sequence: heterogeneous Toeplitz covariance structure, heterogeneous autoregressive order 1 covariance structure, heterogeneous compound symmetry covariance structure, and compound symmetry covariance structure. The Kenward-Roger approximation was used to estimate the denominator degrees of freedom. The treatment group contrast in mean change from baseline at Week 76 with 95% CI was estimated using NCS2. NCS2 assumed missing data were missing at random. Measurements after death were treated as missing.

Percent slowing compared to the placebo group was calculated as: the difference of LS estimates of mean change from baseline between treatment groups at Week 76 divided by the LS estimate of mean change from baseline value of the placebo group.

Reviewer Comment:

- NCS2 is a generalized least squares model of longitudinal data that treated time as a continuous variable in the mean response. NCS2 modeled the temporal mean trend with polynomials, specifically with cubic polynomials, that were spliced together at knot locations. In contrast, mixed model with repeated measures (MMRM) treated time as a categorical variable and did not assume a trend over time. The MMRM had typically been the primary analysis model of longitudinal data in AD studies. However, Sponsor proposed NCS2 as the primary analysis model on the basis that NCS2 attained higher statistical power compared to MMRM (see Section 2 INTRODUCTION for the regulatory

⁶ Investigator sites with <2 subjects were pooled together based on enrollment country.

background). MMRM analysis of iADRS became a sensitivity analysis (Section 3.2.2.3 Sensitivity Analyses).

- Ignoring deaths, a negative outcome, could bias the results. Although the Division had advised Sponsor to account for death in the primary efficacy analysis⁷ prior to the BLA submission, Sponsor instead assumed death was not informative and measurements after death were treated as missing. Sponsor instead pre-specified a sensitivity analysis imputing death as the worst outcome in the NCS2 analysis of iADRS (Section 3.2.2.3 Sensitivity Analyses).

3.2.2.2 Analysis of Select Clinical Endpoints

The following were pre-specified analyses of select clinical endpoints (b) (4).

CDR-SB

Mixed model with repeated measures (MMRM) was fitted to CDR-SB. The change from baseline score in CDR-SB at each scheduled post-baseline visit during the treatment period was the dependent variable. The model included fixed effects terms: treatment, visit, treatment-by-visit interaction, baseline CDR-SB score, baseline score-by-visit interaction, age at baseline, concomitant AChEI and/or memantine use at baseline (yes/no), and pooled investigator. Baseline tau category was included as a fixed effect in the model for the overall population. Visit was treated as a categorical variable. An unstructured covariance matrix was used to model the within-subject variance-covariance errors⁸. The Kenward-Roger approximation was to estimate the denominator degrees of freedom. For MMRM, the primary time point for treatment comparison was Week 76. The treatment group contrast in mean change from baseline with 95% CI was calculated using the MMRM model. MMRM assumed missing data were missing at random. Measurements after death were treated as missing.

iADL, ADAS-Cog13

NCS2, similar to the primary analysis model of iADRS, was used to estimate ADCS-iADL and ADAS-Cog13 change from baseline at Week 76.

iADRS, CDR-SB

Time progression model for the repeated measures (Time-PMRM) (Raket 2022⁹) was used to estimate the slowing, i.e., time saved, of the time progression of the disease from donanemab treatment as compared to the time progression in the placebo group at Week 76 for the endpoints iADRS and CDR-SB. For this analysis, baseline and post-baseline endpoint measurements at the scheduled visits were used as dependent variables, and the model included as covariates the baseline variables: age, concomitant AChEI or memantine use, and pooled investigator. Baseline

⁷ December 22, 2022, Type B meeting minutes. Available in DARRTS.

⁸ If the unstructured variance-covariance structure matrix resulted in a lack of convergence, the following structures would be used in sequence: heterogeneous Toeplitz covariance structure, heterogeneous autoregressive order 1 covariance structure, heterogeneous compound symmetry covariance structure, and compound symmetry covariance structure.

⁹ Raket LL. Progression models for repeated measures: estimating novel treatment effects in progressive diseases. Stat Med. 2022;41(28):5537-5557. <https://doi.org/10.1002/sim.9581>

tau category was included as a covariate in the model for the overall population. Planned visit in weeks from randomization was included as a continuous variable. The intercepts were constrained to be the same between treatment arms. A natural cubic spline model with internal knots at each planned visit was used to interpolate the disease progression between the planned visits for the placebo arm. An unstructured covariance matrix was assumed for the within-subject variance-covariance errors. The donanemab treatment trajectory was estimated assuming the mean disease progression of the treatment group at a given visit can be estimated by the mean disease progression of the placebo group at another time point. Model parameters were estimated using maximum likelihood estimation, and significance testing was done using likelihood ratio tests. The assumption of proportional time slowing was tested, and if the assumption was not met, a model without the proportionality assumption was fitted, estimating a time slowing parameter at each post-baseline visit.

Reviewer Comment: We identified several concerns with the use of Time-PMRM for the analysis and interpretation of iADRS and CDR-SB. As a result, we did not present the results of the Time-PMRM model. We identified the following issues with Time-PMRM:

- Time-PMRM did not define disease progression at the patient level.
- Time-PMRM assumed the mean outcome of the treatment group at a given post-baseline visit can be described as the mean outcome of the placebo group at an earlier visit. This assumption would not hold if the treatment resulted in a worse outcome (harm) than placebo.
- Time-PMRM did not account for death.
- Time-PMRM analyses might have bias and Type I error inflation.
- Interpretation of a time delay would be challenging if there were inconsistent time delay estimates for different endpoints.
- We are unclear that Time-PMRM can reliably estimate time delay using outcome scales iADRS and CDR-SB that did not directly measure time delay.

Time to substantial decline in CDR-G

A Cox proportional hazards (CPH) model estimated the hazard ratio (95% CI) of progressing to substantial decline in CDR-G by Week 76. The analysis modeled time to first occurrence of the event and adjusted for the baseline variables: age, score, and concomitant AChEI and/or memantine use. The model was stratified by pooled investigator sites. For the analysis of overall population, the model was also stratified by the baseline tau category. Ties were handled using the discrete method. Censoring events included death, study discontinuation, and end of study.

3.2.2.3 Sensitivity Analyses

MMRM

The MMRM analysis of iADRS was similar to the MMRM analysis specified for CDR-SB. See Section 3.2.2.2 Analysis of Select Clinical Endpoints.

NCS with 3 degrees of freedom (NCS3)

NCS with three degrees of freedom model (NCS3) was fitted to iADRS. This model assumed two internal knots at equidistant percentiles of the scheduled study visit times. The model had the same set of covariates as described for NCS2, with the exception that three (instead of two) basis functions were included in the model. Similar to NCS2, data were assumed to be missing at random. Measurements after death were treated as missing.

With an additional internal knot, NCS3 can provide a more flexible fit to the data compared to NCS2. NCS3 had a lower Akaike information criterion, which indicated a better fit, but a higher Bayesian information criterion, which indicated a worse fit, when fitted to the Phase 2 study AACG data¹⁰. In a simulation study conducted by Sponsor¹¹, NCS3 showed less bias compared to NCS2 under simulation scenarios using data generated with an MMRM model. FDA requested NCS3 analysis to allow for an assessment of model fit under a different spline assumption compared to NCS2.

ARIA-E/IRR As Censoring Events

The occurrence of Amyloid Related Imaging Abnormality – Edema (ARIA-E) and infusion related reaction (IRR) might potentially lead to functional unblinding of the study treatment. For the sensitivity analysis, iADRS measurements were censored after the first occurrence of ARIA-E or IRR. NCS2 was fitted to this censored dataset. Measurements after death were treated as missing.

Death Imputed As Worst Score

The pre-specified sensitivity analysis imputed the worst possible iADRS score 0 for measurements after death until Week 76 for death cases. NCS2 was then fitted to the data.

During the BLA review, we requested Sponsor to conduct a sensitivity analysis imputing the worst score for measurements after death until Week 76 in the primary analysis of ADAS-Cog13, ADCS-iADL, and time to substantial decline in CDR-G. The statistical analyst conducted a similar sensitivity analysis for CDR-SB.

Rank ANCOVA with Multiply Imputed Data at Week 76

During the BLA review, we requested Sponsor to conduct a non-parametric rank analysis of covariance (ANCOVA) with multiply imputed (MI) data at Week 76 for all randomized subjects in each target population (intermediate tau, overall) to address missing data (assuming data were missing at random) and the potential departure from residual error normality in the primary

¹⁰ December 22, 2022, Type B meeting briefing document. Available in DARRTS.

¹¹ Ibid.

analysis of iADRS and CDR-SB. A rank ANCOVA could be more robust to departures from normal residual errors. See Appendix Section 6.2 Rank ANCOVA for more details.

Tipping Point Analyses

During the BLA review, we requested Sponsor to conduct a tipping point analysis to explore the plausibility of missing data assumptions under which the conclusions change, i.e., under which there is no longer a statistically significant effect shown for the endpoints of iADRS and CDR-SB change from baseline at Week 76 in the intermediate tau and overall population. Data were imputed using the fitted model as was done for the rank ANCOVA (see Appendix Section 6.3 Rank ANCOVA). Decreases on the iADRS scale represented worsening, so a range of delta values from 0 to -5 was explored for donanemab participants in increments of 1. For participants on placebo, delta ranged from 0 to -2 in increments of 1. Increases on CDR-SB represented worsening, so a range of delta values from 0 to 3 in increments of 0.5 was used for donanemab participants and 0 to 1.5 in increments of 0.5 for placebo participants. Delta values were added to each subject's imputed change from baseline. For a subject with imputed data, the delta value was a worse shift compared to observed subjects. Then the rank ANCOVA was conducted on the modified data and Cochran-Mantel-Haenszel test statistics were combined using Rubin's rule. A p-value was provided for each pair of delta values. The sensitivity of the primary analysis result would depend on the plausibility of the missing data assumptions.

The graphical multiplicity testing plan allocated $\alpha = 0.04$ for tests in the intermediate tau population and $\alpha = 0.01$ for tests in the overall population. CDR-SB inherits all the alpha from iADRS in the respective population (see Figure 2). These alpha values were used in the assessment of the tipping point analyses.

3.2.2.4 Analysis of Amyloid PET

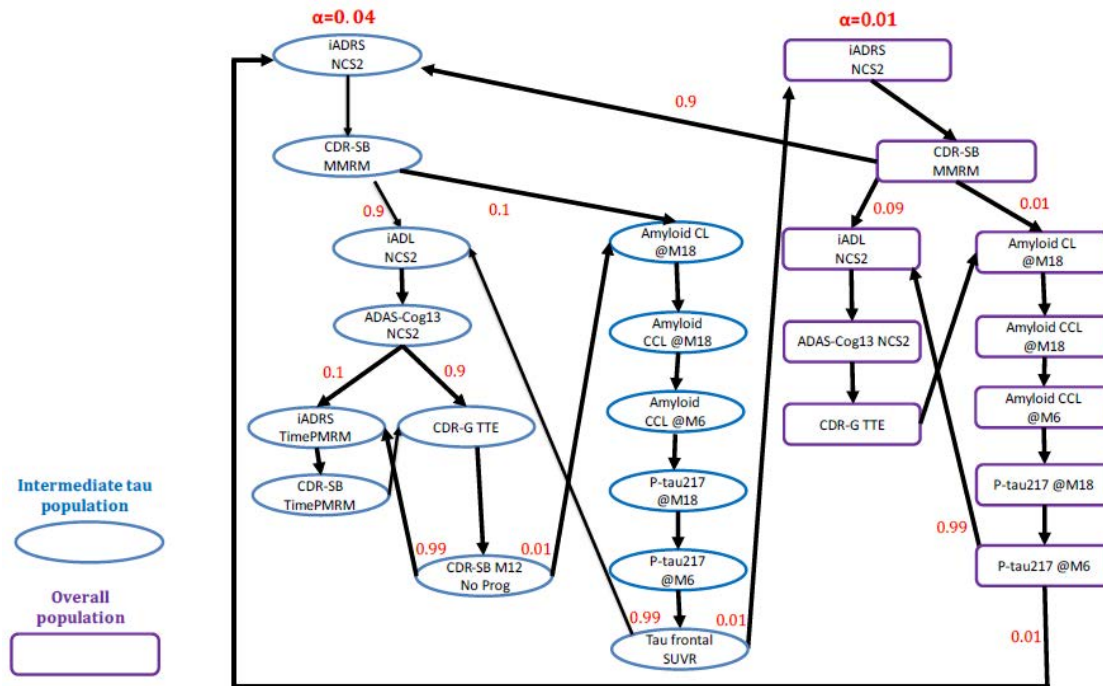
A participant's brain amyloid deposition was measured by amyloid PET imaging at visits: screening, 24, 52, and 76 weeks. Scan measurements were standardized to amyloid centiloid. The change from baseline to the post-baseline visit of the amyloid imaging centiloid was evaluated using an MMRM model. The model included the fixed effects terms: treatment, visit, treatment-by-visit interaction, baseline centiloid, baseline centiloid-by-visit interaction, and age at baseline. Baseline tau category was included as a fixed effect to the model for the overall population. Visit was treated as a categorical variable with values equal to the visit numbers at which amyloid imaging was assessed. For MMRM, the primary time point for treatment comparison was Week 76. The treatment group contrast in least-squares mean progression with 95% CI was calculated using the MMRM model. MMRM assumed missing data were missing at random.

3.2.2.5 Multiplicity Control

A prespecified hypothesis testing plan employed Bretz's graphical approach

(Bretz et al. 2009¹², 2011¹³) to provide control of the study-wise Type I error rate for the primary and key secondary hypotheses at 2-sided level $\alpha=0.05$. For the primary endpoint hypothesis test, the initial 2-sided alpha level was set to 0.04 for the baseline intermediate tau level population and 0.01 for the overall population. The hypothesis testing scheme, alpha recycle, and weights are described in Figure 2.

Figure 2 Hypothesis Testing Scheme to Control Type I Error Rate at Two-sided Alpha Level of 0.05



Source: Statistical Analysis Plan Version 3, Figure AACI.2.1

3.2.2.6 Subgroup Analyses

The pre-specified subgroups included the baseline or screening variables:

- Age: <65, 65-74, ≥ 75 years
- Sex: female, male
- Race: white, black or African American, Asian
- Ethnicity: Hispanic or Latino, not Hispanic or Latino
- ApoE4 Carrier Status: Carrier defined as E2/E4, E3/E4, or E4/E4 genotype; non-carrier defined as all other genotypes
- Number of ApoE4 alleles: 0 (noncarrier), 1 (heterozygote), 2 (homozygote) E4 alleles
- Clinical staging at screening: mild cognitive impairment (MCI) defined as MMSE 27-30, mild AD defined as MMSE 20-26
- Baseline brain tau burden category: intermediate (low/medium), high tau

¹² Bretz F, Maurer W, Brannath W, Posch M. A graphical approach to sequentially rejective multiple test procedures. Stat Med. 2009;28(4):586-604. <https://doi.org/10.1002/sim.3495>

¹³ Bretz F, Posch M, Glimm E, et al. Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes, or parametric tests. Biom J. 2011;53(6):894-913. <https://doi.org/10.1002/bimj.201000239>

- Brain tau screening standard uptake value ratio (SUVR) terciles: SUVR <33 percentile, SUVR 33-67 percentile, SUVR >67 percentile
- BMI: <25, 25- <30, ≥ 30

Additional subgroups of interest, defined post hoc, included region (USA, non-USA) and MMSE (MMSE ≥ 27 , MMSE <27). Lower MMSE scores indicated a greater level of AD severity.

NCS2 analyses were conducted to assess the iADRS subgroup effect. The model setup and covariates were similar to the primary analysis of iADRS (Section 3.2.2.1 Primary Analysis of iADRS) with these additional covariates: subgroup-by-treatment, subgroup-by-basis expansion terms, and subgroup-by-basis expansion-by-treatment interactions. The analyses were conducted for the overall population and the intermediate population separately.

Although Sponsor pre-specified CDR-SB subgroup analysis using NCS2, we conducted instead subgroup analysis of CDR-SB using its primary analysis model MMRM (Section 3.2.2.2 Analysis of Select Clinical Endpoints) to allow comparison between the results of the overall population and subgroups.

3.2.2.7 Sample Size Calculations

The power and sample size determination of the trial were based on the intermediate tau pathology population. The assumptions for the power calculation used the results of the Phase 2 study AACG: the mean progression levels in the placebo and donanemab arms from the MMRM analysis were -10.06 and -6.86 iADRS points over 18 months, respectively, with a standard deviation of 11.06. The assumed discontinuation rate of study AACI was 30%. Multiple longitudinal data sets were simulated, and NCS2 was fit to each sample to determine the power. With a sample size of approximately 1000 randomized participants in the intermediate tau pathology population, NCS2 provided greater than 95% power assuming 0.05 two-sided Type I error of 0.05.

Approximately 1800 participants were planned for randomization in the trial. The Sponsor anticipated that approximately two-thirds of participants had intermediate tau and one-third of participants had high tau pathology.

3.2.3 Patient Disposition, Demographic and Baseline Characteristics

Disposition

See Figure 3 for a flowchart of subject disposition. Of the 8,240 subjects assessed for eligibility, 1,736 subjects (donanemab 860, placebo 876) were randomized. The most common reason for screen failures was failure to meet inclusion/exclusion criteria. A total of 1320/1736 (76%) completed the Double-Blind phase of the study (622/860 (72%) donanemab vs. 698/876 (80%) placebo). The study discontinuation rate was 404/1736 (23%); the donanemab arm had a higher study discontinuation rate compared to the placebo arm (231/860 (27%) vs. 173/876 (20%)). The rates of discontinuation by reason were similar between treatment arms except for withdrawal by subject (13% donanemab vs. 11% placebo) and adverse event including death (8% donanemab vs. 4% placebo). Kaplan-Meier plot of time to study discontinuation showed that subjects treated with donanemab were more likely to discontinue compared to placebo (Figure 4). Twelve patients had not completed the final study visit prior to the database lock and were neither classified as study completers or discontinuers.

Of the randomized subjects, 1727/1736 (99%) received at least 1 dose of treatment (853 donanemab, 874 placebo). Of the randomized subjects, 1289/1736 (74%) subjects completed treatment (595/860 (69%) donanemab vs. 694/876 (79%) placebo). Of the 12 subjects who had not completed the final study visit prior to the database lock, 2 completed treatment. The donanemab arm had a higher treatment discontinuation rate compared to the placebo arm (252/860 (29%) donanemab vs. 176/876 (20%) placebo). The rates of discontinuation by reason were similar between arms except for adverse event including death (13% donanemab vs. 4% placebo). Kaplan-Meier plot of time to treatment discontinuation showed that subjects treated with donanemab were more likely to discontinue compared to placebo (Figure 5). Not counting 19 subjects who were randomized but did not get treatment or who were still on treatment by the data cut-off date, 31/428 (7%) of subjects who discontinued treatment were followed up and completed the study.

Of the randomized subjects, i.e. overall population, 1182/1736 (68%) were subjects in the intermediate tau pathology population. The size of the primary analysis set EES (baseline and at least one post-baseline efficacy scale) in the intermediate tau population and the overall population varied depending on the missingness of the endpoint assessed. For the primary endpoint of iADRS, the EES in the overall population had 1599 subjects (775 donanemab vs. 824 placebo), which is 137 less subjects than the number of randomized subjects. The Safety analysis set had 1727 subjects (853 donanemab vs. 874 placebo).

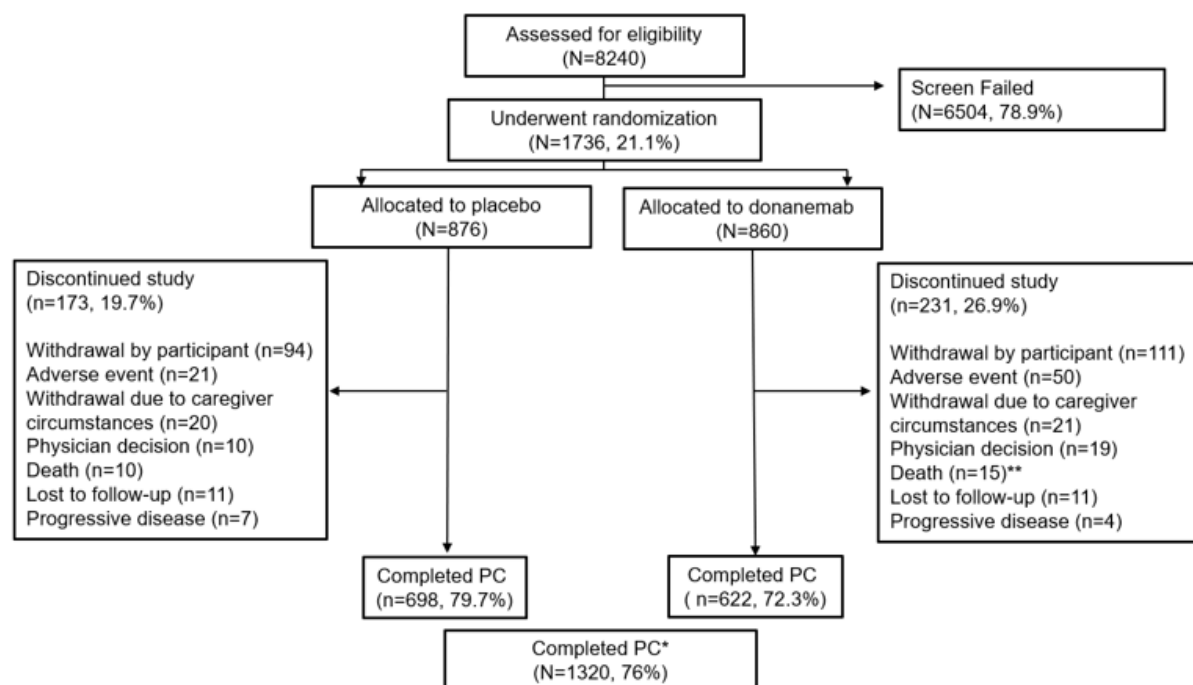
Reviewer Comment: According to the study protocol (Amendment e), subjects who discontinued treatment would remain in the study to be evaluated for efficacy and safety assessments according to schedule. However, only 31/428 (7%) of subjects who discontinued treatment were followed up and completed the study. Study conduct did not appear to adhere to the pre-specified treatment policy strategy estimand, whereby patients with intercurrent events are followed throughout the double-blind period.

Demographics/Baseline Characteristic

At baseline, the overall population (all randomized subjects) comprised 57% female, a mean age of 73 years, were predominately white (92%), and 72% were from the United States, followed by Poland (9%) and Canada (8%). Of all randomized subjects, 70% were ApoE4 carriers, 29% were noncarriers, and <1% were missing; 17% had both ApoE4 alleles (homozygotes) and 54% had one ApoE4 allele (heterozygotes). AchEI or memantine use was 61%. Demographics were generally balanced between treatment arms in the overall population and intermediate tau population (Table 2).

At screening, 84% had MMSE < 27, 16% had MMSE ≥ 27, 1 subject was missing MMSE value. MMSE distributions at screening were similar between treatment arms in the overall and intermediate tau population (results not shown). At baseline, the mean scores (standard deviation (SD)) for the outcome scales in the overall population were: iADRS 103.8 (SD 14.2); CDR-SB 3.9 (SD 2.1); ADAS-Cog13 29.0 (SD 8.9); ADCS-iADL 47.8 (SD 7.9); and MMSE 22.3 (SD 3.9). The CDR-G distribution at baseline for the overall population had <1% at score 0, 60% at score 0.5, 35% at score 1, 3% at score 2, and 1% missing. Baseline scores for these scales were similar between treatment arms in the overall population and intermediate population (Table 3).

Figure 3 Subject Disposition



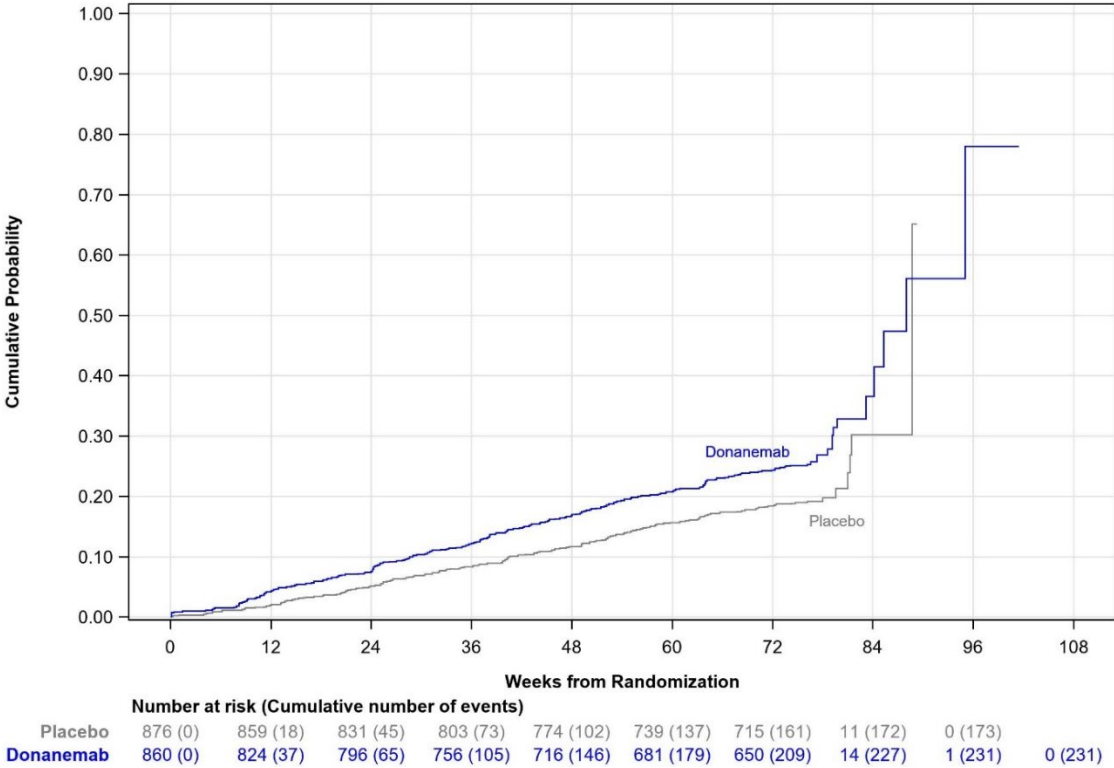
N = number of participants in the population; n = number of participants in the specified category; PC = placebo-controlled

* 12 participants did not complete final visit prior to database lock and are neither considered completers or discontinuers.

** 1 additional death occurred in the +57 day follow-up period.

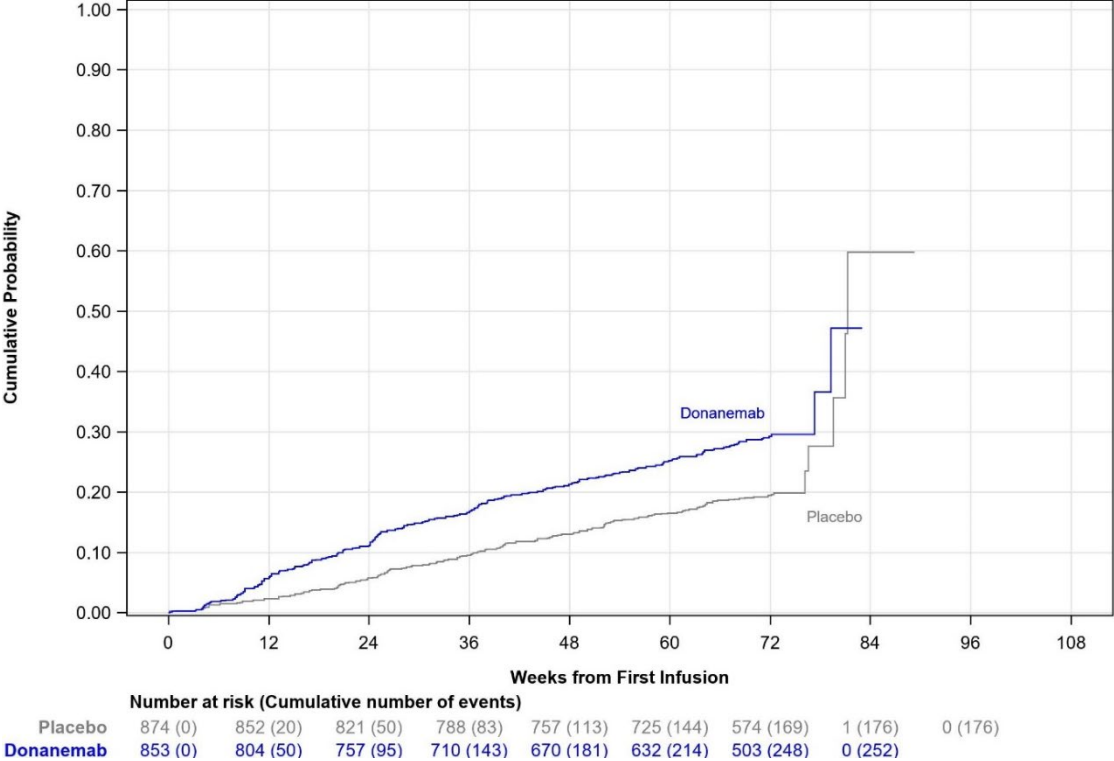
Source: Clinical Study Report Figure AACI.4.1, except eligibility and screen failures, other numbers could be verified by Statistical Analyst

Figure 4 Time to Study Discontinuation



Source: Statistical Analyst

Figure 5 Time to Treatment Discontinuation



The starting numbers at risk comprised subjects who had at least one dose of treatment.

Source: Statistical Analyst

Table 2 Demographics/Baseline Characteristics

	Intermediate Tau Population			Overall Population		
	Placebo (N = 594)	Donanemab (N = 588)	Total (N = 1182)	Placebo (N = 876)	Donanemab (N = 860)	Total (N = 1736)
Demographics^a						
Female sex, n (%)	321 (54.0)	325 (55.3)	646 (54.7)	503 (57.4)	493 (57.3)	996 (57.4)
Mean age, years (SD)	74.3 (5.8)	74.3 (5.7)	74.3 (5.7)	73.0 (6.2)	73.0 (6.2)	73.0 (6.2)
Race, n (%)						
Asian	38 (6.4)	48 (8.2)	86 (7.3)	47 (5.4)	57 (6.6)	104 (6.0)
Black or African American	17 (2.9)	17 (2.9)	34 (2.9)	21 (2.4)	19 (2.2)	40 (2.3)
White	539 (90.7)	522 (88.8)	1061 (89.8)	807 (92.1)	781 (90.9)	1588 (91.5)
Other ^b	0 (0.0)	1 (0.2)	1 (0.1)	1 (0.1)	2 (0.2)	3 (0.2)
Country, n (%)						
Australia	1 (0.2)	9 (1.5)	10 (0.8)	4 (0.5)	13 (1.5)	17 (1.0)
Canada	53 (8.9)	46 (7.8)	99 (8.4)	73 (8.3)	64 (7.4)	137 (7.9)
Czech Republic	8 (1.3)	5 (0.9)	13 (1.1)	10 (1.1)	12 (1.4)	22 (1.3)
Japan	36 (6.1)	40 (6.8)	76 (6.4)	43 (4.9)	45 (5.2)	88 (5.1)
Netherlands	5 (0.8)	6 (1.0)	11 (0.9)	9 (1.0)	13 (1.5)	22 (1.3)
Poland	57 (9.6)	53 (9.0)	110 (9.3)	82 (9.4)	77 (9.0)	159 (9.2)
United Kingdom	17 (2.9)	14 (2.4)	31 (2.6)	23 (2.6)	16 (1.9)	39 (2.2)
United States	417 (70.2)	415 (70.6)	832 (70.4)	632 (72.1)	620 (72.1)	1252 (72.1)
Ethnicity, Hispanic/Latino ^c , n (%)	26 (6.3)	24 (5.8)	50 (6.0)	36 (5.7)	35 (5.7)	71 (5.7)
Education, ≥13 years, n (%)	421 (71.0)	407 (69.2)	828 (70.1)	637 (72.8)	606 (70.5)	1243 (71.7)
APOE4 carrier, n (%)	427 (72.3)	421 (71.7)	848 (72.0)	621 (71.2)	598 (69.8)	1219 (70.5)
APOE genotypes, n (%)						
E2/E2	1 (0.2)	0 (0.0)	1 (0.1)	1 (0.1)	0 (0.0)	1 (0.1)
E2/E3	14 (2.4)	10 (1.7)	24 (2.0)	20 (2.3)	18 (2.1)	38 (2.2)
E2/E4	19 (3.2)	17 (2.9)	36 (3.1)	25 (2.9)	22 (2.6)	47 (2.7)
E3/E3	149 (25.2)	156 (26.6)	305 (25.9)	230 (26.4)	241 (28.1)	471 (27.2)
E3/E4	308 (52.1)	314 (53.5)	622 (52.8)	450 (51.6)	433 (50.5)	883 (51.1)
E4/E4	100 (16.9)	90 (15.3)	190 (16.1)	146 (16.7)	143 (16.7)	289 (16.7)
Missing	3	1	4	4	3	7
AChEI use, n (%)	307 (51.7)	305 (51.9)	612 (51.8)	488 (55.7)	478 (55.6)	966 (55.6)
Memantine use, n (%)	118 (19.9)	112 (19.0)	230 (19.5)	180 (20.5)	172 (20.0)	352 (20.3)
AChEI and/or memantine use, n (%)	341 (57.4)	332 (56.5)	673 (56.9)	538 (61.4)	521 (60.6)	1059 (61.0)

Abbreviations: APOE4 = apolipoprotein E allele 4; AChEI = acetylcholinesterase inhibitor; N = number of participants in the population; n = number of participants in the specified category; PC = placebo-controlled; SD = standard deviation

^a For all categories, the number of subjects with nonmissing data was used as the denominator.

^b Included Multiple and American Indian or Alaska Native.

^c Only included responses from US/Puerto Rico sites.

Source: Clinical Study Report Figure AACI.4.4, verified by Statistical Analyst

Table 3 Baseline Disease Scales and Biomarkers

	Intermediate Tau Population			Overall Population		
	Placebo (N = 594)	Donanemab (N = 588)	Total (N = 1182)	Placebo (N = 876)	Donanemab (N = 860)	Total (N = 1736)
Scale, Mean (SD), Range						
iADRS	105.5 (13.7), 47.0–139.0	105.7 (13.8), 42.0–135.0	105.6 (13.8), 42.0–139.0	103.6 (14.0), 47.0–139.0	104.1 (14.3), 31.0–135.0	103.8 (14.2), 31.0–139.0
CDR-SB	3.7 (2.0), 0.0–15.0	3.7 (2.1), 0.0–11.0	3.7 (2.1), 0.0–15.0	3.9 (2.1), 0.0–15.0	4.0 (2.1), 0.0–11.0	3.9 (2.1), 0.0–15.0
ADAS- Cog ₁₃	27.8 (8.4), 4.0–69.0	27.5 (8.5), 6.0–64.0	27.7 (8.4), 4.0–69.0	29.3 (8.9), 4.0–69.0	28.7 (8.8), 6.0–64.0	29.0 (8.9), 4.0–69.0
ADCS-ADL	66.9 (8.5), 20.0–78.0	66.7 (8.5), 28.0–78.0	66.8 (8.5), 20.0–78.0	66.4 (8.3), 20.0–78.0	66.3 (8.6), 7.0–78.0	66.4 (8.5), 7.0–78.0
ADCS-iADL	48.4 (7.9), 13.0–59.0	48.1 (7.9), 14.0–59.0	48.2 (7.9), 13.0–59.0	47.8 (7.8), 13.0–59.0	47.8 (7.9), 6.0–59.0	47.8 (7.9), 6.0–59.0
MMSE ^a	22.8 (3.8), 6.0–30.0	23.1 (3.6), 12.0–30.0	22.9 (3.7), 6.0–30.0	22.2 (3.9), 6.0–30.0	22.4 (3.8), 11.0–30.0	22.3 (3.9), 6.0–30.0
CDR – Global Score n (%)						
0	3 (0.5)	2 (0.3)	5 (0.4)	4 (0.5)	2 (0.2)	6 (0.4)
0.5	387 (65.5)	382 (65.9)	769 (65.7)	532 (61.2)	514 (60.8)	1046 (61.0)
1	185 (31.3)	177 (30.5)	362 (30.9)	308 (35.4)	304 (36.0)	612 (35.7)
2	16 (2.7)	19 (3.3)	35 (3.0)	25 (2.9)	25 (3.0)	50 (2.9)
Missing	3	8	11	7	15	22
Amyloid PET Centiloids	100.9 (35.1), 36.9–262.8	102.4 (34.7), 33.5–234.1	101.6 (34.9), 33.5–262.8	101.6 (34.5), 36.9–262.8	103.5 (34.5), 33.5–234.1	102.5 (34.5), 33.5–262.8
AD signature weighted neocortical flortaucipir SUVr	1.2 (0.1), 0.8–1.5	1.2 (0.1), 0.9–1.5	1.2 (0.1), 0.8–1.5	1.4 (0.3), 0.8–2.5	1.3 (0.3), 0.9–2.2	1.3 (0.3), 0.8–2.5

Abbreviations: ADAS-Cog₁₃ = Alzheimer's Disease Assessment Scale –13-item Cognitive subscale; ADCS-ADL = Alzheimer's Disease Cooperative Study – Activities of Daily Living scale; ADCS-iADL=Alzheimer's Disease Cooperative Study – instrumental Activities of Daily Living subscale; CDR-SB = Clinical Dementia Rating Scale – Sum of Boxes; iADRS = integrated Alzheimer's Disease Rating Scale; MMSE = Mini-Mental State Examination; MUBADA = multiblock barycentric discriminant analysis (method of analyzing PET images); N = number of participants in the population; n = number of participants in the specified category; PC = placebo-controlled; PET = positron emission tomography; SD = standard deviation; SUVr = standardized uptake value ratio

^a Last nonmissing MMSE prior to or on start of study treatment.

For all scales, the number of subjects with non-missing data was used as the denominator.

Source: Clinical Study Report Figure AACI.4.5, verified by Statistical Analyst

3.2.4 Results and Conclusions

3.2.4.1 Primary Analysis of iADRS

At Week 76, the donanemab arm compared to the placebo arm had statistically significantly less decline as measured by iADRS. The NCS2 difference in adjusted mean change from baseline of iADRS in the intermediate tau population (3.3; 95% CI 1.9, 4.6; $p < 0.001$) and the overall population (2.9; 95% CI 1.5, 4.3; $p < 0.001$) are shown in Table 4 and Figure 6.

Of the randomized subjects in the overall population, 1599/1736 (92%) subjects had both a baseline and post-baseline iADRS score. By Week 76, 1236/1736 (71%) had both a baseline and Week 76 iADRS score. The donanemab arm had higher percentages of missing data at baseline and at Week 76. Measurements after death were treated as missing. See Section 3.2.4.3 Sensitivity Analyses for an evaluation of the impact of missing data and the outcome of death on the primary analysis results.

Additional NCS2 model fit information is provided in Appendix Section 6.5 NCS2 Model Fit Estimates for covariate estimates and covariance matrix estimate and Appendix Section 6.4 NCS2 Model Diagnostic for a discussion of model diagnostics.

Reviewer Comment: The iADRS score ranged from 0 to 141 with lower scores reflecting worse performance. Although the treatment difference of about 3 points was significant, it is small in magnitude. We defer to the Clinical Team's efficacy review on the interpretation of the clinical meaningfulness of the treatment difference.

Table 4 Primary Efficacy Analysis (NCS2) of iADRS Change from Baseline at Week 76 in Evaluable Efficacy Set by Study Population

	Intermediate Tau Population		Overall Population	
	Placebo (N=594)	Donanemab (N=588)	Placebo (N=876)	Donanemab (N=860)
Baseline, n (%) ¹	560 (94%)	533 (91%)	824 (94%)	775 (90%)
Baseline, mean (SD)	106.0 (13.4)	105.9 (13.7)	103.8 (13.9)	104.6 (13.9)
Week 76, n (%) ²	444 (75%)	418 (71%)	653 (75%)	583 (68%)
Week 76, mean (SD)	98.9 (18.0)	101.3 (18.2)	93.8 (20.4)	97.0 (20.9)
Change from baseline, adj. mean ³ (SE)	-9.3 (0.5)	-6.0 (0.5)	-13.1 (0.5)	-10.2 (0.5)
Difference, adj. mean ³ (95% CI)		3.3 (1.9,4.6)		2.9 (1.5, 4.3)
p-value		$p < 0.001$		$p < 0.001$

Abbreviations: CI = 95% confidence interval, N = number of randomized subjects, SD = standard deviation, SE = standard error

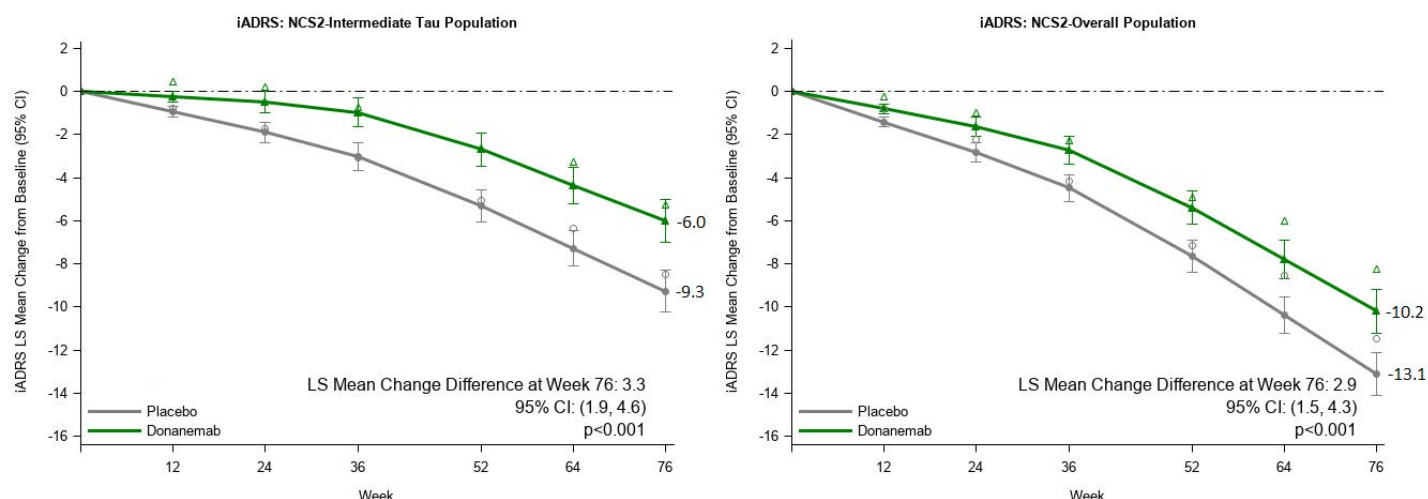
¹ number of randomized subjects with a baseline and at least one post-baseline efficacy score

² number of randomized subjects with a baseline and a Week 76 efficacy score

³Adjusted mean is from natural cubic spline model (NCS2) that adjusted for NCS basis expansion terms, NCS basis expansion term-by-treatment interaction, baseline age, baseline concomitant AChEI or memantine use, and pooled investigator. For the analysis of overall population, baseline tau category was also included as a covariate.

Source: Clinical Study Report Table AACI.5.2, Table AACI.5.3; findings reproduced by the Statistical Analyst

Figure 6 Primary Efficacy Analysis (NCS2) of iADRS Change from Baseline with Observed Mean Change from Baseline in Evaluable Efficacy Set by Study Population



Abbreviations: CI = 95% confidence interval, LS = Least square, NCS2 = Natural Cubic Splines with 2 degrees of freedom, p = p-value

NCS2 estimates with 95% CI are represented as filled symbols with CI bands. Observed estimates are represented as unfilled symbols.

Source: Statistical Analyst

3.2.4.2 Analysis of Select Clinical Endpoints

CDR-SB

At Week 76, the donanemab arm compared to the placebo arm had statistically significantly less decline as measured by CDR-SB. The difference in adjusted mean change from baseline of CDR-SB in the intermediate tau population was -0.7 (95% CI -1.0, -0.4; p < 0.001) and the overall population was -0.7 (95% CI -1.0, -0.5; p < 0.001) (Figure 7, Appendix Table 7). For reference, the range for CDR-SB was 0 to 18 with higher scores indicating a greater level of impairment.

ADCS-iADL

At Week 76, the donanemab arm compared to the placebo arm had statistically significantly less decline as measured by ADCS-iADL. The difference in adjusted mean change from baseline of ADCS-iADL in the intermediate tau population was 1.8 (95% CI 0.9, 2.7; p-value < 0.001) and the overall population was 1.7 (95% CI 0.8, 2.6; p-value < 0.001) (Figure 7, Appendix Table 7). For reference, the range for the iADL score was 0 to 59 with lower scores indicating greater disease severity.

ADAS-Cog13

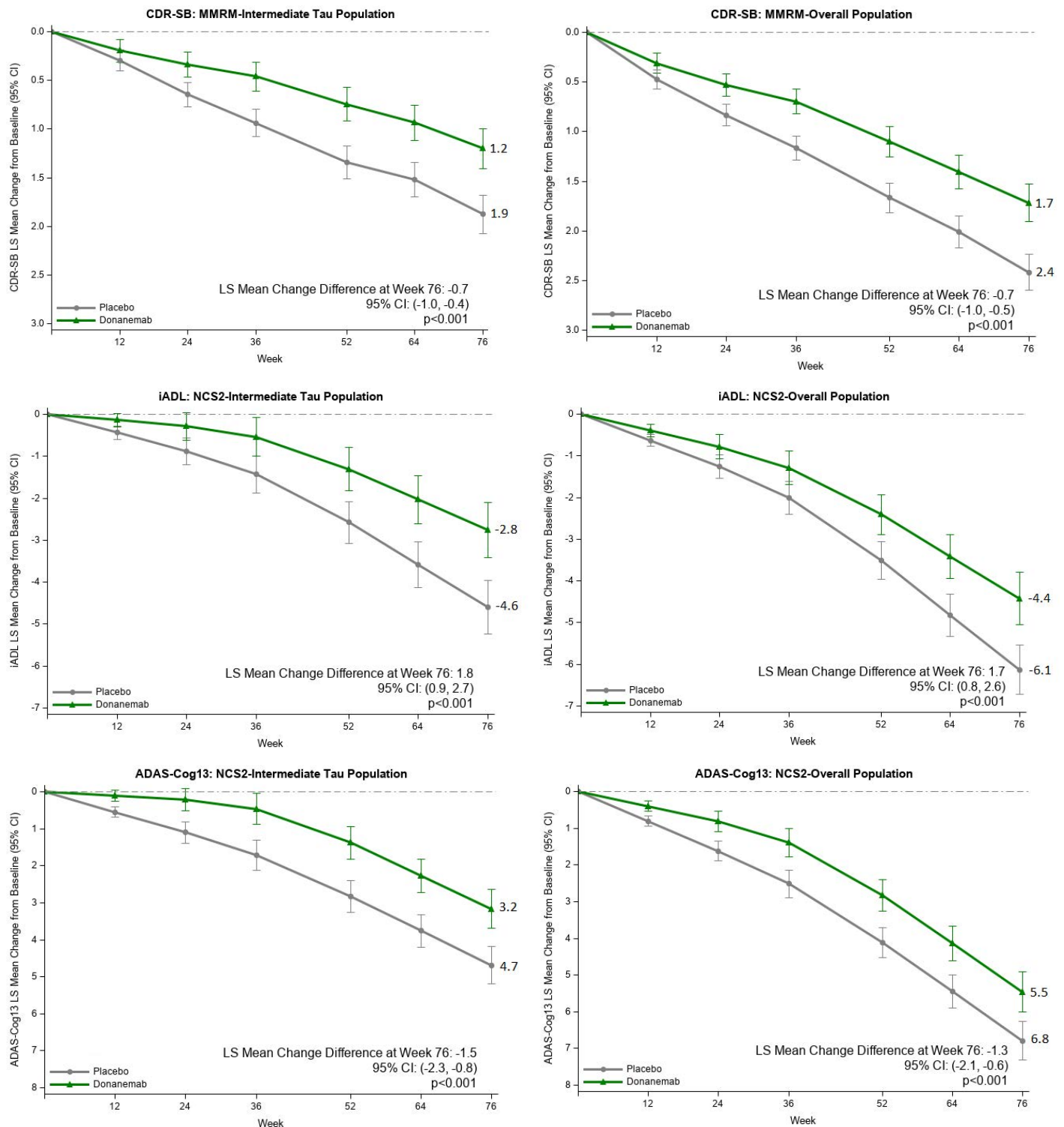
At Week 76, the donanemab arm compared to the placebo arm had statistically significantly less decline as measured by ADAS-COG13. The difference in adjusted mean change from baseline of ADAS-COG13 in the intermediate tau population was -1.5 (95% CI -2.3, -0.8; p-value < 0.001) and the overall population was -1.3 (95% CI -2.1, -0.6; p-value < 0.001) (Figure 7, Appendix Table 7). For reference, the range for ADAS-Cog13 was 0 to 85 with higher scores indicating greater disease severity.

CDR-G

By Week 76, the donanemab arm compared to the placebo arm had statistically significantly less risk for substantial decline in CDR-G. The hazard ratio (HR) in the intermediate tau population was 0.6 (95% CI 0.5, 0.8; p-value < 0.001) and the overall population was 0.6 (95% CI 0.5, 0.8; p-value < 0.001) (Figure 8, Figure 9).

Reviewer Comment: For the clinical endpoints CDR-SB, ADCS-iADL, and ADAS-COG13, the adjusted mean change was significant but small relative to the range of its scale.

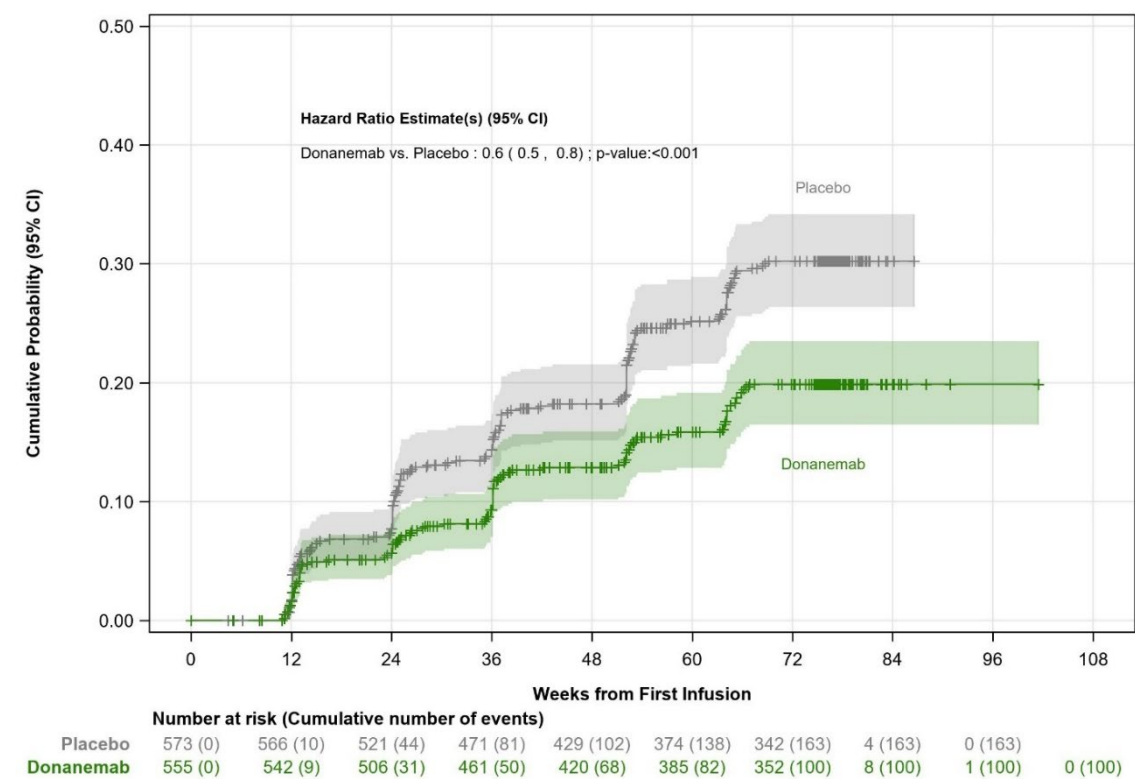
Figure 7 Analysis of Select Clinical Endpoints Change from Baseline at Week 76 in Evaluable Efficacy Set by Study Population



Abbreviations: CI = 95% confidence interval, LS = Least square, MMRM = Mixed Model for Repeated Measures, NCS2 = Natural Cubic Splines with 2 degrees of freedom, p = p-value

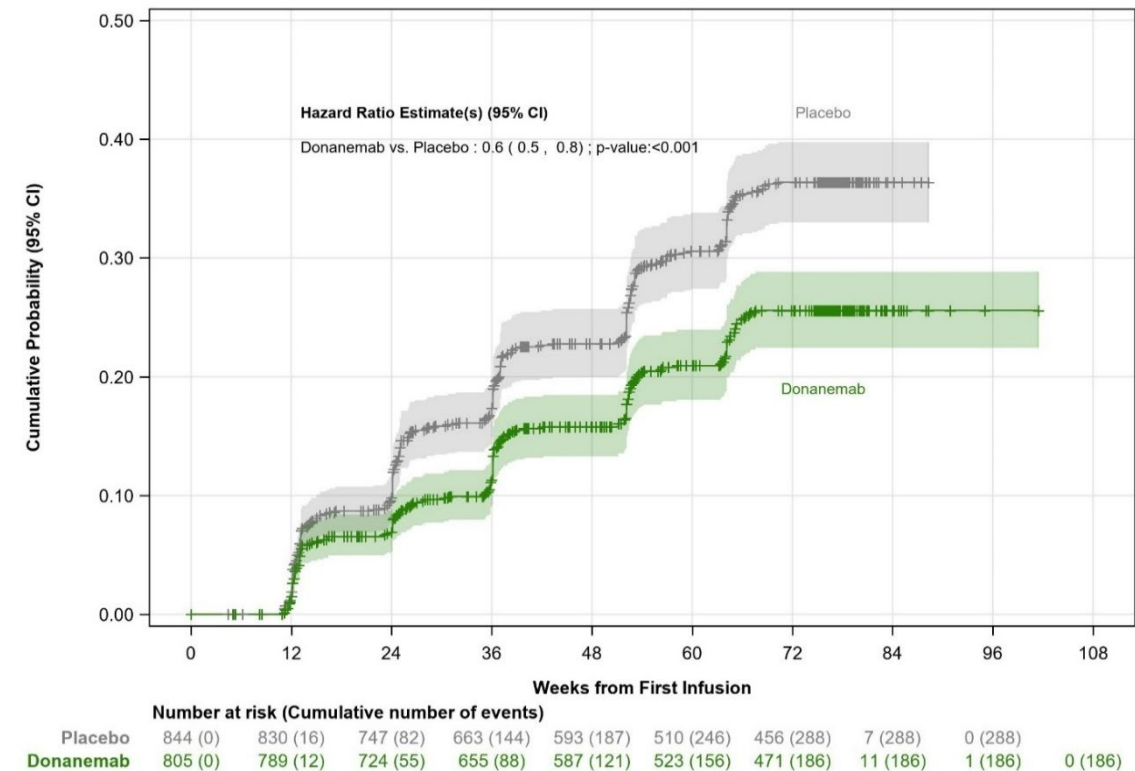
Source: Statistical Analyst

Figure 8 Kaplan-Meier Plot: Cumulative Incidence of Time to Substantial Decline in CDR-G by Week 76 in Evaluable Efficacy Set, Intermediate Tau Population



Abbreviations: CI = 95% confidence interval
The starting number at risk comprised subjects with a baseline and at least one post baseline outcome measurement.
Source: Statistical Analyst

Figure 9 Kaplan-Meier Plot: Cumulative Incidence of Time to Substantial Decline in CDR-G by Week 76 in Evaluable Efficacy Set, Overall Population



Abbreviations: CI = 95% confidence interval
The starting number at risk comprised subjects with a baseline and at least one post baseline outcome measurement.
Source: Statistical Analyst

3.2.4.3 Sensitivity Analyses

3.2.4.3.1 Pre-specified Sensitivity Analyses of iADRS

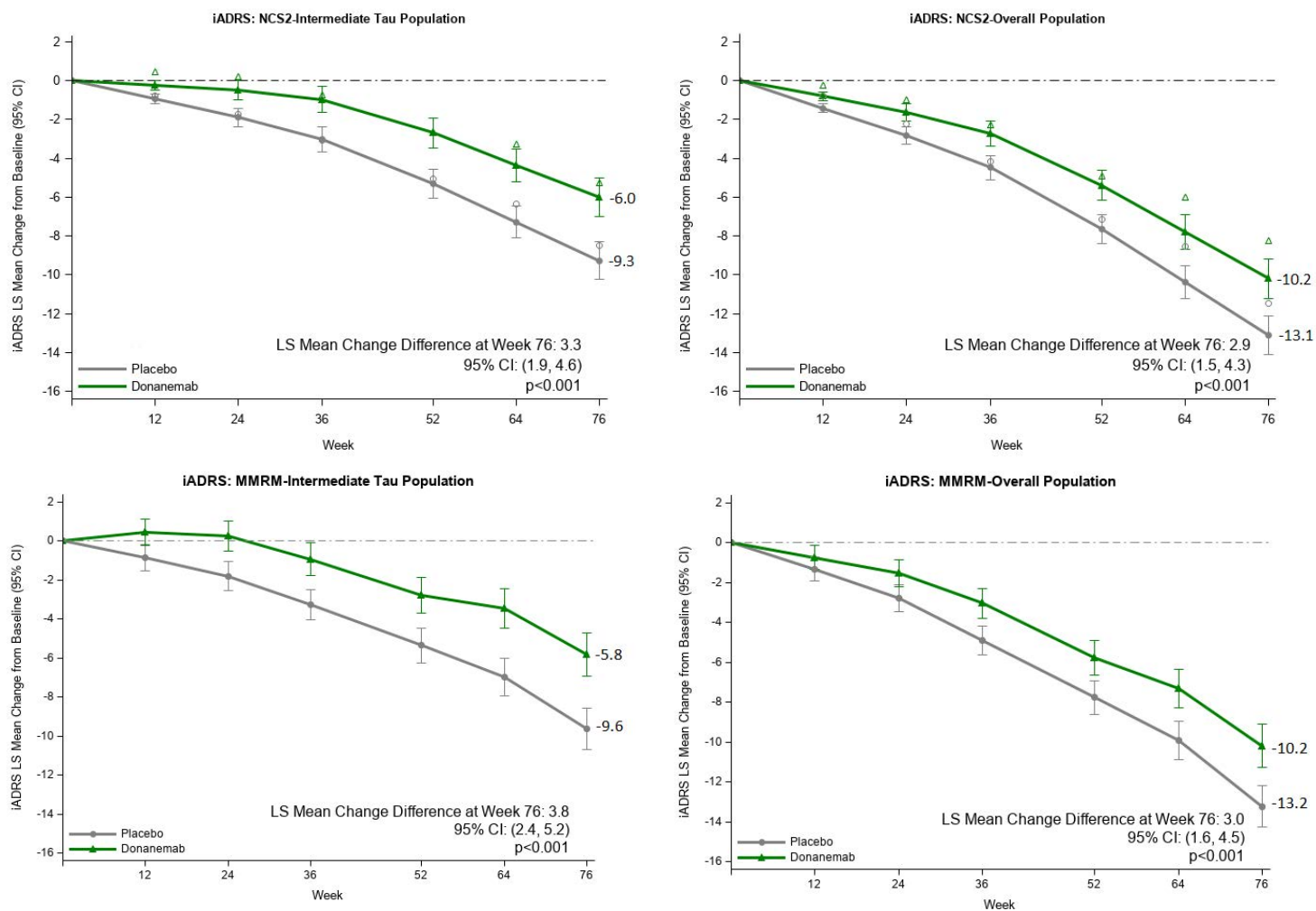
Refer to Figure 10 for the pre-specified sensitivity analysis results.

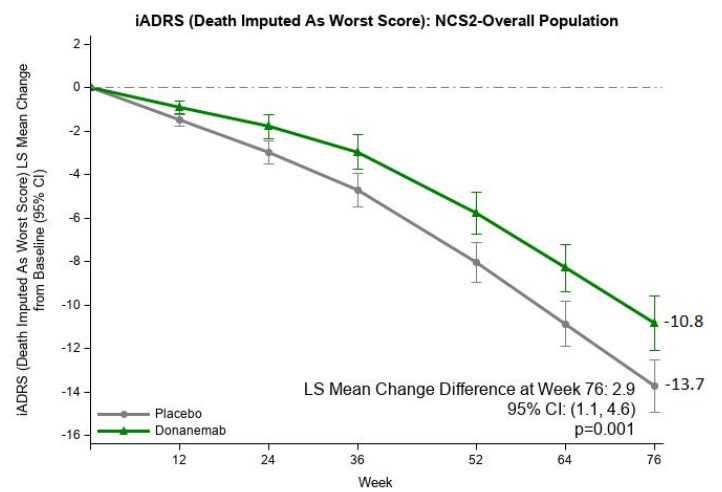
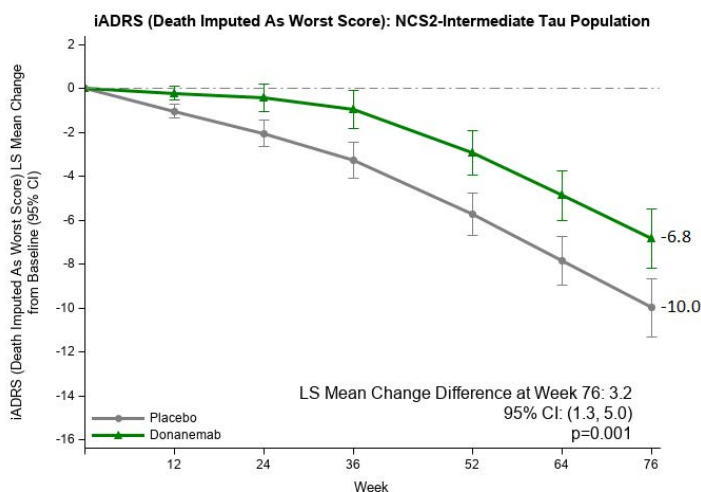
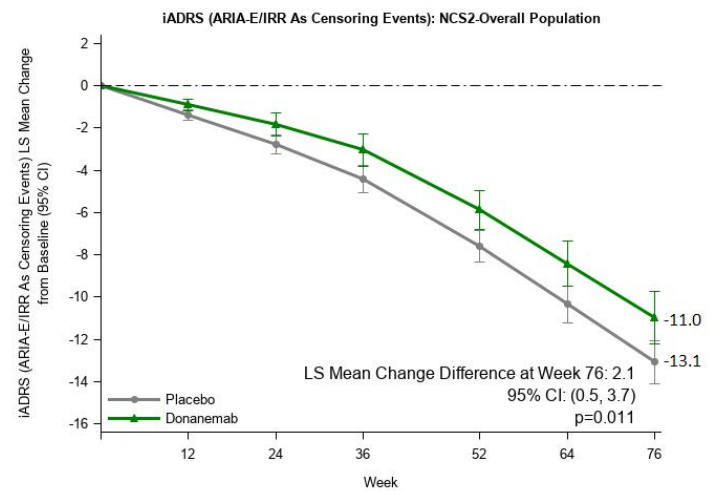
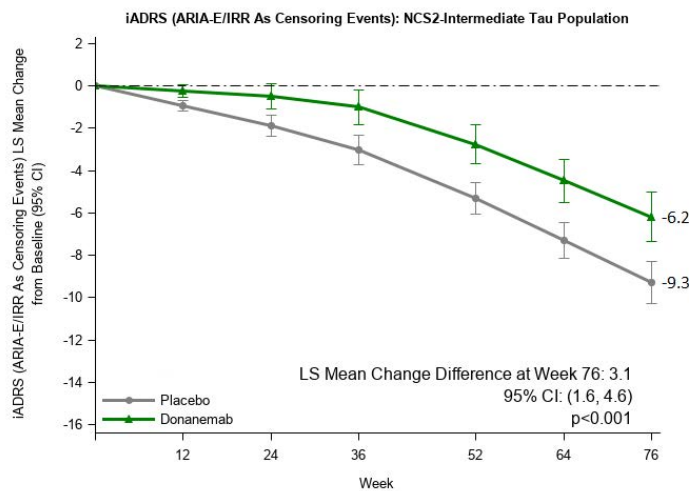
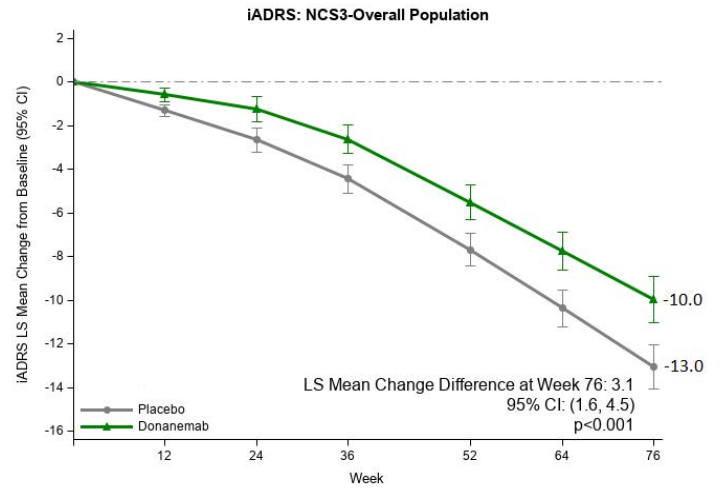
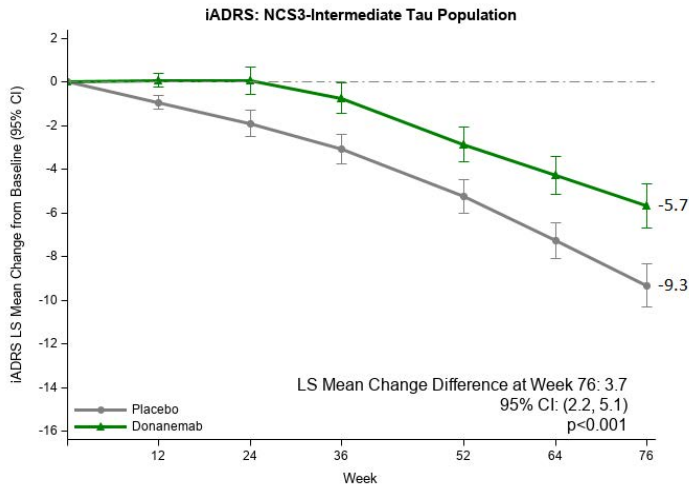
MMRM and NCS3 analyses produced results that were consistent with the primary analysis of iADRS at Week 76 using NCS2. Compared to NCS2, MMRM produced a greater treatment effect estimate in the intermediate tau population (3.8 vs. 3.3), but a similar estimate in the overall population (3.0 vs. 2.9). NCS2 estimates of iADRS mean change from baseline at Week 76 were smaller than the observed mean change at Week 76, a pattern that was also seen in the MMRM analysis (Appendix Figure 13).

To address the potential effect of functional unblinding due to ARIA-E or infusion related reactions, the primary analysis was conducted on a reduced dataset in which all assessments after occurrence of ARIA-E or infusion related reaction were excluded. The results were consistent with the primary analysis of iADRS at Week 76, and therefore, did not suggest a bias because of functional unblinding.

The analysis with death imputed as the worst score for measurements after death also produced similar results to the primary analysis of iADRS change from baseline at Week 76 (see also Section 3.2.4.3.2 Death Imputed as the Worst Score).

Figure 10 Primary (NCS2) and Sensitivity Efficacy Analysis of iADRS Change from Baseline at Week 76 in Evaluable Efficacy Set by Study Population





Abbreviations: CI = 95% confidence interval, LS = Least square, MMRM = Mixed Model for Repeated Measures, NCS2 = Natural Cubic Splines with 2 degrees of freedom, NCS3 = Natural Cubic Splines with 3 degrees of freedom, p = p -value

Model estimates are represented as filled symbols and observed estimates are represented as unfilled symbols.

Source: Statistical Analyst

3.2.4.3.2 Death Imputed as the Worst Score

The analyses of clinical endpoints (iADRS, CDR-SB, ADCS-iADL, ADAS-Cog13, time to substantial decline in CDR-G) with measurements after death imputed as the worst score are shown in Appendix Table 8. The results for CDR-SB and ADCS-iADL change from baseline at Week 76 were similar to the results from their corresponding primary analyses with death as a censoring event (Section 3.2.4.2 Analysis of Select Clinical Endpoints). For ADAS-Cog13 at change from baseline at Week 76, the estimates with death imputed as the worst score in the intermediate tau population (-1.0; 95% CI -2.2, 0.1; $p = 0.07$) and the overall population (-1.0; 95% CI -2.0, 0.0; $p = 0.04$) were not significant based on the alpha allocation of 0.04 and 0.01, respectively (Section 3.2.2.5 Multiplicity Control). Compared to the primary analysis of ADAS-Cog13 change from baseline at Week 76 with death as a censoring event (intermediate tau population -1.5 (95% CI -2.3, -0.8); overall population -1.3 (95% CI -2.1, -0.6)), the sensitivity analysis results were attenuated. Based on the graphical testing hierarchy, results for time to substantial decline in CDR-G by Week 76 would not be statistically significant if ADAS-Cog13 change from baseline at Week 76 were not significant.

3.2.4.3.3 Rank ANCOVA with Multiply Imputed Data at Week 76

iADRS

The analysis set comprised subjects who had non-missing baseline variables: treatment, iADRS score, age, concomitant AChEI and/or memantine use, and pooled investigator. These variables were used in the regression model for imputation. The size of the analysis set was 1684 subjects compared to 1736 total randomized subjects in the overall population. The rank ANCOVA resulted in statistically significant difference between donanemab and placebo for the intermediate tau and overall population (p -values = 0.001, results not shown).

CDR-SB

The analysis set comprised subjects who had non-missing baseline variables: treatment, CDR-SB score, age, concomitant AChEI and/or memantine use, and pooled investigator. These variables were used in the regression model for imputation. The size of the analysis set was 1712 subjects compared to 1736 total randomized subjects in the overall population. The rank ANCOVA resulted in statistically significant difference between donanemab and placebo for the intermediate tau and overall population (p -values < 0.01, results not shown).

Reviewer Comment: The statistically significant rank ANCOVA results indicated that primary analysis results of iADRS and CDR-SB change from baseline at Week 76 were not very sensitive to parametric model assumptions imposed by NCS2 or MMRM.

3.2.4.3.4 *Tipping Point Analyses*

The analysis sets are the same as those used in the rank ANCOVA analyses.

CDR-SB

For both the intermediate tau and overall population, the tipping point was close to the delta pair (0, 0.5) where placebo subjects with missing outcomes had imputed values (delta= 0) similar to observed subject while donanemab subjects with missing outcomes had at least a 0.5-point worse shift on CDR-SB (delta = 0.5) compared to observed subjects. The tipping point was assessed using the alpha allocation (intermediate tau population 0.04, overall population 0.01) where CDR-SB inherited all the alpha from testing iADRS. For other scenarios where placebo subjects with missing outcomes had a worse shift (delta > 0) compared to observed subjects, the delta pairs close to the tipping point were those with a difference of approximately 0.5 point (e.g., (0.5,1.0); (1.0,1.5)). See Table 5 and Table 6 for results.

iADRS

Intermediate tau population: The tipping point was close to the delta pair (0, -4) where placebo subjects with missing outcomes had imputed values (delta= 0) similar to observed subjects while donanemab subjects with missing outcomes had a 4-point worse shift on iADRS (delta = -4) compared to observed subjects. The tipping point was assessed using alpha = 0.04, which was the alpha allocated to testing iADRS in the intermediate tau population. For other scenarios where placebo subjects with missing outcomes had a worse shift (delta > 0) compared to observed subjects, the delta pairs close to the tipping point were those with a difference of approximately 4 points. See Appendix Table 9 for results.

Overall population: The tipping point was close to the delta pair (0, -2) where placebo subjects had imputed values (delta= 0) similar to observed subject while donanemab subjects with missing outcomes had a 2-point worse shift on iADRS (delta = -2) compared to observed subjects. The tipping point was assessed using alpha = 0.01, which was the alpha allocated to testing iADRS in the overall population. For other scenarios where placebo subjects with missing outcomes had a worse shift (delta > 0) compared to observed subjects, the delta pairs close to the tipping point were those with a difference of approximately 2 points. See Appendix Table 10 for results.

Reviewer Comment: The imputation for subjects with missing data (assumed missing at random) used the fitted model where there was a treatment effect favoring donanemab (e.g., intermediate tau population: -0.7 CDR-SB, 3.3 iADRS). A worse shift for a subject in a treatment arm was added to the outcome that was imputed. The primary analysis of CDR-SB change from baseline at Week 76 became non-significant near the delta pair (0, 0.5) where the treatment difference between imputed subjects might still favor donanemab. Also, the tipping point occurred with the minimum within-subject scale increment of 0.5.

Table 5 Tipping Point Analysis of CDR-SB Change from Baseline at Week 76 in Evaluable Efficacy Set, Intermediate Tau Population

Delta for Placebo	Delta for LY3002813	N		p-value*
		Placebo	LY3002813	
0.00	0.00	591	580	.005
0.00	0.50	591	580	.039
0.00	1.00	591	580	.196
0.50	0.00	591	580	<.001
0.50	0.50	591	580	.007
0.50	1.00	591	580	.046
0.50	1.50	591	580	.214
1.00	0.00	591	580	<.001
1.00	0.50	591	580	<.001
1.00	1.00	591	580	.009
1.00	1.50	591	580	.055
1.00	2.00	591	580	.226

* P-value is from Cochran-Mantel-Haenszel raw mean score statistics.

Source: July 20, 2023, Sponsor Response to Information Request Table APP.3.2

Table 6 Tipping Point Analysis of CDR-SB Change from Baseline at Week 76 in Evaluable Efficacy Set, Overall Population

Delta for Placebo	Delta for LY3002813	N		p-value*
		Placebo	LY3002813	
0.00	0.00	868	844	<.001
0.00	0.50	868	844	.004
0.00	1.00	868	844	.036
0.00	1.50	868	844	.225
0.50	0.00	868	844	<.001
0.50	0.50	868	844	<.001
0.50	1.00	868	844	.005
0.50	1.50	868	844	.043
0.50	2.00	868	844	.248
1.00	0.00	868	844	<.001
1.00	0.50	868	844	<.001
1.00	1.00	868	844	<.001
1.00	1.50	868	844	.007
1.00	2.00	868	844	.055
1.00	2.50	868	844	.265
1.50	0.00	868	844	<.001
1.50	0.50	868	844	<.001
1.50	1.00	868	844	<.001
1.50	1.50	868	844	<.001
1.50	2.00	868	844	.011
1.50	2.50	868	844	.070
1.50	3.00	868	844	.295

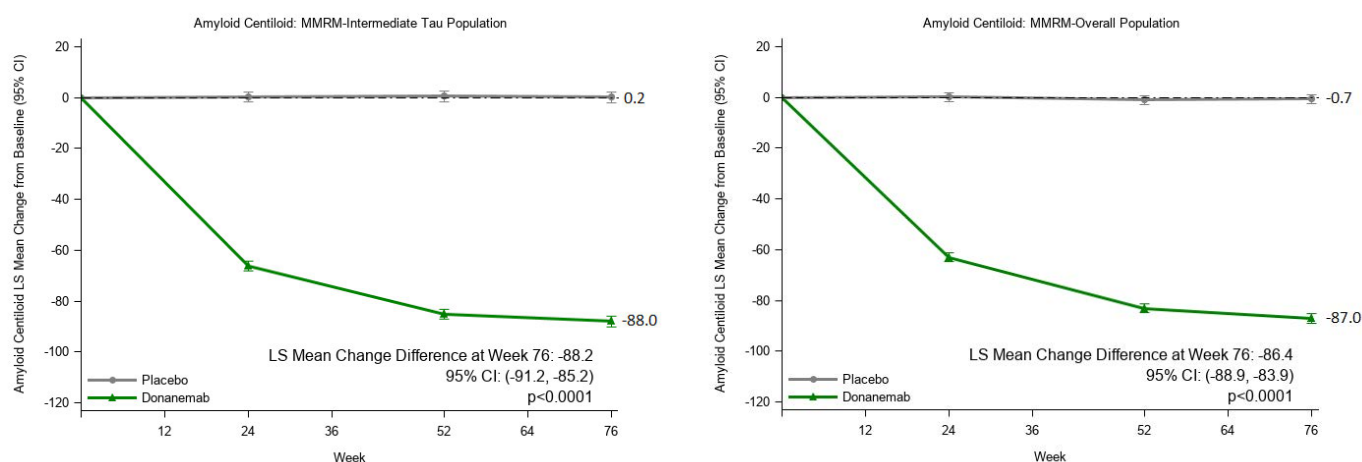
* P-value is from Cochran-Mantel-Haenszel raw mean score statistics.

Source: July 20, 2023, Sponsor Response to Information Request Table APP.3.1

3.2.4.4 Analysis of Amyloid PET

At Week 76, the donanemab arm compared to the placebo arm had statistically significantly more amyloid plaque reduction as measured by amyloid PET (centiloid units). The difference in adjusted mean change from baseline in amyloid centiloids in the intermediate tau population was -88.2 (95% CI -91.2, -85.2; $p < 0.0001$) and the overall population was -86.4 (95% CI -88.9, -83.9; $p < 0.0001$) (Figure 11, Appendix Table 7).

Figure 11 Analysis of Amyloid Change from Baseline at Week 76 in Evaluable Efficacy Set by Study Population



Abbreviations: CI = 95% confidence interval, LS = Least square, MMRM = Mixed Model for Repeated Measures, p = p-value

Source: Statistical Analyst

3.3 Evaluation of Safety

For a review of safety, refer to the Clinical Team's safety review.

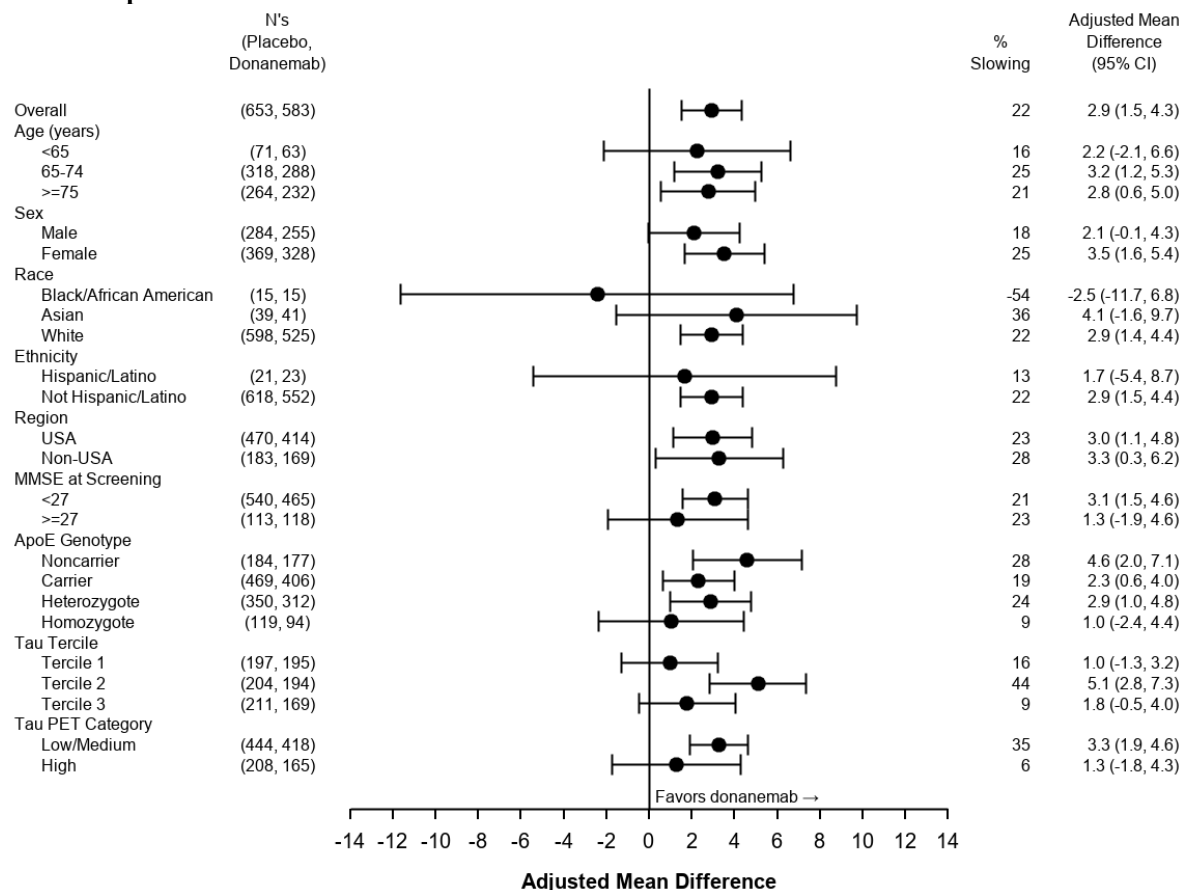
4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

The subgroups of clinical interest were of the baseline variables: age, sex, race, ethnicity, region, MMSE category, ApoE4 carrier status, number of ApoE4 alleles, and tau SUVR tercile. Except for the small (N=30) subgroup of black/African American (-2.5; 95% CI -11.7, 6.8), other subgroup treatment effect estimates of iADRS change from baseline at Week 76 ranged from 1.0 to 5.1, with less precision compared to overall population estimate (2.9; 95% CI 1.5, 4.3). Subgroup analysis results of iADRS are shown in Figure 12 for the overall population.

Except for the small (N=45) subgroup of Hispanic/Latino (1.0; 95% CI -0.8, 2.9), other subgroup treatment effect estimates of CDR-SB change from baseline at Week 76 ranged from -0.1 to -1.0, with less precision compared to overall population estimate (-0.7; 95% CI -0.9, -0.5). Subgroup analysis results of CDR-SB are shown in Appendix Figure 14 for the overall population.

Reviewer Comment: The subgroup results of iADRS and CDR-SB change from baseline at Week 76 were generally consistent with its overall population estimate indicating that donanemab subjects had less decline compared to placebo subjects.

Figure 12 Subgroup Analysis (NCS2) of iADRS Change from Baseline at Week 76 in Evaluable Efficacy Set, Overall Population



Abbreviations: CI = 95% confidence interval, PET = positron emission tomography, N= number of subjects who have a baseline and Week 76 efficacy score.

ApoE genotype (# of E4 alleles): Noncarrier (0), Heterozygote (1), Homozygote (2).

Tau Tercile: Tercile 1 (standard uptake value ratio (SUVR) <33 percentile), Tercile 2 (SUVR 33-67 percentile), Tercile 3 (SUVR >67 percentile).

Ethnicity subgroups are based on responses from both USA and non-USA regions.

Adjusted mean is from natural cubic spline model (NCS2) that adjusted for NCS basis expansion terms, basis expansion terms*treatment, subgroup*treatment, subgroup*basis expansion terms, and subgroup*basis expansion*treatment, baseline age, baseline concomitant AchEI or memantine use, pooled investigator, and baseline tau PET category.

Source: Statistical Analyst

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

A major statistical issue is the sensitivity of the primary analysis of CDR-SB change from baseline at Week 76 to missing data assumptions. Study discontinuation was high with the donanemab arm having a higher rate compared to the placebo arm (27% vs. 20%). During the BLA review, the Sponsor conducted a sensitivity analysis of CDR-SB to varying missing data assumptions using tipping point analysis (Sections 3.2.2.3 Sensitivity Analyses, Section 3.2.4.3.4 Tipping Point Analyses). The primary analysis of CDR-SB change from baseline at Week 76 became non-significant near the delta pair (0, 0.5) where the treatment difference between imputed subjects might still favor donanemab. Also, the tipping point occurred with the minimum within-subject scale increment of 0.5.

Another issue was the handling of death, which resulted in missing data, in the primary analysis of clinical endpoints. Although Sponsor was advised to account for death in the primary efficacy analysis, Sponsor instead assumed death was not informative and treated measurements after death as being missing; thereby, ignoring the severity of death in a patient's disease trajectory. Sponsor instead pre-specified a sensitivity analysis of iADRS change from baseline at Week 76 that imputed measurements after death as the worst outcome. There was not a similar sensitivity analysis pre-specified for CDR-SB, ADAS-Cog13, ADCS-iADL, and time to substantial decline in CDR-G. Ignoring deaths, a negative outcome, could bias the primary analysis result for these endpoints. For ADAS-Cog13 change from baseline at Week 76, the estimates with death imputed as the worst score in the intermediate tau population (-1.0; 95% CI -2.2, 0.1; $p = 0.07$) and the overall population (-1.0; 95% CI -2.0, 0.0; $p = 0.04$) were not statistically significant compared to the alpha allocation of 0.04 and 0.01, respectively. Based on the graphical testing hierarchy, results for time to substantial decline in CDR-G by Week 76 would not be statistically significant if ADAS-Cog13 change from baseline at Week 76 were not significant.

5.2 Collective Evidence

The collective evidence included the Phase 2 study AACG and Phase 3 study AACI, which was the focus of this review. The number of randomized subjects in AACI (1736 subjects) was larger compared to AACG (257 subjects). Another difference between the two studies was the study population: AACG comprised subjects with intermediate tau pathology; AACI comprised subjects with intermediate and high tau pathology. In AACG, donanemab subjects had statistically significantly less decline at Week 76 than placebo subjects as measured by iADRS (MMRM analysis 3.2; 95% CI 0.1, 6.3; p -value = 0.04), thus meeting the applicant's primary objective for the study. The secondary endpoints of CDR-SB, ADAS-cog, MMSE, and ADCS-iADL were not significant based on the pre-specified alpha allocation. For a complete review of study AACG, refer to Massie 2023¹⁴.

Both AACG and AACI studies showed statistically significantly less decline at Week 76 among donanemab subjects as measured by iADRS with similar point estimates around 3.0, with the

¹⁴ Massie, T. Statistical Review and Evaluation of Kisunla (Donanemab) - Biomarker Accelerated Approval. January 17, 2023. Available in DARRTS.

estimate from AACI having more precision because of a larger sample size. The significance of CDR-SB and ADCS-iADL change from baseline at Week 76 in AACI were supportive of a treatment effect.

5.3 Conclusions and Recommendations

Study AACI showed, based on the primary analysis, that the donanemab arm compared to the placebo arm had statistically significantly less decline at Week 76 as measured by iADRS. Compared to the range of the scale (0 to 141), the treatment difference of approximately 3 points was small relatively but significant. The treatment effect estimates for secondary endpoints of CDR-SB and ADCS-iADL change from baseline at Week 76 were small but also significant, providing support for a treatment effect with donanemab.

Sensitivity analyses of iADRS using mixed model with repeated measures, a model with less assumptions and more commonly used as the primary analysis of longitudinal AD data, produced a significant treatment effect. A non-parametric rank analysis of covariance without imposition of parametric model assumption also produced a significant treatment effect. Sensitivity analysis of iADRS with death imputed as the worst score produced a similar treatment effect that was also significant. Study discontinuation was high with the donanemab arm having a higher rate compared to the placebo arm (27% vs. 20%). Tipping point analysis to assess the sensitivity of the primary analysis for CDR-SB to missing data assumptions showed that the results for both the intermediate tau and overall population became non-significant near the delta pair (0, 0.5), where 0.5 is the minimum within-subject scale increment change.

Despite the potential for the primary analysis of CDR-SB to be sensitive to missing data assumptions, study AACI demonstrated evidence of effectiveness. Study AACG provided confirmatory evidence of effectiveness. We recommend approval of donanemab.

For future studies of AD, we recommend that the Division provide specific protocol recommendations for sponsors to proactively follow up study subjects to the end of the trial to minimize missing data. High missing data rates can result in efficacy and safety estimates that are unreliable for decision making. Additionally, pre-specifying a clinically relevant cutoff for a tipping point analysis would allow for a structured assessment of plausible missing data assumptions.

5.4 Labeling Recommendations

None.

6 APPENDICES

6.1 Natural Cubic Spline

Sponsor referenced Donohue et al. (2022)¹⁵ for their use of the natural cubic spline analysis. Natural cubic splines model the temporal mean trend, assuming time to be continuous. A cubic spline is a function defined by cubic polynomials that are spliced together at knot locations, and the resulting function is restricted to be continuous and have continuous first and second derivatives. A natural cubic spline has a further constraint of having a second derivative of zero at the boundaries. A spline function $f(t)$ of time t can be expressed as a linear combination of

basis functions, $b_k(t)$, and weights β_k : $f(t) = \beta_0 + \sum_{k=1}^m \beta_k b_k(t)$. Once the basis functions are determined, the weights can be estimated by maximum likelihood estimation.

The response for subject $i = 1, \dots, N$ and study visit $j = 1, \dots, K$ is modeled by natural cubic spline as:

$$Y_{ij} = \beta_0 + \sum_{k=1}^m \beta_k b_k(t) + 1_i(x) \sum_{k=1}^m \gamma_k b_k(t) + \varepsilon_{ij}$$

where x is the treatment variable, and the indicator $1_i(x)$ is 1 if subject i is in the treatment group and 0 otherwise. The residual vector ε_{ij} of each subject follows a multivariate normal distribution with mean zero and an assumed covariance structure. The resulting curve for the placebo group

is defined by the natural cubic spline $f(t) = \beta_0 + \sum_{k=1}^m \beta_k b_k(t)$; while the natural cubic spline

$g(t) = \sum_{k=1}^m \gamma_k b_k(t)$ represents the treatment group difference over time and is constrained to be zero at time zero. Given the known basis functions, estimation of the unknown parameters (β s and γ s) is accomplished with generalized least squares. Baseline covariates can be added to the model.

Interior knots are typically spaced according to quantiles of t . For a natural cubic spline with 2 degrees of freedom (e.g., 2 knots at the boundaries, 1 knot at the median observation time), there are $m = 2$ basis functions. For a natural cubic spline with 3 degrees of freedom (e.g., 2 boundary knots, 2 internal knots over the time course, usually at 33rd and 67th percentiles), there are $m = 3$ basis functions.

Note that the estimation of the covariance matrix used categorical time.

¹⁵ Donohue MC, Langford O, Insel P, et al. Natural cubic splines for the analysis of Alzheimer's clinical trials [published online August 17, 2022]. <https://doi.org/10.48550/arXiv.2208.08077>

6.2 Rank ANCOVA

Sponsor conducted a nonparametric rank analysis of covariance (ANCOVA) (Quade 1967¹⁶) of multiply imputed iADRS and CDR-SB at Week 76 for the intermediate tau and overall population. The imputation model was a linear regression model adjusting for following baseline variables: iADRS score, pooled investigator, treatment, concomitant AChEI and/or memantine use, and age. Baseline tau category was included in the model for the overall population. To allow for imputation using the above model, the analysis dataset was restricted to all randomly assigned participants with a non-missing value for the variables in the model. One hundred datasets were created that included the observed Week 76 measurements as well as imputed values for the missing data. For participants who died during the trial, the worst possible score was used as their post-mortem measurement instead of the imputed score.

A rank ANCOVA analysis was conducted to assess the treatment effect for each of the imputed datasets as follows:

1. Rank the dependent variable (change in iADRS or CDR-SB at Week 76) within each unique combination of the categorical variables of pooled investigator and concomitant AChEI and/or memantine use at baseline. For the analysis of the overall population, also include the baseline tau category as a categorical variable. That is, the ranking occurred without the treatment arm, age, or baseline value as factors.
2. Conduct linear regression analysis using the ranks of the change from baseline iADRS or CDR-SB as the dependent variable and the continuous covariates (age and baseline iADRS or CDRSB, respectively) as the independent variables, within each unique combination of the categorical variables of pooled investigator and concomitant AChEI and/or memantine use at baseline. Obtain the residuals from the regression.
3. For each of the imputed datasets, apply Cochran-Mantel-Haenszel (CMH) mean score analysis to the obtained residual data to test for the differences between treatment groups across the levels of categorical covariates. To assess if the distribution of the residuals, differs by treatment group, the Hodges-Lehmann estimate, along with the 95% CI, was also calculated to estimate the median differences between treatment groups.
4. Combine the statistics calculated for each imputed dataset from Step 3, according to Rubin's rule. There were 100 datasets. The CMH statistics were normalized with a Wilson-Hilferty transformation before combining to avoid the possible skewness from the chi-square distribution (Goria 1992¹⁷; Wilson and Hilferty 1931¹⁸).

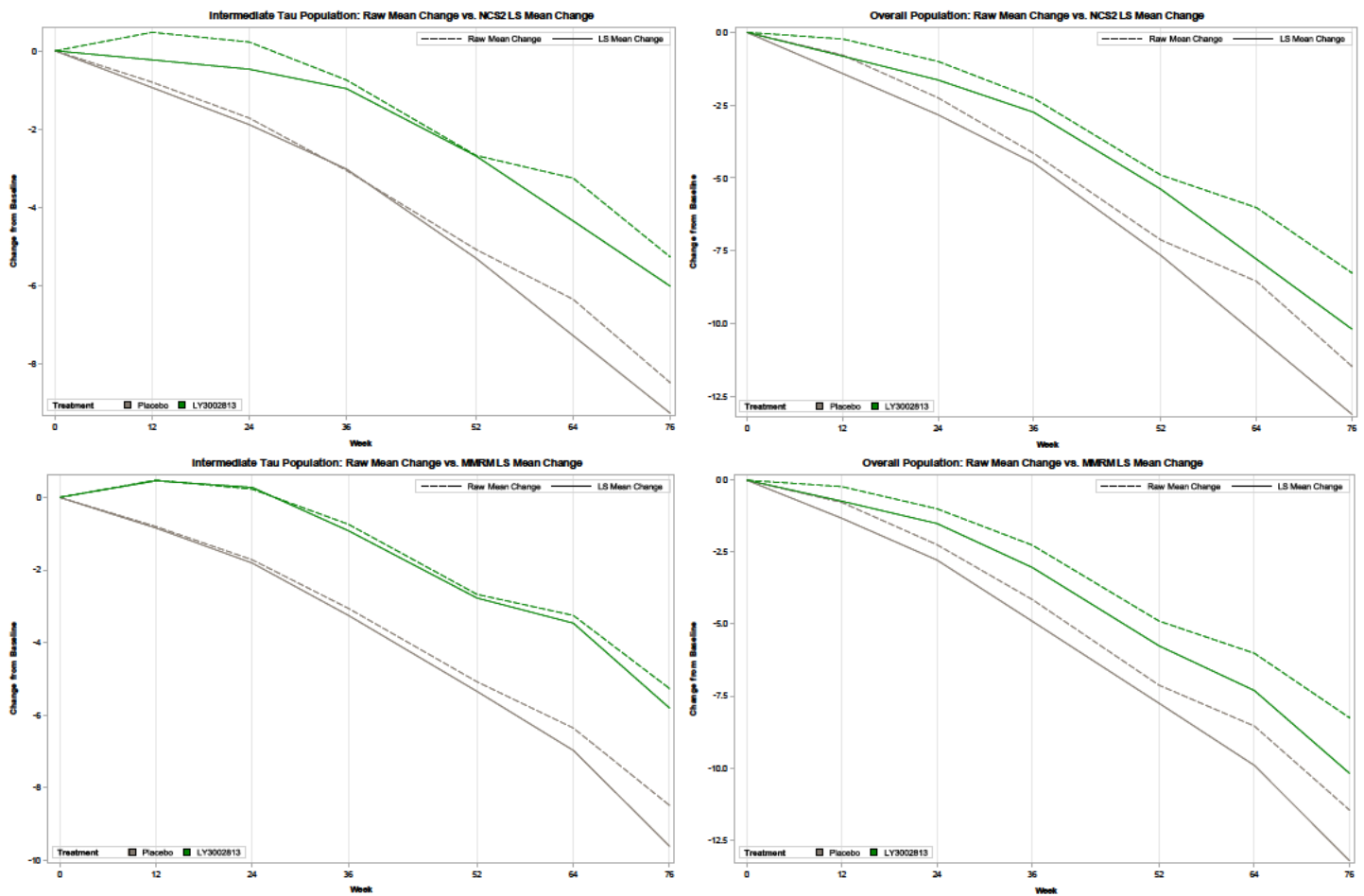
¹⁶ Quade D. Rank analysis of covariance. *J Am Stat Assoc.* 1967;62(320):1187-1200.
<https://doi.org/10.2307/2283769>

¹⁷ Goria MN. On the fourth root transformation of chi-square. *Aust J Stat.* 1992;34(1):55-64.
<https://doi.org/10.1111/j.1467-842X.1992.tb01043.x>

¹⁸ Wilson EB, Hilferty MM. The distribution of chi-squared. *Proc Natl Acad Sci (USA).* 1931;17(12):684-688.
<https://doi.org/10.1073/pnas.17.12.684>

6.3 Results

Figure 13 Natural Cubic Spline vs. Mixed Model for Repeated Measures Analysis of iADRS Change from Baseline in Evaluable Efficacy Set by Study Population



Abbreviations: LS = Least square, LY3002813 = donanemab, MMRM = Mixed Model for Repeated Measures, NCS2 = Natural Cubic Splines with 2 degrees of freedom, p = p-value
Source: September 11, 2023, Sponsor Response to Information Request

Table 7 Analysis of Select Endpoints Change from Baseline at Week 76 in Evaluable Efficacy Set by Study Population

Endpoint		Intermediate Tau Population		Overall Population	
		Placebo (N=594)	Donanemab (N=588)	Placebo (N=876)	Donanemab (N=860)
CDR-SB					
	Baseline, n (%) ¹	569 (96%)	546 (93%)	838 (96%)	794 (92%)
	Baseline, mean (SD)	3.6 (2.0)	3.7 (2.1)	3.9 (2.0)	3.9 (2.1)
	Week 76, n (%) ²	459 (77%)	424 (72%)	672 (77%)	598 (70%)
	Week 76, mean (SD)	5.1 (2.9)	4.6 (2.9)	5.8 (3.2)	5.3 (3.2)
	Difference, adj. mean ³ (95% CI)		-0.7 (-1.0, -0.4)		-0.7 (-1.0, -0.5)
	p-value		p < 0.001		p < 0.001
ADCS-iADL					
	Baseline, n (%) ¹	562 (95%)	535 (91%)	826 (94%)	780 (91%)
	Baseline, mean (SD)	48.6 (7.7)	48.2 (7.9)	48.0 (7.7)	48.0 (7.9)
	Week 76, n (%) ²	451 (76%)	420 (71%)	661 (75%)	591 (69%)
	Week 76, mean (SD)	45.1 (9.8)	46.1 (10.3)	43.3 (10.6)	44.5 (11.1)
	Difference, adj. mean ⁴ (95% CI)		1.8 (0.9, 2.7)		1.7 (0.8, 2.6)
	p-value		p < 0.001		p < 0.001
ADAS-Cog13					
	Baseline, n (%) ¹	570 (96%)	550 (94%)	841 (96%)	797 (93%)
	Baseline, mean (SD)	27.6 (8.2)	27.4 (8.4)	29.2 (8.9)	28.5 (8.8)
	Week 76, n (%) ²	460 (77%)	431 (73%)	677 (77%)	607 (71%)
	Week 76, mean (SD)	31.2 (10.4)	29.8 (10.7)	34.5 (12.0)	32.7 (12.4)
	Difference, adj. mean ⁴ (95% CI)		-1.5 (-2.3, -0.8)		-1.3 (-2.1, -0.6)
	p-value		p < 0.001		p < 0.001
Amyloid PET					
	Baseline, n (%) ¹	556 (94%)	525 (89%)	812 (93%)	765 (89%)
	Baseline, mean (SD)	100.9 (35.3)	103.0 (34.8)	101.8 (34.4)	104.0 (34.4)
	Week 76, n (%) ²	470 (79%)	433 (74%)	690 (79%)	614 (71%)
	Week 76, mean (SD)	101.6 (36.5)	13.4 (22.4)	101.8 (35.7)	15.0 (22.8)
	Difference, adj. mean ³ (95% CI)		-88.2 (-91.2, -85.2)		-86.4 (-88.9, -83.9)
	p-value		<0.0001		<0.0001

Abbreviations: CI = 95% confidence interval, N = number of randomized subjects, SD = standard deviation, SE = standard error

¹ number of randomized subjects with a baseline and at least one post-baseline efficacy score

² number of randomized subjects with a baseline and a Week 76 efficacy score

³Adjusted mean is from the Mixed Model with Repeated Measures (MMRM) that adjusted for treatment, visit, treatment-by-visit interaction, baseline iADRS score, baseline score-by-visit interaction, age at baseline, and concomitant AChEI and/or memantine use at baseline, and pooled investigator. For the analysis of overall population, baseline tau category was also included as a covariate.

⁴Adjusted mean is from natural cubic spline model (NCS2) that adjusted for NCS basis expansion terms, NCS basis expansion term-by-treatment interaction, baseline age, concomitant AChEI and/or memantine use at baseline, and pooled investigator. For the analysis of overall population, baseline tau category was also included as a covariate.

Source: CSR Tables AACI.5.6 - AACI.5.11; findings reproduced by the Statistical Analyst

Table 8 Analysis (Death Imputed as the Worst Score) of Clinical Endpoints at Week 76 in Evaluable Efficacy Set by Study Population

Endpoint ¹		Intermediate Tau Population (N=1182)	Overall Population (N=1736)
iADRS			
	Difference, adj. mean ³ (95% CI)	3.2 (1.3, 5.0)	2.9 (1.1, 4.6)
	p-value	p = 0.001	p = 0.001
CDR-SB			
	Difference, adj. mean ⁴ (95% CI)	-0.5 (-0.9, -0.2)	-0.6 (-0.9, -0.3)
	p-value	p = 0.001	p < 0.001
ADCS-iADL			
	Difference, adj. mean ³ (95% CI)	1.6 (0.5, 2.7)	1.6 (0.6, 2.5)
	p-value	p = 0.004	p = 0.002
ADAS-Cog13			
	Difference, adj. mean ³ (95% CI)	-1.0 (-2.2, 0.1)	-1.0 (-2.0, 0.0)
	p-value	p = 0.07 ^a	p = 0.04 ^a
CDR-G			
	HR ⁵ (95% CI)	0.7 (0.5, 0.9)	0.7 (0.5, 0.8)
	p-value	p = 0.01 ^a	p < 0.001 ^a

Abbreviations: CI = 95% confidence interval, N = number of randomized subjects

¹ Endpoints iADRS, CDR-SB, ADCS-iADL, and ADAS-Cog13 were change from baseline at Week 76. Endpoint CDR-G was time to substantial decline by Week 76.

³Adjusted mean is from natural cubic spline model (NCS2) that adjusted for NCS basis expansion terms NCS basis expansion term-by-treatment interaction, baseline age, concomitant AChEI and/or memantine use at baseline, and pooled investigator. For the analysis of overall population, baseline tau category was also included as a covariate.

⁴Adjusted mean is from Mixed Model with Repeated Measures (MMRM) that adjusted for treatment, visit, treatment-by-visit interaction, baseline iADRS score, baseline score-by-visit interaction, age at baseline, and concomitant AChEI and/or memantine use at baseline, and pooled investigator. For the analysis of overall population, baseline tau category was also included as a covariate.

⁵Time to substantial decline in CDR-G by Week 76 modeled by Cox regression adjusted for baseline age, score, and baseline concomitant AChEI or memantine use. The model was stratified by pooled investigator sites. For the analysis of overall population, the model was stratified by baseline tau category.

^ap-values are not significant based on the testing hierarchy

Source: CSR Table AACI.8.84, Table AACI.8.85; August 11, 2023, Sponsor Response to Information Request, Tables APP.1.-APP.6.

Table 9 Tipping Point Analysis of iADRS Change from Baseline at Week 76 in Evaluable Efficacy Set, Intermediate Tau Population

Delta for Placebo	Delta for LY3002813	N		p-value*
		Placebo	LY3002813	
0.00	0.00	583	569	.001
0.00	-1.00	583	569	.003
0.00	-2.00	583	569	.008
0.00	-3.00	583	569	.015
0.00	-4.00	583	569	.031
0.00	-5.00	583	569	.062
-1.00	0.00	583	569	<.001
-1.00	-1.00	583	569	.002
-1.00	-2.00	583	569	.004
-1.00	-3.00	583	569	.008
-1.00	-4.00	583	569	.017
-1.00	-5.00	583	569	.034
-1.00	-6.00	583	569	.064
-2.00	0.00	583	569	<.001
-2.00	-1.00	583	569	<.001
-2.00	-2.00	583	569	.002
-2.00	-3.00	583	569	.004
-2.00	-4.00	583	569	.009
-2.00	-5.00	583	569	.018
-2.00	-6.00	583	569	.035
-2.00	-7.00	583	569	.070

* P-value is from Cochran-Mantel-Haenszel raw mean score statistics.

Source: July 20, 2023, Sponsor Response to Information Request Table APP.2.2

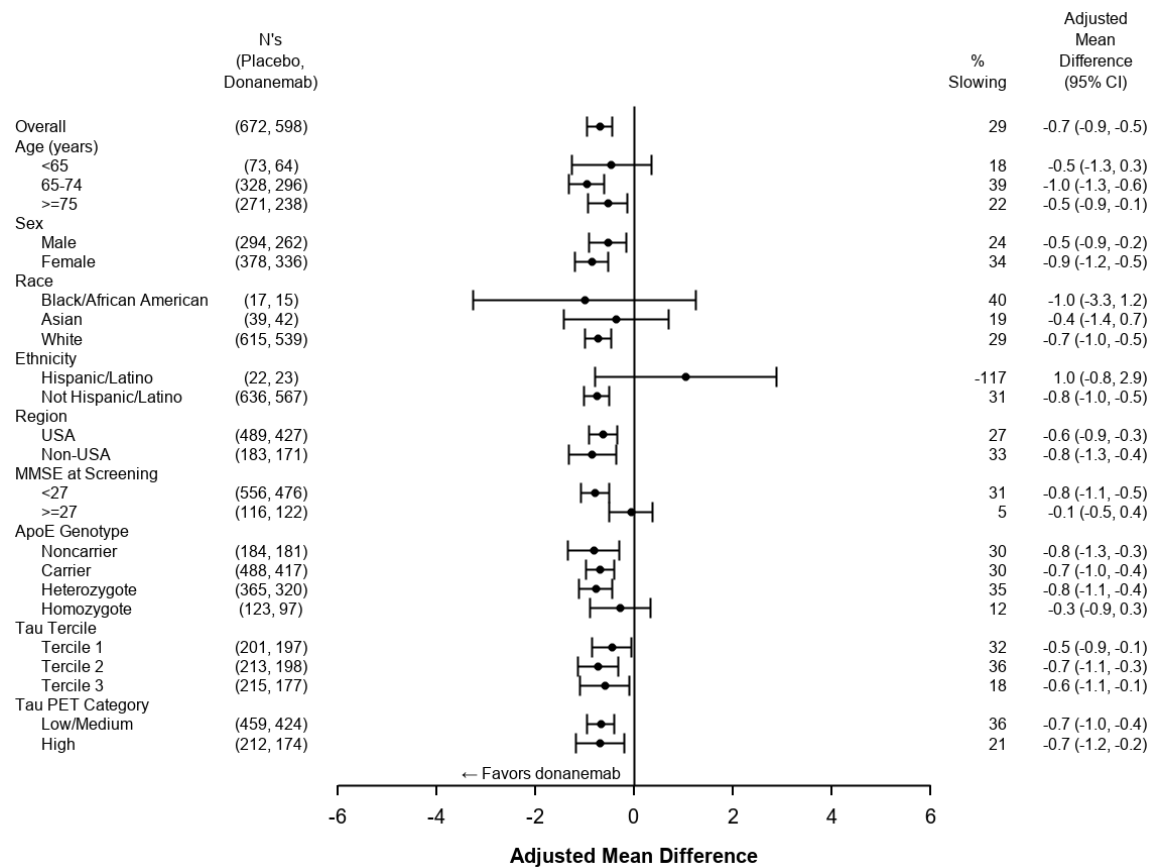
Table 10 Tipping Point Analysis of iADRS Change from Baseline at Week 76 in Evaluable Efficacy Set, Overall Population

Delta for Placebo	Delta for LY3002813	N		p-value*
		Placebo	LY3002813	
0.00	0.00	855	829	.001
0.00	-1.00	855	829	.003
0.00	-2.00	855	829	.008
0.00	-3.00	855	829	.019
0.00	-4.00	855	829	.041
0.00	-5.00	855	829	.085
-1.00	0.00	855	829	<.001
-1.00	-1.00	855	829	.002
-1.00	-2.00	855	829	.004
-1.00	-3.00	855	829	.009
-1.00	-4.00	855	829	.021
-1.00	-5.00	855	829	.046
-1.00	-6.00	855	829	.092
-2.00	0.00	855	829	<.001
-2.00	-1.00	855	829	<.001
-2.00	-2.00	855	829	.002
-2.00	-3.00	855	829	.005
-2.00	-4.00	855	829	.011
-2.00	-5.00	855	829	.025
-2.00	-6.00	855	829	.052

* P-value is from Cochran-Mantel-Haenszel raw mean score statistics.

Source: July 20, 2023, Sponsor Response to Information Request Table APP.2.1

Figure 14 Subgroup Analysis (MMRM) of CDR-SB Change from Baseline at Week 76 in Evaluable Efficacy Set, Overall Population



Abbreviations: CI = 95% confidence interval, PET = positron emission tomography, N= number of subjects who have a baseline and Week 76 efficacy score.

ApoE genotype (# of E4 alleles): Noncarrier (0), Heterozygote (1), Homozygote (2).

Tau Tercile: Tercile 1 (standard uptake value ratio (SUVR) <33 percentile), Tercile 2 (SUVR 33-67 percentile), Tercile 3 (SUVR >67 percentile).

Ethnicity subgroups are based on responses from both USA and non-USA regions.

Adjusted mean is from the MMRM that adjusted for visit, visit*treatment, baseline CDR-SB score, baseline score*visit, baseline age, baseline concomitant AChEI or memantine use, pooled investigator, and baseline tau PET category.

Source: Statistical Analyst

6.4 Model Diagnostics

To check the model assumptions of NCS2 and MMRM, Sponsor provided model diagnostics¹⁹ for the analysis of clinical endpoints iADRS, CDR-SB, ADCS-iADL, and ADAS-COG13. The diagnostics and a summary of their conclusions are as follows:

- Assessment of model fit and influential observations using a spaghetti plot of estimated mean change and subjects' score change over time
 - The mean change estimates appeared to represent most subjects' score change over time. Based on a qualitative visual check, there were no highly influential observations.
- Assessment of the assumption of normally distributed residual errors using a quantile-quantile plot
 - Residuals appeared approximately normal for most analyses, except for NCS2 analysis of iADRS and ADCS-iADL. However, these plots contained residuals from all time points, which were correlated. The sample sizes of the analyses (iADRS:1599, iADL:1606) were large and the residuals were expected (according to the Central Limit Theorem) to be normal as the sample size increased. Note that plotting the quantile-quantile plot for each scheduled visit, instead of combining all visits, would provide for a more useful diagnostic check.
- Assessment of the assumption of residual errors having a zero mean and constant variance using a plot of residuals versus fitted values for each scheduled visit. This assessment was conducted for iADRS and CDR-SB only.
 - For each scheduled visit, the residuals had an approximate zero mean and constant variance across fitted values.

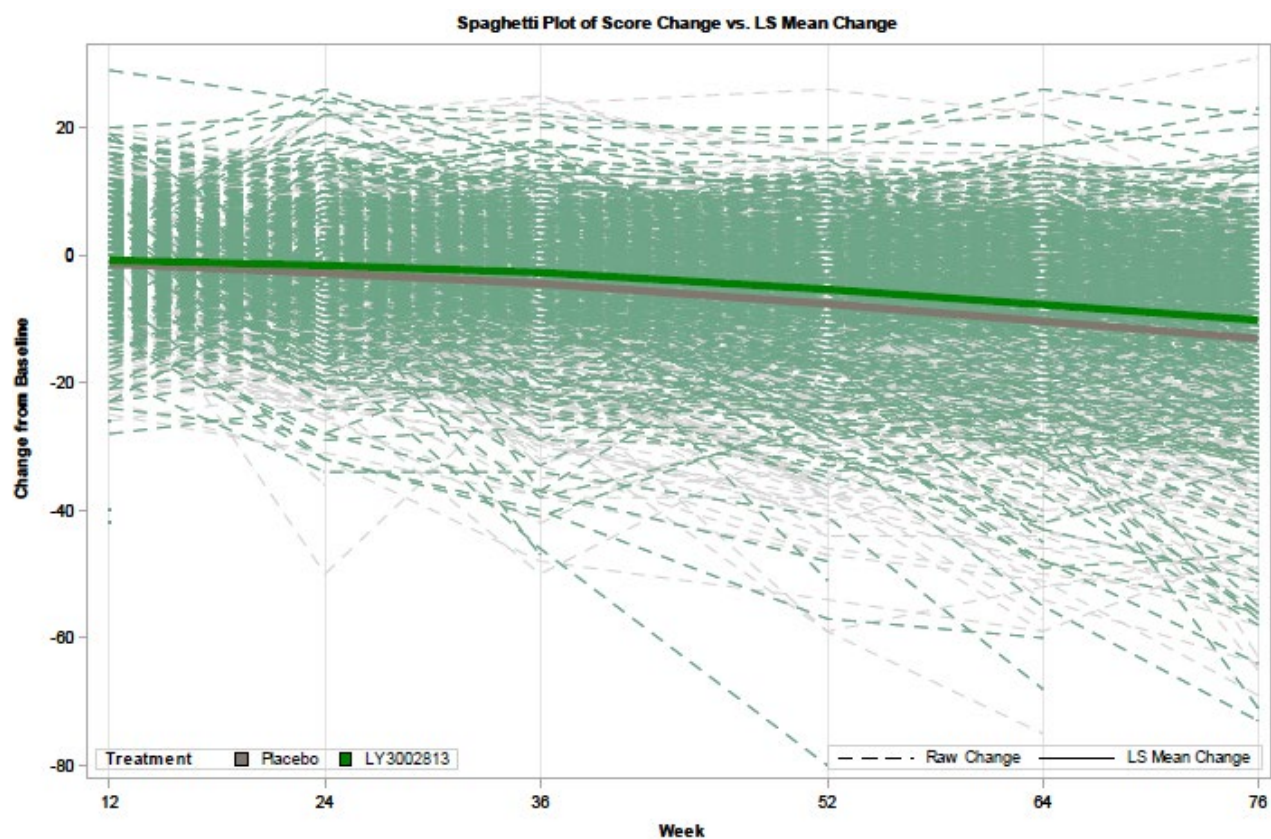
For illustrative purposes, we provided diagnostic plots in Figures 15-17 from the NCS2 analysis of iADRS change from baseline at Week 76 in the overall population. Diagnostics plots for the intermediate population and other endpoints were not shown.

Sponsor provided diagnostics²⁰ to assess the assumption of proportional hazards in the Cox regression analysis of time to substantial decline in CDR-G by Week 76. The diagnostics included a Kaplan-Meier plot and a log(-log(survival probability)) vs. log (covariate) plot for each categorical covariate. The plots did not indicate that the assumption of proportional hazards was violated (results not shown).

¹⁹ September 11, 2023, and October 25, 2023, Sponsor Response to Information Request. Available in DAARTS.

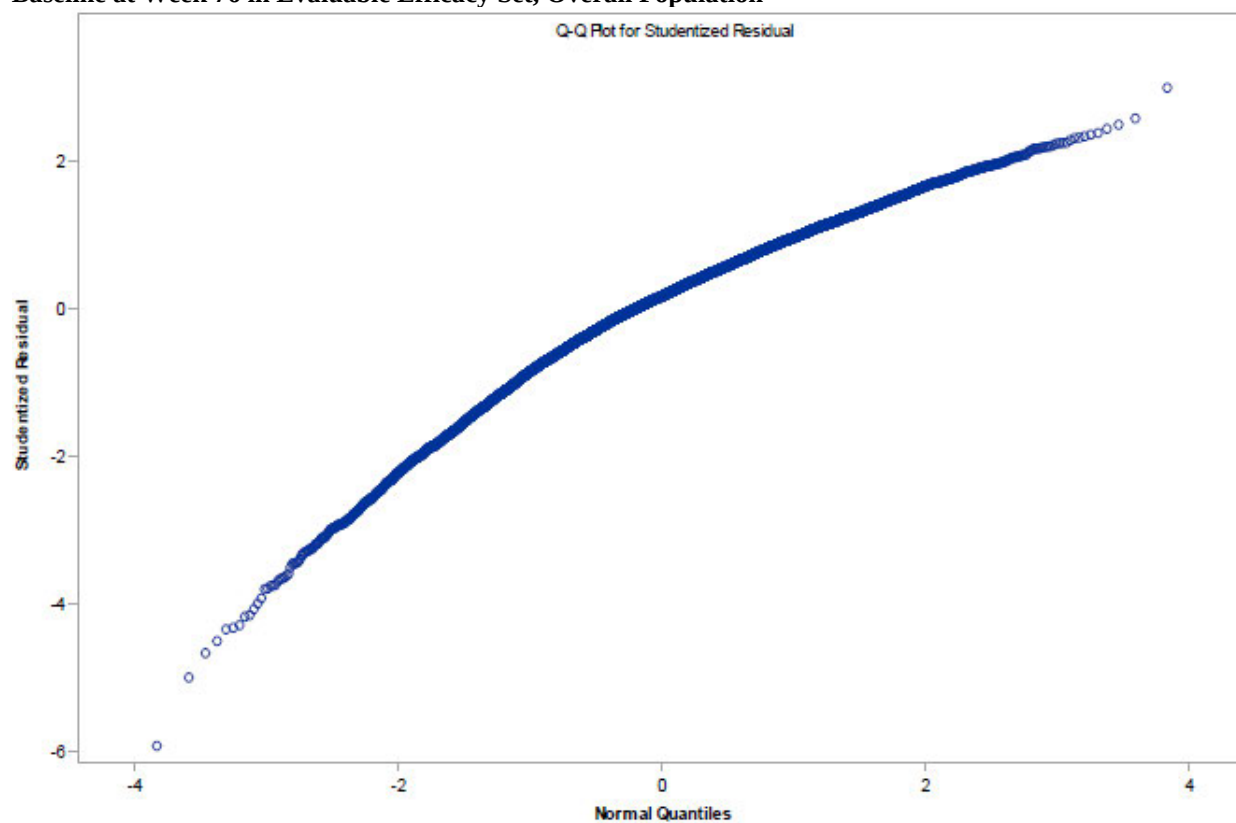
²⁰ July 13, 2023, Sponsor Response to Information Request. Available in DAARTS.

Figure 15 Subjects' Score Change and Estimated Mean Change from the Primary Efficacy Analysis (NCS2) of iADRS Change from Baseline at Week 76 in Evaluable Efficacy Set, Overall Population



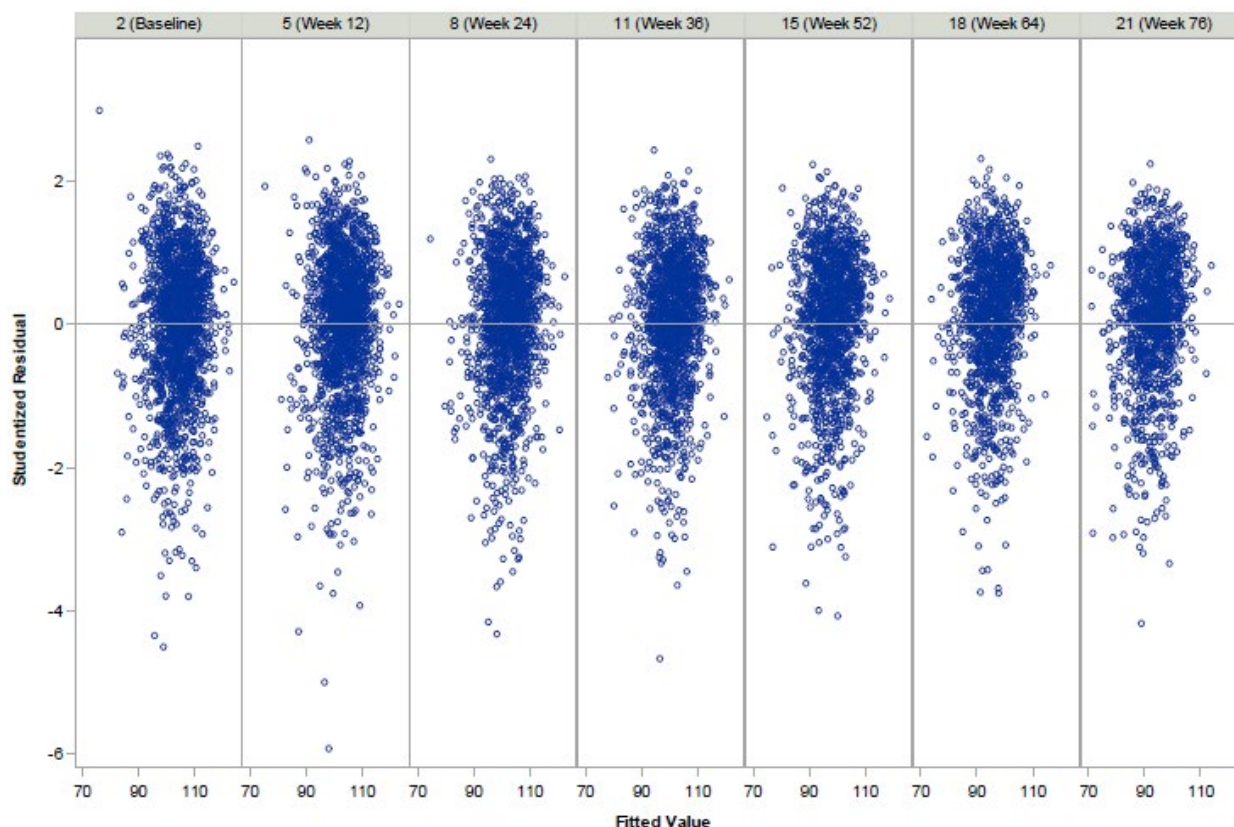
Source: September 11, 2023, Sponsor Response to Information Request.

Figure 16 Residual Quantile-Quantile Plot from Primary Efficacy Analysis (NCS2) of iADRS Change from Baseline at Week 76 in Evaluable Efficacy Set, Overall Population



Source: September 11, 2023, Sponsor Response to Information Request.

Figure 17 Residuals versus Fitted Values by Scheduled Visit from the Primary Efficacy Analysis (NCS2) of iADRS Change from Baseline at Week 76 in Evaluable Efficacy Set, Overall Population



Source: October 25, 2023, Sponsor Response to Information Request.

6.5 NCS2 Model Fit Estimates

For the NCS2 analysis of iADRS change from baseline at Week 76, the estimate of the covariance matrix showed increasing variance over time and high covariance between visits (Appendix Table 12). Sponsor posited²¹:

...in longitudinal follow-up studies of progressive diseases, like Study I5T-MC-AACI (AACI) in Alzheimer's disease (AD) as participants worsen over the time, often at different rates resulting in increasing variance over time. [...] Increasing covariance over time, specifically the covariances between baseline and other scheduled follow-up visits, is due to the setup of the NCS2 model [that did not adjust for baseline scores] in this study.

Sponsor conducted an exploratory analysis by including the interaction term of baseline iADRS score and the spline basis expansion functions. The covariance matrix (Appendix Table 13) with adjustment for baseline iADRS showed smaller estimates of variance (that still increased with time) and covariance compared to the pre-specified primary analysis estimate (Appendix Table 12). The covariance with adjustment for baseline iADRS decreased over time for the comparison of baseline to later visits, Week 12 to later visits, and Week 24 to later visits. With adjustment for baseline iADRS, the estimated difference in mean change from baseline of iADRS at Week

²¹ October 25, 2023, Sponsor Response to Information Request. Available in DAARTS.

76 in the overall population (2.7; 95% CI 1.4, 4.0; $p < 0.001$) was similar to the primary analysis result (2.9; 95% CI 1.5, 4.3; $p < 0.001$). Model fit and assumptions were not checked for this exploratory analysis.

Table 11 Covariate Estimates - Primary Efficacy Analysis (NCS2) of iADRS Change from Baseline at Week 76 in Evaluable Efficacy Set, Overall Population

Solution for Fixed Effects									
Effect	ACHEI and/or Memantine Use at Baseline	Pooled Investigator Group 1	Screening Tau Category	trtn_	Estimate	Standard Error	DF	t Value	Pr > t
Intercept					131.82	4.2712	1419	30.86	<.0001
spl_1					-0.06781	0.009797	1487	-6.92	<.0001
spl_2					-0.00233	0.000255	1388	-9.14	<.0001
spl_1*trtn_				1	-0.04976	0.01360	1471	-3.66	0.0003
spl_1*trtn_				2	0	.	.	.	.
spl_2*trtn_				1	0.000398	0.000349	1323	1.14	0.2542
spl_2*trtn_				2	0	.	.	.	.
AGE					-0.3836	0.05563	1418	-6.90	<.0001
ACHMEM	N				5.9714	0.7017	1421	8.51	<.0001
ACHMEM	Y				0	.	.	.	.
INVIDGR1		AUS			-3.8353	3.5215	1449	-1.09	0.2763
INVIDGR1		CAN			0.3588	3.3873	1407	0.11	0.9156
INVIDGR1		CZE			-23.4346	8.9631	1422	-2.61	0.0090
INVIDGR1		ENTID_	(b) (6)		7.0671	5.0250	1380	1.41	0.1598
INVIDGR1		ENTID_			-1.8205	3.9487	1399	-0.46	0.6449
INVIDGR1		ENTID_			-14.9837	4.6657	1348	-3.21	0.0014
INVIDGR1		ENTID_			-4.2196	3.4895	1384	-1.21	0.2268
INVIDGR1		ENTID_			4.5264	3.3744	1381	1.34	0.1800
INVIDGR1		ENTID_			-5.2273	3.1916	1400	-1.64	0.1017
INVIDGR1		ENTID_			-5.1578	3.3668	1371	-1.53	0.1258
INVIDGR1		ENTID_			-4.8407	6.1501	1417	-0.79	0.4314
INVIDGR1		ENTID_			3.3762	4.1418	1381	0.82	0.4151
INVIDGR1		ENTID_			-3.7032	3.6405	1419	-1.02	0.3092
INVIDGR1		ENTID_			-7.5356	2.7771	1408	-2.71	0.0067
INVIDGR1		ENTID_			9.1904	3.2682	1378	2.81	0.0050
INVIDGR1		ENTID_			0.6604	3.7876	1410	0.17	0.8616
INVIDGR1		ENTID_			-0.3459	3.0985	1400	-0.11	0.9111
INVIDGR1		ENTID_			3.0515	3.6368	1416	0.84	0.4016
INVIDGR1		ENTID_			-0.3246	5.4736	1357	-0.06	0.9527
INVIDGR1		ENTID_			1.0616	4.6922	1426	0.23	0.8210
INVIDGR1		ENTID_			-0.1456	4.1275	1343	-0.04	0.9719

Solution for Fixed Effects									
Effect	ACHeI and/or Memantine Use at Baseline	Pooled Investigator Group 1	Screening Tau Category	trtn_	Estimate	Standard Error	DF	t Value	Pr > t
INVIDGR1		ENTID_	(b) (6)		-0.2797	4.1218	1320	-0.07	0.9459
INVIDGR1		ENTID_			-3.2878	3.4380	1489	-0.96	0.3391
INVIDGR1		ENTID_			-14.5944	8.9664	1422	-1.63	0.1038
INVIDGR1		ENTID_			-2.5376	5.1256	1450	-0.50	0.6206
INVIDGR1		ENTID_			-4.8042	4.1466	1363	-1.16	0.2468
INVIDGR1		ENTID_			-14.6215	4.6970	1424	-3.11	0.0019
INVIDGR1		ENTID_			-6.2713	3.6485	1434	-1.72	0.0859
INVIDGR1		ENTID_			-7.3695	3.2519	1325	-2.27	0.0236
INVIDGR1		ENTID_			-2.7702	2.8642	1476	-0.97	0.3336
INVIDGR1		ENTID_			-2.0649	2.5763	1423	-0.80	0.4230
INVIDGR1		ENTID_			2.6767	3.6207	1377	0.74	0.4599
INVIDGR1		ENTID_			12.6317	5.4766	1357	2.31	0.0212
INVIDGR1		ENTID_			-11.1864	5.0619	1438	-2.21	0.0273
INVIDGR1		ENTID_			1.2294	3.9444	1382	0.31	0.7553
INVIDGR1		ENTID_			0.3409	2.7672	1404	0.12	0.9020
INVIDGR1		ENTID_			2.3831	3.1044	1389	0.77	0.4428
INVIDGR1		ENTID_			-2.2809	4.2740	1472	-0.53	0.5937
INVIDGR1		ENTID_			-2.6396	3.0193	1392	-0.87	0.3821
INVIDGR1		ENTID_			3.4435	4.3911	1369	0.78	0.4331
INVIDGR1		ENTID_			3.1196	4.1464	1364	0.75	0.4520
INVIDGR1		ENTID_			-4.7596	5.5041	1401	-0.86	0.3873
INVIDGR1		ENTID_			4.7650	3.3018	1422	1.44	0.1492
INVIDGR1		ENTID_			1.7599	5.0387	1384	0.35	0.7269
INVIDGR1		ENTID_			-1.6951	3.1704	1362	-0.53	0.5930
INVIDGR1		ENTID_			-7.6658	2.6347	1453	-2.91	0.0037
INVIDGR1		ENTID_			-1.4236	6.1226	1365	-0.23	0.8162
INVIDGR1		ENTID_			1.2070	2.1426	1395	0.56	0.5733
INVIDGR1		ENTID_			-1.7115	5.0423	1387	-0.34	0.7343
INVIDGR1		ENTID_			0.8702	3.3935	1418	0.26	0.7976
INVIDGR1		ENTID_			-0.2534	6.1055	1365	-0.04	0.9669
INVIDGR1		ENTID_			-2.2628	2.7694	1416	-0.82	0.4140
INVIDGR1		ENTID_			-3.4433	2.7723	1394	-1.24	0.2144
INVIDGR1		ENTID_			-4.3615	4.3978	1398	-0.99	0.3215
INVIDGR1		ENTID_			-15.8549	6.1062	1365	-2.60	0.0095
INVIDGR1		ENTID_			6.2712	3.9709	1433	1.58	0.1145
INVIDGR1		ENTID_			6.2082	5.0211	1352	1.24	0.2165
INVIDGR1		ENTID_			3.3658	2.6547	1389	1.27	0.2051

Solution for Fixed Effects									
Effect	ACHeI and/or Memantine Use at Baseline	Pooled Investigator Group 1	Screening Tau Category	trtn_	Estimate	Standard Error	DF	t Value	Pr > t
INVIDGR1		ENTID_	(b) (6)		-11.8046	4.7727	1462	-2.47	0.0135
INVIDGR1		ENTID_			-6.9942	6.1501	1417	-1.14	0.2556
INVIDGR1		ENTID_			-4.6665	4.2056	1461	-1.11	0.2674
INVIDGR1		ENTID_			-9.5848	3.6400	1422	-2.63	0.0085
INVIDGR1		ENTID_			-0.08532	5.0192	1367	-0.02	0.9864
INVIDGR1		ENTID_			1.9262	2.1916	1445	0.88	0.3796
INVIDGR1		ENTID_			-10.8272	4.7010	1427	-2.30	0.0214
INVIDGR1		ENTID_			2.8841	4.6739	1377	0.62	0.5373
INVIDGR1		ENTID_			-3.9746	4.3920	1393	-0.90	0.3656
INVIDGR1		ENTID_			-5.2958	4.7065	1437	-1.13	0.2607
INVIDGR1		ENTID_			8.0501	5.0363	1387	1.60	0.1102
INVIDGR1		ENTID_			7.3186	6.0803	1315	1.20	0.2289
INVIDGR1		ENTID_			-9.2004	6.0752	1314	-1.51	0.1302
INVIDGR1		ENTID_			-0.4035	4.4208	1444	-0.09	0.9273
INVIDGR1		ENTID_			9.8825	3.7774	1394	2.62	0.0090
INVIDGR1		ENTID_			2.1030	6.0740	1315	0.35	0.7292
INVIDGR1		ENTID_			2.2348	4.4208	1449	0.51	0.6133
INVIDGR1		ENTID_			2.0578	5.6470	1494	0.36	0.7156
INVIDGR1		ENTID_			4.6696	4.7066	1430	0.99	0.3213
INVIDGR1		ENTID_			-13.7922	3.9423	1380	-3.50	0.0005
INVIDGR1		ENTID_			0.1386	4.4454	1470	0.03	0.9751
INVIDGR1		ENTID_			1.4709	3.9686	1398	0.37	0.7110
INVIDGR1		ENTID_			2.5205	4.6479	1347	0.54	0.5877
INVIDGR1		ENTID_			-3.0576	3.5268	1423	-0.87	0.3861
INVIDGR1		ENTID_			2.5432	3.7985	1438	0.67	0.5033
INVIDGR1		ENTID_			2.6472	5.5185	1429	0.48	0.6315
INVIDGR1		ENTID_			2.2951	4.3632	1345	0.53	0.5990
INVIDGR1		ENTID_			-2.7008	5.0601	1434	-0.53	0.5936
INVIDGR1		ENTID_			-8.6880	4.3959	1371	-1.98	0.0483
INVIDGR1		ENTID_			-11.4800	4.1878	1460	-2.74	0.0062
INVIDGR1		ENTID_			-13.1748	6.2482	1494	-2.11	0.0351
INVIDGR1		ENTID_			-3.5148	5.5498	1453	-0.63	0.5266
INVIDGR1		ENTID_			-13.3269	5.5242	1434	-2.41	0.0160
INVIDGR1		ENTID_			-4.1010	4.3499	1319	-0.94	0.3460
INVIDGR1		ENTID_			3.3148	4.2151	1473	0.79	0.4318
INVIDGR1		ENTID_			3.5296	12.2841	1392	0.29	0.7739
INVIDGR1		ENTID_			-4.3555	3.4185	1466	-1.27	0.2028

Solution for Fixed Effects									
Effect	ACheI and/or Memantine Use at Baseline	Pooled Investigator Group 1	Screening Tau Category	trtn_	Estimate	Standard Error	DF	t Value	Pr > t
INVIDGR1		ENTID_	(b) (6)		-6.3016	5.0082	1352	-1.26	0.2085
INVIDGR1		ENTID_			1.9897	6.1309	1387	0.32	0.7456
INVIDGR1		ENTID_			1.8565	6.0726	1314	0.31	0.7599
INVIDGR1		ENTID_			0.5734	4.7541	1457	0.12	0.9040
INVIDGR1		ENTID_			7.1692	5.0283	1379	1.43	0.1542
INVIDGR1		ENTID_			-0.09186	6.9986	1313	-0.01	0.9895
INVIDGR1		ENTID_			8.2256	5.4972	1395	1.50	0.1348
INVIDGR1		ENTID_			-0.5796	3.8241	1467	-0.15	0.8795
INVIDGR1		ENTID_			-5.2559	5.0223	1366	-1.05	0.2955
INVIDGR1		ENTID_			-0.2577	5.0135	1351	-0.05	0.9590
INVIDGR1		ENTID_			-2.8640	6.1182	1365	-0.47	0.6398
INVIDGR1		ENTID_			5.7852	4.1370	1343	1.40	0.1622
INVIDGR1		ENTID_			0.9349	3.4371	1483	0.27	0.7856
INVIDGR1		ENTID_			2.9157	5.0477	1417	0.58	0.5636
INVIDGR1		ENTID_			-3.8067	2.6049	1448	-1.46	0.1441
INVIDGR1		ENTID_			4.1257	4.1473	1367	0.99	0.3200
INVIDGR1		ENTID_			7.5475	2.5649	1474	2.94	0.0033
INVIDGR1		ENTID_			3.5224	3.2627	1364	1.08	0.2805
INVIDGR1		ENTID_			6.4033	2.6228	1390	2.44	0.0148
INVIDGR1		ENTID_			8.9109	4.9936	1316	1.78	0.0746
INVIDGR1		ENTID_			0.1509	5.0398	1395	0.03	0.9761
INVIDGR1		ENTID_			-1.4901	4.1159	1320	-0.36	0.7174
INVIDGR1		ENTID_			8.7012	4.9926	1316	1.74	0.0816
INVIDGR1		ENTID_			6.9117	5.5524	1463	1.24	0.2134
INVIDGR1		ENTID_			0.4881	4.1790	1445	0.12	0.9070
INVIDGR1		ENTID_			8.5493	5.0725	1442	1.69	0.0921
INVIDGR1		ENTID_			8.6183	6.0780	1314	1.42	0.1564
INVIDGR1		ENTID_			3.4777	3.0350	1444	1.15	0.2520
INVIDGR1		ENTID_			4.3141	6.2094	1461	0.69	0.4873
INVIDGR1		ENTID_			3.6704	6.0772	1314	0.60	0.5460
INVIDGR1		ENTID_			-0.9799	5.0992	1447	-0.19	0.8476
INVIDGR1		ENTID_			5.3605	6.1414	1398	0.87	0.3829
INVIDGR1		ENTID_			-2.2193	4.1271	1343	-0.54	0.5909
INVIDGR1		ENTID_			-0.1225	6.2182	1483	-0.02	0.9843
INVIDGR1		ENTID_			-4.1244	5.5887	1489	-0.74	0.4606
INVIDGR1		ENTID_			6.5671	3.7890	1391	1.73	0.0833
INVIDGR1		ENTID_			7.8040	6.0804	1314	1.28	0.1996

Solution for Fixed Effects									
Effect	ACHEI and/or Memantine Use at Baseline	Pooled Investigator Group 1	Screening Tau Category	trtn_	Estimate	Standard Error	DF	t Value	Pr > t
INVIDGR1		ENTID_ (b) (6)			-4.4299	4.6540	1360	-0.95	0.3413
INVIDGR1		ENTID_			3.8344	2.4028	1395	1.60	0.1108
INVIDGR1		ENTID_			-0.7256	2.2219	1469	-0.33	0.7440
INVIDGR1		ENTID_			-1.4501	6.1215	1387	-0.24	0.8128
INVIDGR1		ENTID_			1.5447	4.3637	1319	0.35	0.7234
INVIDGR1		ENTID_			5.1254	6.1135	1365	0.84	0.4020
INVIDGR1		JPN			-3.8673	2.3042	1412	-1.68	0.0935
INVIDGR1		NLD			-0.8752	8.8145	1470	-0.10	0.9209
INVIDGR1		POL			14.7255	8.7443	1427	1.68	0.0924
INVIDGR1		USA			0	.	.	.	.
BLTAUCAT			High		-4.6592	0.7334	1424	-6.35	<.0001
BLTAUCAT			Low/Medium		0	.	.	.	.

Abbreviations: spl_1 = spline basis function 1, spl_2 = spline basis function 2, trtn_n = treatment (donanemab, placebo), AGE = baseline age, ACHMEM = concomitant acetylcholinesterase inhibitor (AChEI) and/or memantine use at baseline (yes/no), INVIDGR1 = pooled investigator, BLTAUCAT = Baseline tau category, DF = degrees of freedom

Table 12 Covariance Matrix Estimate - Primary Efficacy Analysis (NCS2) of iADRS Change from Baseline at Week 76 in Evaluable Efficacy Set for the Overall Population

Estimate of the Variance-covariance Matrix							
Timepoint	2 (Baseline)	5 (Week 12)	8 (Week 24)	11 (Week 36)	15 (Week 52)	18 (Week 64)	21 (Week 76)
2 (Baseline)	160.14						
5 (Week 12)	150.21	194.22					
8 (Week 24)	156.83	182.91	226.52				
11 (Week 36)	162.44	192.25	216.60	261.10			
15 (Week 52)	175.80	211.18	235.93	262.65	323.47		
18 (Week 64)	184.21	223.69	253.49	283.21	328.49	385.19	
21 (Week 76)	198.31	238.18	270.94	309.18	353.43	393.76	455.00

Source: Statistical Analyst

Table 13 Covariance Matrix Estimate - NCS2 Analysis of iADRS Change from Baseline at Week 76 Including Baseline Score Covariates in Evaluable Efficacy Set, Overall Population

Estimate of the Variance-covariance Matrix							
Timepoint	2 (Baseline)	5 (Week 12)	8 (Week 24)	11 (Week 36)	15 (Week 52)	18 (Week 64)	21 (Week 76)
2 (Baseline)	162.29						
5 (Week 12)	133.96	163.07					
8 (Week 24)	121.32	134.28	160.18				
11 (Week 36)	106.51	125.52	132.88	160.64			
15 (Week 52)	90.0948	117.60	124.81	135.29	168.14		
18 (Week 64)	73.5648	107.58	120.00	134.07	151.57	187.09	
21 (Week 76)	65.4108	101.59	116.61	139.46	154.86	173.63	211.91

Source: October 25, 2023, Response to Information Request.

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

/s/

THANH V TRAN
11/09/2023 12:22:59 PM

XIAOFENG WANG
11/09/2023 01:05:00 PM

JOHN P LAWRENCE
11/12/2023 04:54:57 PM

HSIEN MING J HUNG
11/13/2023 08:24:04 AM



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 #: 761248 (0002)

Drug Name: Kisunla (Donanemab)

Indication(s): Alzheimer's

Applicant: Eli Lilly

Date(s): September 27, 2021

Review Priority: Standard

Biometrics Division: I

Statistical Reviewer: Tristan Massie, Ph.D.

Concurring Reviewers: James (Hsien-Ming) Hung, Ph.D., Division Director

Medical Division: Division of Neurology I

Clinical Team: Kevin Krudys, Ph.D.
Ranjit Mani, M.D.

Project Manager: Andrew Papanastasiou

Keywords: Biomarker Accelerated Approval

Table of Contents

1	EXECUTIVE SUMMARY	5
2	INTRODUCTION	6
2.1	OVERVIEW	6
2.2	DATA SOURCES.....	7
3	STATISTICAL EVALUATION	7
3.1	DATA AND ANALYSIS QUALITY	7
3.2	EVALUATION OF EFFICACY	7
3.2.1.1	Study Design and Endpoints	7
3.2.1.2	Statistical Methodologies	7
3.2.1.3	Patient Disposition, Demographic and Baseline Characteristics	12
3.2.1.4	Results and Conclusions.....	15
3.2.1.4.1	Sponsor's Results	15
3.2.1.4.2	Reviewer's Results	18
3.3	EVALUATION OF SAFETY	19
4	FINDINGS IN SPECIAL/SUBGROUP POPULATIONS.....	20
4.1	GENDER, RACE, AGE, AND GEOGRAPHIC REGION	20
4.1.1	<i>Gender, Race, and Age</i>	20
4.1.2	<i>Geographic Region</i>	20
4.1.2.1	Individual Sites	20
4.2	OTHER SPECIAL/SUBGROUP POPULATIONS.....	21
5	SUMMARY AND CONCLUSIONS	21
5.1	STATISTICAL ISSUES	21
5.2	COLLECTIVE EVIDENCE	21
5.3	CONCLUSIONS AND RECOMMENDATIONS	22

LIST OF TABLES

Table 1 Double Blind Phase 2 Placebo Controlled Study Characteristics	6
Table 2 Study AACG Baseline Demographics and Characteristics	14
Table 3 Study AACG Primary Analysis of IADRS	16
Table 4 Study AACG MMRM Analysis of Amyloid Biomarker Results.....	18

LIST OF FIGURES

Figure 1 Study AACG Patient Disposition.....	13
Figure 2 Study AACG Amyloid Biomarker Results.....	17
Figure 3 Relationship between CDRSB and PET Amyloid Change at Week 76 for Donanemab monotherapy vs. Placebo by APOE.....	19

1 EXECUTIVE SUMMARY

The applicant submitted Study AACG to evaluate donanemab for treatment of Alzheimer's disease. In this double blind phase 2 study, the applicant's primary endpoint was an Alzheimer's disease composite, Integrated Alzheimer's Disease Rating scale (iADRS).

The study initially had three arms: placebo, LY3002813 monotherapy and LY3002813 in combination with LY3202626. The combination arm was stopped after 15 patients were randomized to it due to evolving evidence on the safety and efficacy of BACE inhibitors, like LY3202626, as a class.

As prespecified by the applicant in the original protocol, the trial was to be considered a success if LY3002813 alone and/or in combination with LY3202626 administered for up to 76 weeks decreased the cognitive and/or functional decline (based on iADRS) in patients with early symptomatic Alzheimer's Disease. Note that the Division had disagreed that the iADRS would be acceptable as a primary endpoint. At Week 76, donanemab-treated participants had statistically significantly less decline in cognition/function than placebo-treated participants as assessed by the iADRS, thus meeting the applicant's primary objective of the study. The treatment difference on LS mean change $\pm$ SE was 3.20 ± 1.56 ($p=.042$).

The applicant specified several key secondary endpoints including change from baseline in the following efficacy measures: ADAS-cog, CDRSB, MMSE and ADCS-iADL, followed by several other secondary endpoints including brain amyloid plaque deposition and change in brain tau deposition. A graphical multiplicity adjustment approach was specified to control for multiplicity of testing the primary iADRS and the ADAS-cog, CDRSB, MMSE and ADCS-iADL, but this plan did not include a pathway for multiplicity adjusted testing of change in brain amyloid deposition. All endpoints not included in this graphical approach should be considered exploratory.

None of the secondary endpoints, ADAS-cog, CDRSB, MMSE and ADCS-iADL, met the alpha allocated to them based on the pre-specified graphical approach, though ADAS-Cog13 suggested a treatment effect [LS mean change difference (SE) = -1.86 (0.898), nominal $p=.040$].

Among the secondary endpoints, the applicant focused in the application on the effect of donanemab compared to placebo on brain amyloid pathophysiology at Week 76 of treatment in subjects with Early Alzheimer's Disease as measured by amyloid positron emission tomography (PET). The mixed model for repeated measures analyses of change from baseline in brain amyloid levels as measured by PET SUVR normalized to whole cerebellum mask at Week 76 estimated a reduction of 85 centiloids for LY3002813-M compared to placebo (nominal p-value of <0.001).

The impact of amyloid reduction on the clinical outcome is uncertain. The correlation at the patient level is 0.098 between change in iADRS and change in Amyloid PET SUVR at Week 76 within the LY3002813 monotherapy arm. This correlation is small. Furthermore, there is no apparent treatment effect on the clinical endpoint in APOE4 non-carriers despite the numerically greater amyloid reductions compared to APOE4 carriers (see Figure 3).

2 INTRODUCTION

2.1 Overview

Donanemab is an IgG1 antibody directed at the N3pG A β epitope that is present only in brain amyloid plaques (Bridel et al. 2017). Donanemab binds to deposited amyloid plaque and aids its removal through microglial-mediated phagocytosis (DeMattos et al. 2012).

The IND for this drug development is IND 109157. There is one completed phase 2 study, study AACG. The primary endpoint was the IADRS at 18 months. The effect(s) on Amyloid imaging was a secondary endpoint.

Table 1 Double Blind Phase 2 Placebo Controlled Study Characteristics

Study Name	Phase and Design	Treatment Period	Follow-up Period	# of Subjects per Arm	Study Population
AACG	2	18 months Donanemab could be decreased to 700 or 0 if PET showed Amyloid Negativity at Week 24 or Week 52	18 months	ITT/PET PI N=126 Donanemab-M N=131 1400 mg Q4W Donanemab-C N=15 discontinued combination arm with a BACE inhibitor(LY3202626)	Alzheimer's

Note that a phase 3 Placebo-controlled, double-blind, parallel-group, 18-month study to confirm safety and efficacy of donanemab in subjects with early Alzheimer's disease involving approximately 1800 subjects is expected to complete in early 2023.

2.2 Data Sources

The amyloid imaging data set for the completed phase 2 study, AACG, is located as follows: \\CDSESUB1\evsprod\BLA761248\0002\m5\datasets\i5t-mc-aacg\analysis\adam\datasets. The primary endpoint <\\CDSESUB1\evsprod\BLA761248\0002\m5\datasets\i5t-mc-aacg\analysis\adam\datasets\adiadras.xpt>.

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

The submitted data and analysis quality appear adequate.

3.2 Evaluation of Efficacy

AACG

Phase 2, double-blind, placebo-controlled, study to evaluate the safety and efficacy of donanemab in patients with early symptomatic AD (prodromal AD and mild dementia due to AD), to assess whether removal of existing amyloid plaque can slow the progression of disease as assessed by clinical measures and biomarkers of disease pathology and neurodegeneration over up to 72 weeks of treatment, followed by final study assessments at 76 weeks.

18 December 2017 (first participant visit)

Study Completion: 04 December 2020 (last participant visit)

3.2.1.1 Study Design and Endpoints

3.2.1.2 Statistical Methodologies

Study AACG Statistical Analysis Plan dated Dec 7, 2021

Overview of Study Design

Study AACG was a multicenter, randomized, double-blind, placebo-controlled, Phase 2 study of donanemab (LY3002813) in participants with early symptomatic AD. The 133-week study

included

- a screening period of up to 9 weeks
- a treatment period of up to 72 weeks with final evaluations that occurred 4 weeks later at Week 76, and
- a 48-week immunogenicity and safety follow-up period.

Participants who met entry criteria were randomized in a 1:1 ratio to 1 of the following treatments:

- Donanemab (LY3002813): IV donanemab (700 mg Q4W for the first 3 doses, then 1400 mg Q4W) for up to 72 weeks, or
- Placebo: IV placebo Q4W for up to 72 weeks.

Note: Under previous versions of the protocol, as a third treatment group, donanemab was administered in combination with 12 mg of LY3202626, an inhibitor of BACE1. However, as of amendment (d), the combination therapy and the corresponding oral placebo for blinding of the oral LY3202626 were discontinued from the study. Participants in the donanemab-C group remained blinded and were allowed to continue in the study, receiving monthly infusions of donanemab without the co-administration of the oral agent.

Determination of Sample Size

Approximately 250 subjects were to be enrolled and randomized in a 1:1 ratio to the 2 treatment arms (placebo and LY3002813). It was expected that approximately 200 subjects would complete the double-blind treatment period of the study (approximately 100 per treatment arm). This sample size was expected to provide approximately 84% power to demonstrate that the active treatment arm had a ≥ 0.6 posterior probability of slowing down iADRS progression over placebo by at least 3 points. The assumptions for the power calculation are that mean progression levels in the placebo and LY3002813 arms are approximately 12 and 6 points (50% slowing) over 18 months, respectively, with common standard deviation of 17. If the active treatment arm is placebo-like with no efficacy, the probability of passing the efficacy criterion specified above (i.e., false positive) is approximately 6%. The simulation for the power calculation and sample size determination was carried out in FACTS Version 6.0.

Note that the sample size was originally determined under the originally planned 3 arm design (with the ultimately discontinued combination arm) as follows.

Approximately 375 subjects were to be enrolled and randomized in a 1:1:1 ratio to the three treatment arms (placebo, LY3002813-M, and LY3002813-C). It was expected that approximately 300 subjects would complete the double-blind treatment period of the study (approximately 100 per treatment arm). This sample size was expected to provide approximately 86% power to demonstrate that at least one active treatment arm has a ≥ 0.6 posterior probability of slowing down iADRS progression over placebo by at least 3 points. The assumption for this power calculation was that mean progression levels in placebo, LY3002813-M and LY3002813-C arms would be approximately 12, 8 (33% slowing), and 6 points (50% slowing) over 18 months, respectively, with common standard deviation of 17. If both active treatment arms were placebo-like with no efficacy, the probability of passing the efficacy criterion specified above (i.e., false positive) is approximately 11% for the originally planned three arm design.

Reviewer's Comment: The original protocol planned to evaluate whether either of two experimental arms showed a difference compared to placebo. As shown above the type I error rate of the original plan with three arms was approximately 11% and no multiplicity adjustment for declaring the study a success on the basis of either investigational agent arm was specified in the protocol. This study was originally planned in the context of a development plan including two follow-up phase 3 studies.

Full Analysis Set: All randomized patients to LY-M or placebo with a baseline and at least one post-baseline iADRS result

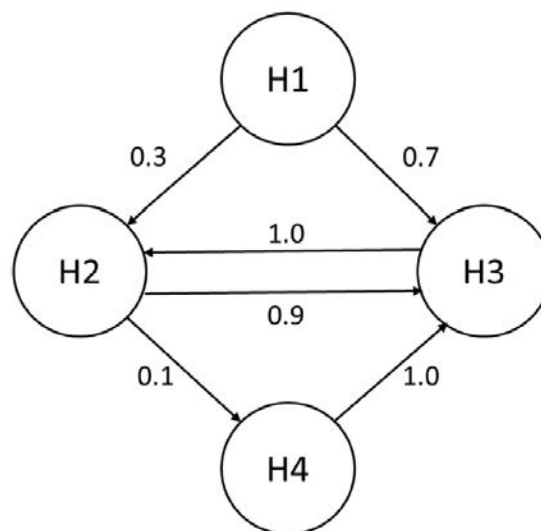
Primary Outcome and Methodology

The primary analysis is considered a treatment policy strategy as outlined by ICH E9 guidelines. The primary objective of this study is to test the hypothesis that IV infusion of donanemab slows the cognitive and/or functional decline of AD as measured by the composite measure iADRS compared with placebo in patients with early symptomatic AD. This was to be assessed using an MMRM analysis comparing LY-M vs. PL. The change from baseline score on the iADRS at each scheduled postbaseline visit during the treatment period was to be the dependent variable. The model for the fixed effects was to include the following terms: baseline score, pooled investigator, treatment, visit, treatment-by-visit interaction, baseline-by-visit interaction, concomitant AChEI and/or memantine use at baseline (yes/no), and age at baseline. Visit was to be considered a categorical variable. The null hypothesis is that the contrast between the donanemab group versus placebo at the last visit equals 0. An unstructured covariance matrix was to be used to model the within subject variance-covariance errors. If the unstructured covariance structure matrix resulted in a lack of convergence, the following tests were to be used in sequence: heterogeneous Toeplitz covariance structure, heterogeneous autoregressive covariance structure, heterogeneous compound symmetry covariance structure, compound symmetry covariance structure. The Kenward-Roger approximation was to be used to estimate the denominator degrees of freedom. The primary time point for treatment comparison was to be at Week 76. The treatment group contrast in least-squares mean progression and its associated p-

value and 95% CI were to be calculated for the treatment comparison of donanemab versus placebo using the MMRM model specified above. In addition, Bayesian posterior probability of the active treatment arm being superior to placebo by at least a margin of interest (25% slowing of placebo progression) was also to be calculated using a non-informative prior.

Gated Secondary Efficacy Endpoints

Bretz's graphical approach was to be utilized to provide strong control of the study-wise type I error rate for the primary and key secondary hypotheses at alpha level of 0.05 (Bretz et al. 2009, 2011). Assuming the primary analysis is statistically significant, the MMRM analyses described for the primary analysis were to be conducted on the CDR-SB, ADAS-Cog13, iADL and MMSE scores and statistical significance was to be determined based on the following multiplicity graph of hypotheses regarding the IV infusion of donanemab slowing the cognitive and/or functional decline of AD:



- H1. MMRM analysis of CDR-SB at Week 76
- H2. MMRM analysis of ADAS-Cog13 total score at Week 76
- H3. MMRM analysis of iADL subscore at Week 76
- H4. MMRM analysis of MMSE at Week 76

Analysis of Amyloid PET

At baseline, standardized uptake value ratio (SUVr) and centiloid value were to be calculated as a ratio of the composite summary region that is an average of 6 different cortical regions (anterior cingulate, posterior cingulate, medial orbital frontal, lateral temporal, lateral parietal, precuneus) with whole cerebellum as a reference region (SUVRCAL). The SUVRCAL was also

used at visits 8 and 15 in order to determine if the amyloid level had been reduced enough in LY-treated patients to reduce the dose of LY.

To evaluate the amyloid PET longitudinal change from baseline (SUVRCLAA), an MMRM analysis was to be used to compare change from baseline in SUVR at 76 weeks in the Evaluable Efficacy dataset. The model was to include the fixed, categorical effects of treatment, visit, and treatment-by-visit interaction, as well as continuous effects of baseline SUVR and age at baseline. Visit was to be considered a categorical variable with values equal to the visit numbers at which tau imaging was assessed. The null hypothesis is that the difference in LSM between the donanemab and placebo equals zero. Annualized change in the composite summary SUVRCLAA for each patient was to be calculated using the change at the last post-baseline visit.

The annualized change was to be compared between the treatment groups with an ANCOVA. The ANCOVA model was to include the fixed effect of treatment as well as continuous effects of baseline amyloid PET and age at baseline. The null hypothesis is that the difference in LSM between donanemab and placebo equals zero.

Spearman's rank correlation coefficient was to be calculated to assess the relationship of change in amyloid PET at Week 76 to change from baseline to Week 76 in iADRS, CDR-SB, ADCSiADL and MMSE by treatment. Partial correlation analyses were to be conducted using only patients who have the amyloid PET and scale result at Week 76 and adjusted for APOE4 carrier (Y/N), age and sex.

Additionally, a linear regression analysis of change in iADRS versus change in amyloid PET endpoint in centiloids (SUVRCLAA) was to be graphed. The graph was to include the regression line, p-value for the explanatory variable (amyloid PET), and color to differentiate patients who are 'amyloid negative' (amyloid PET endpoint < 24.1 centiloids). Change and annualized change analyses described above were to be repeated on the Full Analysis Set (Pooled LY). Treatment comparisons were to be made between LY (pooled LY-M and LY-C) and placebo.

Analysis of Tau PET

To evaluate the change from baseline in tau imaging parameters, an MMRM analysis was to be used to compare change from baseline in global tau load measurement at 76 weeks in the Evaluable Efficacy dataset. The model was to include the fixed, categorical effects of treatment, visit, and treatment-by-visit interaction, as well as continuous effects of baseline global tau load measurement and age at baseline. Visit was to be considered a categorical variable with values equal to the visit numbers at which tau imaging is assessed. The null hypothesis is that the difference in LSM between donanemab and placebo equals zero. The same MMRM analysis was to be conducted on the tau PET MUBADA.

Change from baseline and annualized change from baseline analyses were to be conducted on global tau load measurement computed from the Tau IQ algorithm. The annualized change was to be compared between the treatment groups with an ANCOVA on the full efficacy dataset. The ANCOVA model was to include the fixed effect of treatment as well as continuous effects of baseline global tau PET value and age at baseline. The null hypothesis is that the difference in LSM between donanemab and placebo equals zero.

To assess the relationship of biomarker with cognition and function by treatment, Spearman's rank correlation coefficient was to be obtained on change from baseline to Week 76 for the global tau load computed from the Tau IQ algorithm and with change from baseline to Week 76 for CDR-SB, iADRS, ADAS-Cog13, ADCS-iADL, and MMSE. Partial correlation analyses by treatment group were to be conducted using only patients who have the clinical outcome and global tau load measurement at Week 76; adjusted for APOE4 carrier (Y/N), age and sex.

Change and annualized change analyses described above were to be repeated on the Full Analysis Set (Pooled LY). Treatment comparisons were to be made between LY (pooled LY-M and LY-C) and placebo.

3.2.1.3 Patient Disposition, Demographic and Baseline Characteristics

Patient Disposition

LY3002813 in combination with LY3202626

Study AACG was originally designed as a three arm study comparing LY3002813 alone (LY-M) and in combination with LY3202626 (LY-C) versus placebo (PL). Study AACG was amended to drop the LY-C arm after a small number of patients had been enrolled to one of those three arms. Patients randomized to LY-C were allowed to continue in the study getting monthly infusions of LY3002813 without the co-administration of the oral LY3202626. For this clinical study report, all analyses of the Randomized Population include all three arms (LY-M, LY-C, and PL).

Study disposition

In this study, it was possible to discontinue treatment but remain in the study.

A total of 187 participants (donanemab-M, n = 94; placebo, n = 93) completed the study. A total of 69 participants were discontinued from the study (donanemab-M, n = 37; placebo, n = 32).

A statistically significantly higher proportion of participants discontinued from the study due to AEs in the donanemab-M group (15.3%) than in the placebo group (4.8%) ($p=.007$). [Figure 1]

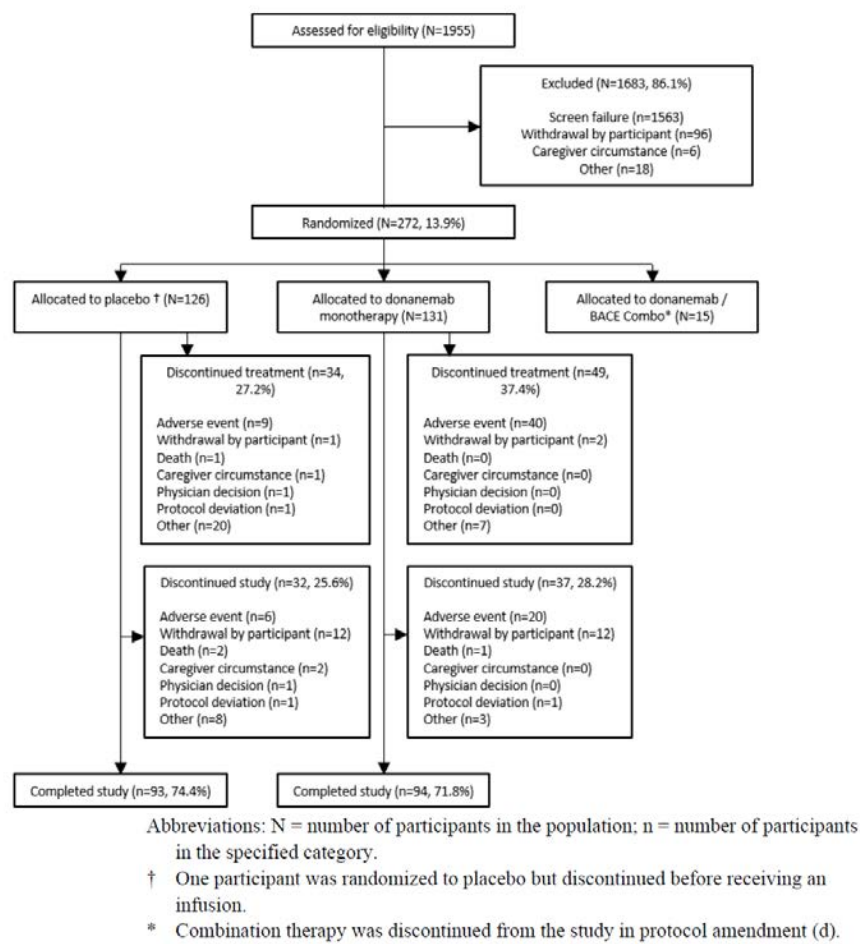
Treatment disposition

A total of 173 participants (donanemab-M, n = 82; placebo, n = 91) completed treatment. A total of 83 participants were discontinued from treatment (donanemab-M, n = 49; placebo, n = 34).

As shown in Figure 1, a statistically significantly higher proportion of participants discontinued from treatment due to AEs in the donanemab-M group (30.5%) than in the placebo group (7.2%) ($p<.001$). A statistically significantly lower proportion of participants discontinued from treatment due to other reasons in the donanemab-M group (5.3%) than in the placebo group

(16.0%) (p=.007).

Figure 1 Study AACG Patient Disposition



Note: Figure copied from page 38 of sponsor’s study report

Baseline Demographics and Characteristics

Baseline demographics for the placebo and donanemab-M groups, respectively, for

- mean age were 75.4 and 75.0 years
- proportions female sex were 51.6% and 51.9%
- proportions white race were 96.0% and 93.1%, and
- proportions Apo E4 carrier were 74.2% and 72.5%.

Table 2 presents the demographics and other baseline characteristics. Additionally, 67.9% of donanemab-M-treated participants and 65.9% of placebo-treated participants were on stable doses of acetylcholinesterase inhibitors or memantine at study baseline. At the time of study initiation, the study consisted of 3 groups, including a combination group of donanemab with BACE1 inhibitor. This group was discontinued early in the study, with 15 participants

randomized to that group. In the safety population, 125 participants received placebo and 131 participants received donanemab-M.

There were no noticeable imbalances in baseline demographics characteristics between the treatment groups.

Table 2 Study AACG Baseline Demographics and Characteristics

	Placebo (N = 126)	Donanemab-M (N = 131)	Donanemab-C (N = 15)	Total [†] (N = 272)
Demographics				
Female sex, n (%)	65 (51.6)	68 (51.9)	12 (80.0)	145 (53.3)
Mean age, years (SD)	75.4 (5.4)	75.0 (5.6)	75.2 (6.0)	75.2 (5.5)
Race, n (%)				
Asian	2 (1.6)	1 (0.8)	0 (0)	3 (1.1)
Black or African American	3 (2.4)	5 (3.8)	0 (0)	8 (2.9)
White	121 (96.0)	122 (93.1)	15 (100.0)	258 (94.9)
Other*	0 (0)	3 (2.3)	0 (0)	3 (1.1)
Ethnicity, Hispanic/Latino [#] , n (%)	3 (2.4)	5 (3.8)	1 (6.7)	9 (3.3)
Education, ≥13 years, n (%)	102 (81.0)	97 (74.0)	10 (66.7)	209 (76.8)
APOE4 carrier, n/N (%)	92/124 (74.2)	95/131 (72.5)	10/15 (66.7)	197/270 (73.0)
E2/E3, n (%)	1 (0.8)	1 (0.8)	0 (0)	2 (0.7)
E2/E4, n (%)	2 (1.6)	2 (1.5)	0 (0)	4 (1.5)
E3/E3, n (%)	31 (25.0)	35 (26.7)	5 (33.3)	71 (26.3)
E3/E4, n (%)	62 (50.0)	68 (51.9)	7 (46.7)	137 (50.7)
E4/E4, n (%)	28 (22.6)	25 (19.1)	3 (20.0)	56 (20.7)
AChEI use, n (%)	74 (58.7)	78 (59.5)	10 (66.7)	162 (59.6)
Memantine use, n (%)	28 (22.2)	31 (23.7)	1 (6.7)	60 (22.1)
AChEI and/or Memantine use, n (%)	83 (65.9)	89 (67.9)	10 (66.7)	182 (66.9)
Scale, Mean (SD), range				
iADRS	105.9 (13.2), 67.0–139.0	106.2 (13.0) ^a , 60.0–130.0	108.9 (10.2), 88.0–123.0	106.2 (13.0) ^b , 60.0–139.0
CDR-SB	3.4 (1.7), 0.5–8.0	3.6 (2.1), 0.5–11.0	3.6 (2.1), 1.5–8.0	3.5 (1.9), 0.5–11.0
ADAS-Cog ₁₃	27.5 (7.6), 5.0–47.0	27.6 (7.7), 10.0–51.0	27.3 (7.3), 13.0–39.0	27.6 (7.6), 5.0–51.0
ADCS-ADL	67.0 (8.1), 40.0–78.0	67.4 (8.6) ^a , 28.0–78.0	69.9 (5.6), 56.0–78.0	67.3 (8.2) ^b , 28.0–78.0
ADCS-iADL	48.4 (7.5), 24.0–59.0	48.9 (7.6) ^a , 21.0–59.0	51.2 (5.5), 38.0–59.0	48.8 (7.5) ^b , 21.0–59.0
MMSE	23.7 (2.9) ^c , 16.0–29.0	23.6 (3.1) ^d , 14.0–29.0	21.5 (5.0) ^e , 13.0–30.0	23.5 (3.1) ^f , 13.0–30.0

	Placebo (N = 126)	Donanemab-M (N = 131)	Donanemab-C (N = 15)	Total [†] (N = 272)
Amyloid PET Centiloids, Mean (SD), range	101.1 (33.3), 38.7–225.2	107.6 (36.0), 41.0–251.4	99.2 (35.6), 42.1–167.1	104.2 (34.8), 38.7–251.4
Flortaucipir PET global tau load, Mean (SD), range	0.46 (0.15) ^g , 0.2–0.9	0.47 (0.19) ^h , 0.1–1.2	0.45 (0.15), 0.2–0.7	0.46 (0.17) ⁱ , 0.1–1.2

Abbreviations: APOE4 = apolipoprotein E allele 4; AChEI = acetylcholinesterase inhibitor;

ADAS-Cog₁₃ = Alzheimer's Disease Assessment Scale – Cognitive 13-item subscale;

ADCS-ADL = Alzheimer's Disease Cooperative Study – Activities of Daily Living scale;

ADCS-iADL=Alzheimer's Disease Cooperative Study – instrumental Activities of Daily Living subscale;

iADRS = integrated Alzheimer's Disease Rating Scale; MMSE = Mini-Mental State Examination;

CDR-SB = Clinical Dementia Rating Scale – Sum of Boxes; PET = positron emission tomography; N = number of participants in the population; n = number of participants in the specified category; SD = standard deviation.

a Donanemab monotherapy N = 130.

b Total N = 271

c Placebo N = 121

d Donanemab monotherapy N = 126

e Donanemab/BACE Combo N = 14

f Total N = 261

g Placebo N = 124

h Donanemab monotherapy N = 130

i Total N = 269.

* Includes Multiple & American Indian or Alaska Native.

† Includes participants in the combo group.

Number of participants with nonmissing data, used as denominator

Note: This table was copied from page 41-42 of sponsor's study report

3.2.1.4 Results and Conclusions

3.2.1.4.1 Sponsor's Results

MMRM Analysis of iADRS

The primary objective of this study was to test the hypothesis that IV infusions of donanemab administered up to 72 weeks in the double-blind treatment period would slow the cognitive and functional decline of AD as measured by the composite measure iADRS (Wessels et al. 2015) at 76 weeks compared with placebo in participants with early symptomatic AD. The primary statistical analysis is using an MMRM analysis comparing donanemab-M versus placebo.

Table 3 summarizes the MMRM analysis of change from baseline on the iADRS. At Week 76, donanemab-treated participants had statistically significantly less decline in cognition/function than placebo-treated participants as assessed by the iADRS, thus meeting the primary objective of the study. The LS mean change difference $\pm$ SE was 3.20 ± 1.56 ($p=.042$). The analysis shows an approximately 32% reduction in cognitive/functional decline for donanemab-treated participants compared with placebo.

Table 3 Study AACG Primary Analysis of IADRS

Analysis Value				Treatment Difference (vs. PLACEBO)					
Treatment	Timepoint	n	Mean (SD)	LS Mean Change (SE)	LS Mean Difference (SE)	95% CI	p-value	Effect Size	Percent of Reduction
Placebo (N=120)	Baseline	120	106.06 (13.050)						
	5 (Week 12)	113	104.68 (13.980)	-0.07 (0.684)					
	8 (Week 24)	110	104.45 (13.952)	-1.18 (0.666)					
	11 (Week 36)	103	102.57 (16.259)	-3.17 (0.730)					
	15 (Week 52)	90	99.26 (16.080)	-6.70 (0.929)					
	18 (Week 64)	90	98.32 (17.864)	-8.34 (1.038)					
	21 (Week 76)	91	96.69 (18.593)	-10.06 (1.141)					
LY3002813-M (N=125)	Baseline	125	106.30 (12.730)						
	5 (Week 12)	120	105.45 (13.881)	0.42 (0.673)	0.49 (0.877)	(-1.24, 2.22)	0.575	0.074	665.759
	8 (Week 24)	112	105.34 (13.922)	0.12 (0.663)	1.30 (0.854)	(-0.38, 2.99)	0.129	0.204	110.394
	11 (Week 36)	102	103.91 (14.248)	-0.73 (0.732)	2.44 (0.956)	(0.55, 4.33)	0.011	0.351	77.020
	15 (Week 52)	88	102.22 (15.927)	-3.03 (0.933)	3.67 (1.256)	(1.19, 6.15)	0.004	0.415	54.760
	18 (Week 64)	89	101.16 (17.126)	-4.92 (1.038)	3.42 (1.415)	(0.63, 6.21)	0.017	0.345	40.969
	21 (Week 76)	93	99.29 (17.494)	-6.86 (1.135)	3.20 (1.560)	(0.12, 6.27)	0.042	0.289	31.793

Variance-covariance structure = Unstructured.

Abbreviations: CI = confidence interval; LS Mean = least-squares mean; n = number of patients at each visit with non-missing values; N = number of patients in the full analysis set; SD = standard deviation; SE = standard error.

LSmean change from baseline, SE, 95% CI and p-value are derived using MMRM with factors for treatment, visit, treatment-by-visit interaction, pooled investigator, AChEI and/or memantine use at baseline, and covariates for baseline score, age at baseline, and baseline score-by-visit interaction.

Effect Size = Treatment difference from LS Mean Difference/SD from VAR-COV matrix at each timepoint.

Percent of Reduction = ((LS Mean Change of placebo - LS Mean Change of LY)/LS Mean Change of placebo) *100.

Program: /lillyce/prd/ly3002813/i5t_mc_aacg/intrm3/programs/primary/tfl1/t_rm_iadrs13_fas.sas

Output: /lillyce/prd/ly3002813/i5t_mc_aacg/intrm3/output/shared/t_rm_iadrs13_fas.rtf

Dataset: /lillyce/prd/ly3002813/i5t_mc_aacg/intrm3/data/analysis/shared/adam/adsl, adiadras.sas7bdat.

Note: This Table was copied from page 50 of Sponsor's study report

For the key secondary endpoints at Week 76 LS Mean differences compared to placebo analyzed by MMRM were as follows.

CDRSB: -0.36 (S.E.=0.239), p=0.139.

ADAS-Cog13: -1.86 (S.E.= 0.898), p=.040

ADCS-iADL: 1.21 (SE= 1.009), p=.230

MMSE: 0.64 (SE = 0.525), p=.227.

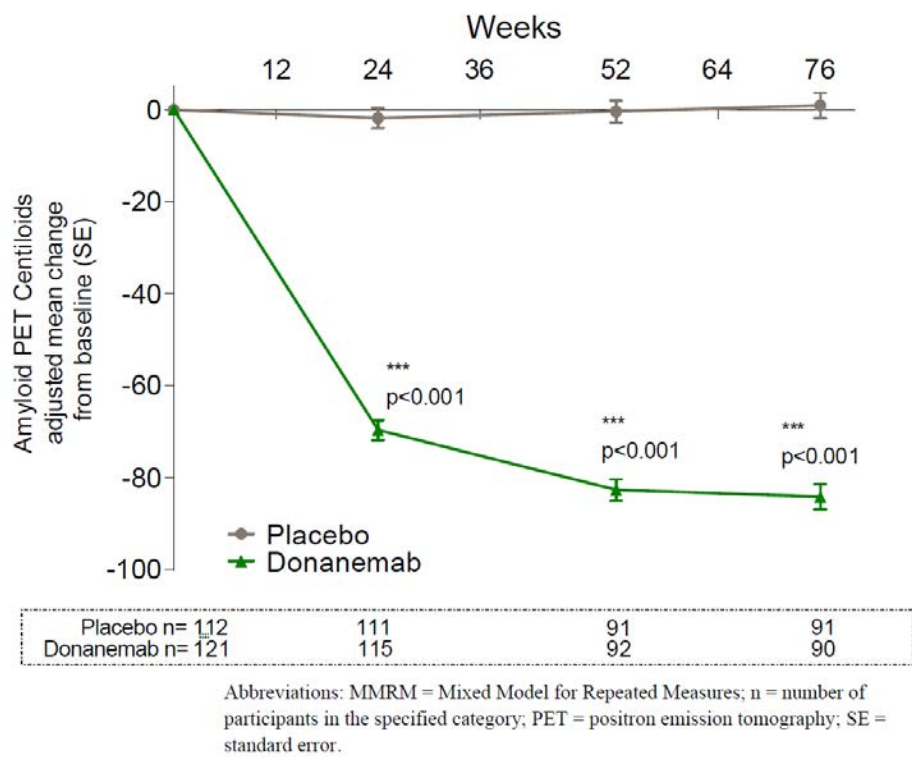
According to the prespecified graphical approach to account for multiplicity of the four key secondary endpoints, none of these endpoints are statistically significant when adjusted for multiplicity. The adjusted p-value for ADAS-Cog13 according to the prespecified graphical approach is p=0.139.

Analysis of Amyloid PET

Table 4 summarizes the MMRM analysis of change from baseline on the amyloid PET Centiloid measurement. At Week 76, donanemab-treated participants had statistically significantly less A β in the brain than placebo-treated participants as assessed by the amyloid PET Centiloid measurement (Figure 2). Slight increases in amyloid levels were observed in placebo-treated participants. The LS mean change difference $\pm$ SE was -85.06 $\pm$ 3.867, p<.001 (Table 4). The MMRM analysis for the amyloid PET Centiloid measurement also showed statistically

significantly lesser A β in the brain of donanemab-treated participants than placebo-treated participants at Week 24 and Week 52.

Figure 2 Study AACG Amyloid Biomarker Results



Note: This figure was copied from page 70 of sponsor's study report

Table 4 Study AAG MMRM Analysis of Amyloid Biomarker Results

Treatment	Timepoint	n	Analysis Value		Treatment Difference (vs. Placebo)				
			Mean (SD)	LS Mean Change (SE)	LS Mean Change Difference (SE)	95% CI	p-value	Effect Size	Percent of Reduction
Placebo (N=112)	Baseline	112	103.09 (33.841)						
	8 (Week 24)	111	102.08 (36.607)	-1.82 (2.251)					
	15 (Week 52)	91	102.97 (32.818)	-0.35 (2.420)					
	21 (Week 76)	91	103.86 (35.046)	0.93 (2.739)					
LY3002813-M (N=121)	Baseline	121	107.18 (33.938)						
	8 (Week 24)	115	37.07 (30.934)	-69.64 (2.206)	-67.83 (3.155)	(-74.04, -61.61)	<0.001	-2.863	-3732.020
	15 (Week 52)	92	24.71 (29.459)	-82.65 (2.392)	-82.30 (3.408)	(-89.02, -75.59)	<0.001	-3.385	-23778.416
	21 (Week 76)	90	21.27 (28.769)	-84.13 (2.723)	-85.06 (3.867)	(-92.68, -77.43)	<0.001	-3.109	9176.493

Variance-covariance structure = Unstructured.

Abbreviations: CI = confidence interval; LOCF = last observation carried forward; LS Mean = least-squares mean; n = number of patients at each visit with non-missing values; N = number of patients in the full analysis set; SD = standard deviation; SE = standard error. LS mean change from baseline, SE, 95% CI and p-value are derived using mixed model repeated measures methodology with factors for treatment, visit, treatment-by-visit interaction, and covariates for baseline score and age at baseline. Effect Size = Treatment difference from LS Mean Difference/SD from VAR-COV matrix at each timepoint. Percent of Reduction = ((LS Mean Change of placebo - LS Mean Change of LY)/LS Mean Change of placebo) *100.

Note: This table was copied from page 70 of sponsor's study report

On the SUVR scale the LY3002814-M difference from placebo in Amyloid change was -0.47 at Week 76.

There was not a statistically significant correlation between change from baseline to endpoint in amyloid PET biomarker values and change from baseline to endpoint in clinical efficacy scales. Within the LY3002813 arm the sponsor estimated a patient level partial correlation of 0.098 between change in iADRS and Amyloid change at Week 76.

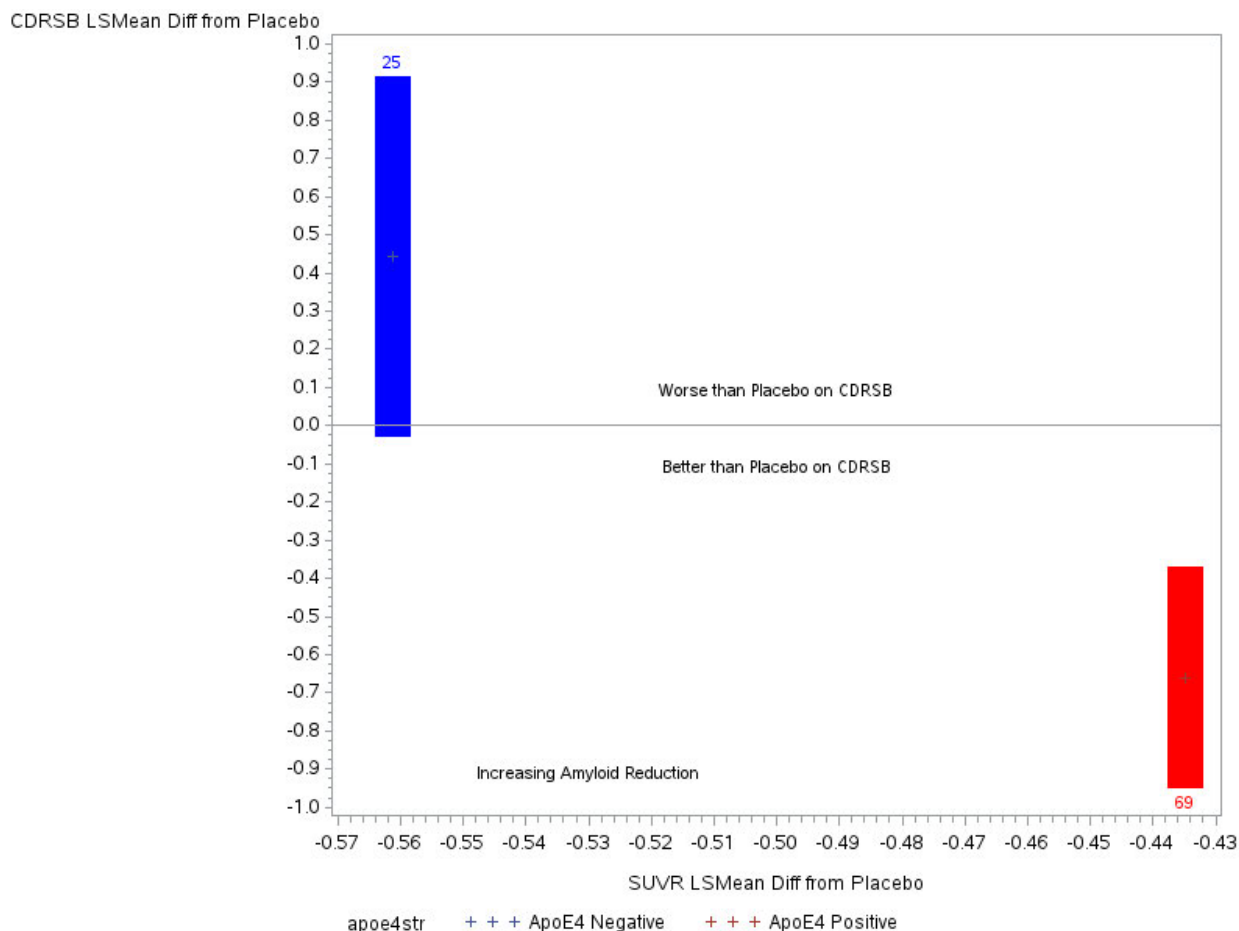
3.2.1.4.2 Reviewer's Results

The reviewer verified the sponsor's results. At Week 76, donanemab-treated participants had statistically significantly less decline in cognition/function than placebo-treated participants as assessed by the iADRS, thus meeting the applicant's primary objective of the study. The LS mean change difference $\pm$ SE was 3.20 ± 1.56 ($p=.042$). For the key secondary CDRSB at week 76 the MMRM estimated treatment difference was -0.36 (0.239) (-0.83, 0.12), $p=0.139$. There was a significant interaction at Week 76 between APOE status and treatment effect on iADRS

[carriers -5.5 (SE=1.9) and non-carriers +3.7 (SE=3.1), the difference of 9.2 between the treatment effects in carriers versus non-carrier had interaction $p=0.0083$]. There also appeared to be a large interaction at Week 76 between APOE status and treatment effect on CDRSB [carriers -.66(SE=.28) and non-carriers +.44 (SE=.46), the difference of 1.10 between the treatment effects in carriers versus non-carrier had interaction $p=0.0427$]. Figure 3 shows that while APOE4 non-carriers had slightly more amyloid reduction than carriers their treatment effect on CDRSB at Week 76 was numerically worse than placebo. If reduction in amyloid is likely to

predict change on clinical endpoint, then one would not expect non-carriers to be numerically worse than placebo on CDRSB and iADRS at Week 76 (i.e., non-carriers have slightly more reduction in Amyloid PET than carriers and while admittedly the sample size is smaller for non-carriers the clinical treatment effect is numerically worse than placebo).

Figure 3 Relationship between CDRSB and PET Amyloid Change at Week 76 for Donanemab monotherapy vs. Placebo by APOE



The scale of the x-axis in Figure 3 is SUVRCLAA. In terms of centiloid scale the differences from placebo were -103 for APOE4 non-carriers and -80 for APOE4 carriers.

3.3 Evaluation of Safety

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

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race, Age, and Geographic Region

4.1.1 Gender, Race, and Age

Fifty two percent of the full analysis set population were Female. Mean Amyloid difference compared to placebo at Week 76 in Females was -83 [S.E.=5] and in Males was -89 [S.E.=5].

Less than 5% were less than 65 years. Overall, there appeared to be a bigger amyloid reduction compared to placebo at Week 76 in Age ≥ 75 (-95 [S.E.=5] than in Age <75 (-73 [S.E.=6]).

Sample sizes were too small to do meaningful subgroup analyses by Race because ninety five percent were white race.

Seventy three percent were APOE carriers. Overall, there was a bigger amyloid reduction compared to placebo at Week 76 in non-carriers (-103 [S.E.=7] than in carriers (-80 [S.E.=5]).

On the iADRS, primary clinical endpoint, there was a bigger placebo progression at Week 76 in ApoE4 carriers than in non-carriers (-11.3[SE=1.3] vs. -6.0[SE=2.1]). The Donanemab-M ApoE4 non-carriers were numerically worse than placebo non-carriers (-3.7 [SE=2.9]) while Donanemab-M carriers were better than placebo carriers (+5.7 [SE=1.8]).

There were similar trends on the CDRSB at Week 76 with an estimated difference of Donanemab-M minus placebo of -0.64 [SE=.28] in carriers vs. +.34 [SE=.46], i.e., numerically worse than placebo, in non-carriers.

4.1.2 Geographic Region

All of the PET SUVR data was collected in North America, and 92% was in the US, so there are no meaningful subgroup analyses by region or country.

4.1.2.1 Individual Sites

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

4.2 Other Special/Subgroup Populations

No other subgroups were examined.

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

There was an interaction noted between ApoE4 status and treatment group for the Week 76 treatment group difference on iADRS with non-carriers being numerically worse than placebo. Although sample size is smaller for non-carriers, this pattern of slightly more reduction of Amyloid on SUVR than carriers but being numerically worse than placebo at Week 76 in non-carriers does not align with treatment effect on Amyloid being reasonably likely to predict treatment effect on the clinical endpoint at Week 76 over all patients, e.g., regardless of ApoE4 status.

The study initially had three arms; placebo, LY3002813 monotherapy and LY3002813 in combination with LY3202626. The combination arm was stopped after 15 patients were randomized to it.

The applicant specified several key secondary endpoints including change from baseline in the following efficacy measures: ADAS-cog, CDRSB, MMSE and ADCS-iADL, followed by several other secondary endpoints including brain amyloid plaque deposition and change in brain tau deposition. A graphical approach was specified to control for multiplicity of the primary iADRS and the ADAS-cog, CDRSB, MMSE and ADCS-iADL, but this plan did not include a pathway for multiplicity adjusted testing of change in brain amyloid deposition. All secondary endpoints not included in the graphical approach for multiplicity adjustment should be considered exploratory. Note also that the Division had not agreed that the iADRS would be acceptable as a primary endpoint.

5.2 Collective Evidence

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

5.3 Conclusions and Recommendations

The applicant submitted Study AACG to evaluate donanemab for treatment of Alzheimer's disease. In this double blind phase 2 study, the applicant's primary endpoint was an Alzheimer's disease composite, Integrated Alzheimer's Disease Rating scale (iADRS).

The study initially had three arms: placebo, LY3002813 monotherapy and LY3002813 in combination with LY3202626. The combination arm was stopped after 15 patients were randomized to it, due to evolving evidence on the safety and efficacy of BACE inhibitors, like LY3202626, as a class.

As prespecified by the applicant in the original protocol, the trial was to be considered a success if LY3002813 alone and/or in combination with LY3202626 administered for up to 76 weeks decreased the cognitive and/or functional decline (based on iADRS) in patients with early symptomatic Alzheimer's Disease. Note that the Division had disagreed that the iADRS would be acceptable as a primary endpoint.

At Week 76, donanemab-treated participants had statistically significantly less decline in cognition/function than placebo-treated participants as assessed by the iADRS, thus meeting the applicant's primary objective of the study. The treatment difference on LS mean change $\pm$ SE was 3.20 ± 1.56 ($p=.042$).

The applicant specified several key secondary endpoints including change from baseline in the following efficacy measures: ADAS-cog, CDRSB, MMSE and ADCS-iADL, followed by several other secondary endpoints including brain amyloid plaque deposition and change in brain tau deposition. A graphical multiplicity adjustment approach was specified to control for multiplicity of testing the primary iADRS and the ADAS-cog, CDRSB, MMSE and ADCS-iADL, but this plan did not include a pathway for multiplicity adjusted testing of change in brain amyloid deposition. All endpoints not included in this graphical approach should be considered exploratory.

None of the secondary endpoints, ADAS-cog, CDRSB, MMSE and ADCS-iADL, met the alpha allocated to them based on the pre-specified graphical approach, though ADAS-Cog13 suggested a treatment effect [LS mean change difference (SE) = -1.86 (0.898), nominal $p=.040$].

Among the secondary endpoints, the applicant focused in the application on the effect of donanemab compared to placebo on brain amyloid pathophysiology at Week 76 of treatment in subjects with Early Alzheimer's Disease as measured by amyloid positron emission tomography (PET). The mixed model for repeated measures analyses of change from baseline in brain amyloid levels as measured by PET SUVR normalized to whole cerebellum mask at Week 76 estimated a reduction of 85 centiloids for LY3002813-M compared to placebo (nominal p-value of <0.001).

The impact of amyloid reduction on the clinical outcome is uncertain. The correlation at the patient level is 0.098 between change in iADRS and change in Amyloid PET SUVR at Week 76 within the LY3002813 monotherapy arm. This correlation is small. Furthermore, there is no apparent treatment effect on the clinical endpoint in APOE4 non-carriers despite the numerically greater amyloid reductions compared to APOE4 carriers (see Figure 3).

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

/s/

TRISTAN S MASSIE
01/17/2023 04:46:44 PM

HSIEN MING J HUNG
01/17/2023 04:51:43 PM