



August 11, 2021

Comments Regarding the Centers for Medicare and Medicaid Services' National Coverage Determination Analysis for Monoclonal Antibodies Directed Against Amyloid for the Treatment of Alzheimer's Disease (CAG-00460N)

Public Citizen, a consumer advocacy organization with more than 500,000 members and supporters nationwide, submits these comments regarding the Centers for Medicare and Medicaid Services' (CMS') national coverage determination analysis for Monoclonal Antibodies Directed Against Amyloid for the Treatment of Alzheimer's Disease (CAG-00460N).

The Food and Drug Administration's (FDA's) decision to approve aducanumab for treatment of Alzheimer's disease showed a stunning disregard for science and eviscerated the agency's standards for approving new drugs. Because of this reckless action, the agency's credibility has been irreparably damaged.

The approval of aducanumab was based on seriously flawed post hoc analyses of two identical phase 3 trials that were stopped early because a preliminary review of the data found that the trials, if continued to completion, were unlikely to show the drug benefitted Alzheimer's disease patients. Moreover, the integrity of the FDA's review of the marketing application for aducanumab was dangerously corrupted by the unprecedented and inappropriately close collaboration between Biogen and the FDA during the analyses of data from the key clinical trials of the drug after the termination of the phase 3 clinical trials because of futility.

CMS must not compound the FDA's egregious error in approving aducanumab on June 7, 2021. Given the lack of scientific evidence that aducanumab provides any clinically meaningful benefit in terms of cognitive function outcomes in Alzheimer's disease patients, the drug cannot possibly be deemed reasonable and necessary for treatment of such patients. Public Citizen therefore urges CMS to issue a national coverage determination that excludes aducanumab from coverage under the Medicare program.

Detailed Discussion

Unprecedented and inappropriate close collaboration between the FDA and Biogen

As we explained in detail in our December 9, 2020, letter to the Department of Health and Human Services Office of Inspector General,¹ there was unprecedented close collaboration between the FDA and Biogen before and after the submission of the company's biologics license application (BLA) for aducanumab for treatment of Alzheimer's disease.

¹ Public Citizen. Letter to the U.S. Department of Health and Human Services Office of Inspector General. December 9, 2020. <https://www.citizen.org/wp-content/uploads/2560.pdf>. Accessed August 10, 2021.

Such close collaboration — which was made readily transparent in press releases and presentation documents issued by Biogen and in the unprecedented joint briefing document prepared by the FDA and Biogen for the FDA's Peripheral and Central Nervous System (PCNS) Drugs Advisory Committee meeting on November 6, 2020 — dangerously compromised the independence and objectivity of senior staff and clinical reviewers in the agency's Office of Neuroscience (ON) in the Center for Drug Evaluation and Research's (CDER's) Office of New Drugs during the agency's review of Biogen's BLA for aducanumab and key data from the two identical pivotal phase 3 clinical trials of the drug. ON Director Billy Dunn supervised the FDA team conducting this review and likely played a key role in the close FDA–Biogen collaboration.

The FDA's unbridled enthusiasm for aducanumab documented in the PCNS Drugs Advisory Committee meeting joint briefing document and echoed in Dr. Dunn's presentation at the November 2020 advisory committee meeting was unsupported by any objective review of data from the pivotal phase 3 clinical trials, which had been terminated early because a planned prespecified interim analysis showed the trials were unlikely to yield evidence that the drug was effective for treating Alzheimer's disease.

Stunning new disclosures in a detailed exposé published on June 29, 2021, by STAT² appear to provide further evidence that starting more than two years ago, key FDA staff in CDER's ON who were responsible for the review of Biogen's BLA for aducanumab for treatment of Alzheimer's disease engaged in an unprecedented and inappropriately close collaboration with the company in the analysis of data from the key clinical trials of the drug. Among the most troubling disclosures in the STAT article were the following:

- In early May 2019 — shortly after Biogen and its partner Eisai had announced the decisions to terminate the two pivotal phase 3 clinical trials of aducanumab after a prespecified interim futility analysis and to end development of the drug — Biogen Chief Scientist Al Sandrock reached out to CDER's ON Director, Dr. Dunn, with whom he “already had a longstanding professional relationship,” and sat down with him for an “off-the-books” meeting while the two were attending a neurology conference in Philadelphia. “Sandrock wanted to let Dunn know that Aduhelm — publicly declared ineffective — might actually be slowing the progression of Alzheimer's... And wanted to know if Dunn would be open to helping find a way to get the drug approved.”
- “It was clear that Billy Dunn was an ally, so the job for Biogen became figuring out how to support his efforts within the FDA,” a former Biogen employee told STAT.”
- Following Sandrock's meeting with Dunn, Biogen “mounted a secret campaign, code-named ‘Project Onyx,’ to resurrect the drug and convince the FDA to give it the green light. Central to their mission was an inside ally: Billy Dunn, the agency's top regulator of Alzheimer's drugs.”

² Feuerstein A, Herper M, Garde D. Inside ‘Project Onyx’: How Biogen used an FDA back channel to win approval of its polarizing Alzheimer's drug. June 29, 2021. STAT. <https://www.statnews.com/2021/06/29/biogen-fda-alzheimers-drug-approval-aduhelm-project-onyx/>. Accessed August 10, 2021.

- “The FDA’s support grew quickly. By June 2019, only a month after the crucial meeting with Dunn, agency officials in his Office of Neuroscience were so willing to advance Aduhelm that they proposed as one option a regulatory shortcut called ‘accelerated approval,’ according to meeting minutes read to STAT. The move stunned even Biogen’s top executives, who had considered that out of the question for a host of reasons, including the fact that the FDA had never used the [accelerated approval] pathway for an Alzheimer’s treatment.

“This disclosure contradicts what the FDA has said publicly in recent weeks about how it came to consider the use of accelerated approval for Aduhelm. An internal review document made public by the FDA last week claimed officials inside the agency first raised the possibility of an accelerated approval of Aduhelm during a meeting held this past March 31.”

- “To some inside Biogen, the FDA’s tone made approval seem inevitable.”
- “The signal of support provided by the FDA’s Dunn to Sandrock during their off-the-books meeting in early May 2019 triggered Project Onyx and the revival of Aduhelm. By mid-May, Biogen was sharing clinical data and other information with FDA officials.”
- “All of the investment that Biogen put into courting Dunn appeared to be paying off. Dunn’s Office of Neuroscience offered Biogen a road map to Aduhelm’s approval, suggesting five different scenarios or options for ways the drug could be reviewed by the agency to allow it to reach the market. Only one of those five options even contemplated Biogen having to conduct another clinical trial before approval. Three would result in the drug’s immediate approval.”
- “After the June 14, 2019, meeting [between Biogen and the FDA], Biogen and the FDA established a ‘working group collaboration’ consisting of company employees and agency review staff. The group met or communicated almost daily in June, July, and August of 2019, working to collect and analyze Aduhelm data for inclusion in the planned marketing submission. The group decided to pursue a standard FDA approval based on data on how patients had fared on cognitive surveys.”
- “‘I knew from the interest levels within FDA that the agency was always going to find a way to approve Aduhelm,’ said a former Biogen employee with knowledge of its interactions with Dunn and other FDA officials during this time period.”

The circumstances described in the STAT exposé, if confirmed, provide overwhelming additional unequivocal evidence of the unprecedented and inappropriately close collaboration between the FDA and Biogen before the submission of the company’s BLA for aducanumab for treatment of Alzheimer’s disease. They also paint a damning picture of FDA drug regulators who surrendered their independence and objectivity, essentially began working on behalf of Biogen, and fostered regulatory capture at the agency.

Lack of scientific evidence that aducanumab provides any clinically meaningful benefit

After completion of a first-in-human phase 1 clinical trial that tested single doses of aducanumab ranging from 0.3 to 60 milligrams/kilogram (mg/kg) in 53 subjects with mild-to-moderate Alzheimer's disease (Study 101) and a subsequent phase 1b randomized, placebo-controlled trial that tested aducanumab at fixed dosages ranging from 1 to 10 mg/kg every four weeks for 14 doses in 196 subjects with prodromal Alzheimer's disease or mild Alzheimer's disease dementia (Study 103), Biogen in 2015 launched two identical large, phase 3, multicenter, randomized, double-blind, placebo-controlled clinical trials to evaluate the safety and efficacy of two dosing regimens of aducanumab (Study 301 [also called ENGAGE] and Study 302 [also called EMERGE]).³ Studies 301 and 302 enrolled 1,653 and 1,643 subjects, respectively, with mild cognitive impairment due to Alzheimer's disease or mild Alzheimer's disease dementia.

On March 21, 2019, Biogen and its partner, Eisai, issued a press release announcing the decision to terminate both pivotal phase 3 trials testing aducanumab after a prespecified interim futility analysis by an independent data-monitoring committee indicated that the trials were unlikely to meet their primary efficacy endpoint upon completion.⁴ That action should have marked the end of aducanumab as a potential treatment for Alzheimer's disease, at least as it pertains to the clinical trials thus far completed.

On October 22, 2019, Biogen shocked the medical community when it issued another press release announcing plans to seek FDA approval for aducanumab to treat patients with early Alzheimer's disease based on a series of post hoc analyses of data from Studies 301 and 302, including additional data collected after the interim futility analysis and the announced termination of the trials.⁵ The company stated in the press release that these new analyses had been "conducted by Biogen in consultation with the FDA." In an October 22, 2019, slide presentation for investors, Biogen similarly noted that the company "consulted with external advisors and the FDA to better understand these different results" and that "[a]fter consulting with the FDA, **we believe that the totality of these data support a regulatory filing**" [emphasis in original].⁶

The primary efficacy outcome for Studies 301 and 302 was the change from baseline at 18 months (78 weeks) of treatment in the Clinical Dementia Rating Sum of Boxes (CDR-SB) score, which provides an assessment of cognition and function and has a scale ranging from 0 to 18. A difference of 1 to 2 points in CDR-SB scores is considered to be the minimal clinically

³ Food and Drug Administration and Biogen. Combined FDA and Applicant briefing document for the November 6, 2020, meeting of the Peripheral and Central Nervous System Drugs Advisory Committee meeting regarding NDA/BLA# 761178, Aducanumab. <https://www.fda.gov/media/143502/download>. Accessed August 10, 2021. PDF pages 22-23.

⁴ Biogen. Biogen and Eisai discontinue phase 3 ENGAGE and EMERGE trials of aducanumab in Alzheimer's disease. March 21, 2019. <https://investors.biogen.com/news-releases/news-release-details/biogen-and-eisai-discontinue-phase-3-engage-and-emerge-trials>. Accessed August 10, 2021.

⁵ Biogen. Biogen plans regulatory filing for aducanumab in Alzheimer's disease based a new analysis of larger dataset from phase 3 studies. November 28, 2019. <https://investors.biogen.com/news-releases/news-release-details/biogen-plans-regulatory-filing-aducanumab-alzheimers-disease>. Accessed August 10, 2021.

⁶ Biogen. Aducanumab update. October 22, 2019. <https://investors.biogen.com/static-files/5a31a1e3-4fbb-4165-921a-f0ccb1d64b65>. Accessed August 10, 2021.

meaningful difference in clinical outcome assessments for Alzheimer's disease clinical trials.⁷ The post hoc analysis of Study 301 revealed no statistically significant differences in CDR-SB scores at 18 months between the aducanumab group subjects and the placebo group subjects for either the low-dose or high-dose regimens.⁸ The post hoc analysis of Study 302 showed no statistically significant differences in CDR-SB scores at 18 months between the aducanumab group subjects and the placebo group subjects for the low-dose regimen but did find a mean difference in the CDR-SB of 0.39 (95% confidence interval -0.69 to -0.09) between the aducanumab group subjects and the placebo group subjects for the high-dose regimen,⁹ a difference that was statistically significant ($p=0.01$) but not clinically meaningful and highly likely to be a false-positive signal.

The statistical review and evaluation document written by Tristan Massie, Ph.D., Mathematical Statistician, Division of Biometrics I, Office of Biostatistics, Office of Translational Science, CDER provides the most objective FDA assessment of the data from Studies 301 and 302.¹⁰ Notably, Dr. Massie is not under the supervisory chain of Dr. Dunn. Dr. Massie's review clearly indicates that he was one of the few FDA staff involved in the review of Biogen's BLA who did not succumb to the regulatory capture that compromised the independence and objectivity of the FDA's overall review of the BLA for aducanumab. The following are representative excerpts from the executive summary of Dr. Massie's review that highlight some of the numerous serious flaws he found in the post hoc data analyses of Studies 301, 302, and 103 that were conducted by Biogen (in collaboration with *other* FDA staff):

The two phase 3 studies (study 301, study 302) were stopped early for futility...when both studies had reached 50% completion since **it was estimated based on the interim study-pooled estimate of the treatment effects that both studies individually had <20% chance of success for either dose if completed.** Following a futility press release announcement and collection of subsequent study closeout follow up data, the sponsor requested a meeting to discuss the two trials final data after discovering that despite the futility conclusion, **the final analysis on face showed a statistically significant effect for the high dose in one of the two trials ($p=0.01$) but not the other ($p=0.83$).**

Inconsistency on many levels summarizes the final clinical efficacy data from these trials. Because the two phase 3 studies were terminated for futility, **the NDA [new drug application] package doesn't contain a single phase 3 study that was fully completed according to the plan. In fact, almost 50% are missing the Week 78 time point assessment of CDRSB [Clinical Dementia Rating Sum of Boxes] which is the only timepoint that shows any significance and that is only significant in one of the two studies (the first study, study**

⁷ Andrews JS, Desai U, Kirson NY, et al. Disease severity and minimal clinically important differences in clinical outcome assessments for Alzheimer's disease clinical trials. *Alzheimers Dement (N Y)*. 2019 Aug 2;5:354-363.

⁸ Food and Drug Administration. Statistical review(s) for application number 761178Orig1s000, aducanumab. May 11, 2021. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2021/761178Orig1s000StatR_Redacted.pdf. Accessed August 10, 2021. PDF page 28.

⁹ *Ibid.* PDF page 28.

¹⁰ *Ibid.*

301, high dose is numerically worse than placebo at Week 78 on the primary endpoint). A chance worse placebo response in study 302 than was observed in study 301 could explain the significance of study 302 ($p=0.01$).

This BLA submission does not have a situation such as just one study in existence and for which that study is strong. **We have a second large adequate well controlled study that directly contradicts the first and is not even close to significance $p=0.8252$. If one has two studies and takes the best and pretends like it's the only study, one's estimate is most likely biased and misleading.** In the opportunity to complete subset of 302 the high dose vs. placebo has a p-value of 0.0368 for the CDRSB at Week 78 and, even in the ITT [intention-to-treat] population, there was no significance before Week 78. It is not justifiable to search for patients in 301 who are similar to 302 because that may have selection bias and presumes that 302 is right and 301 is wrong, for which there is no justification (without resorting to post-hoc analyses which are at best exploratory). Any selection of patients would need a proper placebo control. The overall 301 primary result is the only valid well controlled, multiplicity adjusted, randomization validated analysis of 301 (and it had a substantial sample size).

The sponsor tries to discount study 301 due to post-hoc defined "rapid progressors". Rapid progressors are likely part of the reality of Alzheimer's and after the fact it is too late to address them in a completed large randomized study. Study 302 could just as well be the outlier relative to the true proportion of outliers in the natural progression. In fact, the range of CDRSB changes in Study 301 at 18 months appears consistent with the Alzheimer's Disease Neuroimaging Study study...The sponsor also tries to use 301 to find a subgroup similar to Study 302, i.e., a subgroup showing efficacy in 301 but this relies on post-hoc non-randomized comparisons (Figure 14 on page 49). These analyses hide the fact that the post-hoc matched placebo progresses faster as the number of 10 mg/kg doses increases in these post-randomization event defined subgroups and such post-hoc matching can never equal a true randomization backed analysis. **The only valid analysis of Study 301 is the prespecified randomization supported analysis of study 301 which failed for the high dose ($p=0.83$) and this study outcome should not be discounted without an extremely compelling reason (which there is not).**

The sponsor argues, relying on non-randomized comparisons, that the high dose arm was challenged by intermediate dosing rather than full dosing in some patients. This can be countered by the fact that the low dose was numerically better than the high dose in Study 301, a comparison supported by randomization, and the low dose was also numerically better than high dose in study 302 in the subset after the mid-study protocol amendment increasing the maximum high dose for APOE [apolipoprotein E] carriers. Furthermore, the APOE non-carriers have less treatment effect on all four primary and key secondary efficacy endpoints despite having 10 mg/kg dosing from study start and less ARIA

[amyloid-related imaging abnormalities] adverse events than APOE carriers, so fewer dose reductions due to ARIA.¹¹

[Emphasis added]

Dr. Massie made the following conclusions:

In summary, the totality of the data does not seem to support the efficacy of the high dose. There is only one positive study at best and a second study which directly conflicts with the positive study. **Both studies were not fully completed as they were terminated early for futility** and had sporadic unblinding for dose management of ARIA cases which was much higher in the drug group(s). **The Amyloid PET [positron emission tomography] substudy data suggested a larger effect in APOE- (non-carriers) which is the opposite of what was observed for the overall clinical outcome data. Within the high dose group (or high and low combined) at the patient level there is no compelling correlation between the Week 78 change in the primary biomarker A β [amyloid-beta] SUVR [standard uptake value ratio, a measure of amyloid-beta in the brain on PET scans] in the Composite region of interest with reference in the cerebellum and the Week 78 Change from baseline in CDRSB (see the biomarker section 3.2.1.4.2.2). For these reasons, substantial evidence has not been met in this application.**¹²

[Emphasis added]

In a prerecorded presentation posted on the FDA's website for the agency's November 6, 2020, PCNS Drugs Advisory Committee meeting regarding aducanumab, Dr. Massie offered the following harsh critique of the post hoc analyses of clinical trial data that had been conducted collaboratively by Biogen and FDA staff:

In the past couple years in the scientific community the issue of reproducibility of experiments has been raised. In this case we do not have a single strong study in isolation. On the contrary[,] we actually have a second trial in which the purported effective dose was in the wrong direction compared to placebo, i.e., numerically worse than placebo. Under the null if winning in just one study out of two was enough[,] then the chance of falsely rejecting the null would be .0975 across the two studies. Furthermore, if we select only the better study[,] our estimate is very likely biased, and we already know not consistently repeatable in our experience. **Thus[,] excluding data from a large trial without sufficient**

¹¹ *Ibid.* PDF pages 8-10.

¹² *Ibid.* PDF page 11.

justification is unscientific, statistically inappropriate and misleading.^{13,14}
[Emphasis added]

During the November 6, 2020, PCNS Drugs Advisory Committee meeting, many committee members voiced critical comments and pointed questions that reflected deep skepticism about the post hoc analyses that had been conducted collaboratively by Biogen and the FDA. Some illustrative examples of such comments and questions include the following:

Scott Emerson, M.D., Ph.D., Professor Emeritus of Biostatistics, University of Washington, Seattle, Washington:

This analysis seems to be subject to the **Texas sharpshooter fallacy**, a name for the joke of someone first firing a shotgun at a barn and then painting a target around the bullet holes. So, understanding the sampling scheme for the presented results is all important. Can you clarify...the extent to which the selection of data — that is, which study and what dataset — was prespecified, and if they were prespecified, what's the evidence that the discordant results are truly uncommon under the null hypothesis? ...¹⁵

These decisions [regarding the analyses] were made after you had the results that [Study] 302 and [Study] 301 were discordant. **So...it was not prespecified at the very beginning of the trial that [Study] 302 was going to be the only study analyzed, correct?** ...¹⁶

P-values are meant to capture the possibility that there might be randomization imbalances. We'll come back far later to whether you can take a post-randomization variable and exclude them. I don't believe you can. You apparently believe you [can] **with some complicity from the FDA clinical staff**, though not the FDA statistician as near as I can tell...¹⁷

Well again, we can talk about the sampling of [Study] 302 and what the true sampling distribution was for [Study] 302, or we could talk about [Study] 302 with [Study] 301, taking both results. One result is saying [Study] 302 is the best of two possible studies; that's one sampling distribution. Another is saying we're going to look at Study 302, just [Study] 302. And recognize that [Study] 301 carries the exact same weight and eventually would be taken care of in a meta-

¹³ Food and Drug Administration. FDA pre-recorded presentation slides for the November 6, 2020 meeting of the Peripheral and Central Nervous System Drugs Advisory Committee. <https://www.fda.gov/media/143504/download>. August 10, 2021.

¹⁴ Food and Drug Administration. FDA pre-recorded presentation transcripts for the November 6, 2020 meeting of the Peripheral and Central Nervous System Drugs Advisory Committee. <https://www.fda.gov/media/143505/download>. Accessed August 10, 2021.

¹⁵ Food and Drug Administration. Webcast recording of the November 6, 2020, meeting of the Peripheral and Central Nervous System Drugs Advisory Committee. <https://collaboration.fda.gov/p2uew93ez7dw/>. Accessed November 29, 2020. Approximately 01:00:27 to 01:01:02.

¹⁶ *Ibid.* Approximately 01:02:21 to 01:02:38.

¹⁷ *Ibid.* Approximately 01:06:08 to 01:06:31.

analysis. **My interpretation is the FDA wants us to imagine that we can look at [Study] 302, just those results, but we need to recognize that that is the best of two studies conducted concurrently**, and again if Dr. Dunn will tell me that what his persuasive evidence means, I heard 'persuasive evidence' far more often than what any results were — just conclusions — but if he'll tell me what his persuasive evidence is in terms of the P-value that he was looking at on that primary endpoint. I realize there's totality of evidence, but I just want to know was he taking into account that that was a p-value that was approximately .024 **or was he taking into account the erroneous conclusion that that was a p-value of .012...**¹⁸

I'll note that **I was very disturbed by...some of the FDA's interpretation of [Study] 301** by starting out with the assumption that the treatment works and now trying to say why do we get null results in [Study] 301. **Usually, we start off saying the treatment doesn't work and are these [results] compatible with that.** I spoke to this earlier, about if you assume the treatment doesn't work, then it's not that rare to have some strong results on one of the trials and just completely nothing results [on another trial]. And that's happened to me many times in my life when I monitored trials at the same time... Lastly, **I was very, very, very disturbed by some of the analyses that were considered.** I was glad to hear Dr. Dunn soften what they were doing and try to make clear, but I will just state that...some 20 years ago I was involved as **an expert witness in a scientific misconduct trial** of, as it turns out, an Alzheimer's disease researcher, who was removing data that...she didn't like and just seeing what happens, and **that's just never acceptable to do.** So, for the most part, the sensitivity analyses were sometimes just completely unnecessary, they were just reproducing the statistics we already had...¹⁹

You know, if you thought that I was being critical, you're absolutely correct. On page 226 has one of the lines that I felt was bad... You start off by saying if [aducanumab is] effective, then it follows that's reflective of the two effects, and their patients in Study 301 who, based on certain characteristics, should show response. Okay, the flip side is that — I...again didn't have time earlier, but I was going to ask for the analysis in which you added into [Study] 302 the patients who weren't represented that...were rapid progressors, perhaps owing to the drug itself. And you never did that analysis, **so you were not at all symmetric, and you certainly were not starting off with saying, could these results be explained by a null affect,** in which case you'd say, yeah you know what, nothing was going on in [Study] 301, that's the truth, and in [Study] 302 why did we get aberrant results. And so, the truth is probably somewhere in between about the way to do it. **But there was just no question that all of this was just terrifically one-sided, and again, I'm highly critical of the fact that the FDA presentation today was so heavily weighted to just giving the same**

¹⁸ *Ibid.* Approximately 03:40:08 to 03:41:24.

¹⁹ *Ibid.* Approximately 05:44:46 to 05:46:21.

conclusions that the sponsor did and that there was not presentation by the [FDA] statistician, who had done a careful analysis and made many points that I was very glad to see that the committee read.²⁰

Chiadi U. Onyike, M.D., M.H.S., Associate Professor of Psychiatry and Behavioral Sciences, Division of Geriatric Psychiatry and Neuropsychiatry, Department of Psychiatry and Behavioral Sciences, The Johns Hopkins University School of Medicine, Baltimore, Maryland:

My question — I'll set aside for the moment the idea that the post hoc analysis seeking to disqualify the observations of Study 301 are okay — with that in mind,... you put forward certain explanations for the discordance in the results between the two studies, 301 and 302. **What you haven't discussed is the possibility that the placebo groups differed.**²¹

G. Caleb Alexander, M.D., M.S., Professor of Epidemiology and Medicine, Johns Hopkins Bloomberg School of Public Health, Center for Drug Safety and Effectiveness, Baltimore, Maryland:

I do want to say that it seems to me that **there is an extraordinary amount of explaining around the contrary findings**, and I think Dr. [Craig] Mallinckrodt [Biostatistician, Biogen], you recently said that, you... use the word 'causal' in referring to rapid progressors and dosing differences as explaining the failure of 301. And I just...don't see it... With rapid progressors, we're talking about a difference of four or five people in a group containing 500 or more, and this theory of rapid progressors was introduced, I believe, only post hoc, and other methods of examining outliers, other outlier analyses, that may be more suitable — such as robust regression or trim means — also failed to replicate the findings of [Study] 302 in looking at [Study] 301... [I]t reminds me a little bit of a separate committee where there was a subset of individuals that appeared to be responding particularly well, and I think a member of the committee used the term 'super responders.' **And so, I understand the appeal of trying to identify and explain away the null findings, but I don't think that the evidence is there...** So I want to turn then to placebo response, and while you provided some helpful information, you didn't include...the graphical illustration that I think is most troublesome to me and which I'm sure you're familiar with, which was included in the biostatistical review by the FDA...²²

I wanted to ask questions of the FDA earlier, and it's relevant to this question, and I guess the bottom line is that **I find the materials that the FDA has provided strikingly incongruent, and I have a very hard time understanding**, after carefully reviewing what I thought was a very well done and well-articulated

²⁰ *Ibid.* Approximately 05:57:12 to 05:58:40.

²¹ *Ibid.* Approximately 01:10:36 to 01:11:07.

²² *Ibid.* Approximately 01:23:39 to 01:25:28.

[FDA] biostatistical review, which convincingly argued the evidence was ‘at best compellingly conflicted,’ **how the FDA could conclude that there are substantial evidence of effectiveness** and, in particular, that Study 302 provides ‘a robust and exceptionally persuasive study,’ and **it just feels to me like the audio and the video on the TV are out of sync. And there are literally a dozen different red threads that suggest concerns about the consistency of evidence.** A dozen — I mean for every point that you can find suggesting support, there is another point or two that raises concern. So, there’s only one time point with statistically significant different findings from placebo...²³

I just have a few brief comments here about Study 103, but I do think that it’s one of these settings where...**it felt to me like the briefing materials really selectively identified lines of argument which would be supportive of [Study] 302 and then just sort of set aside a similar greater number of lines of argument that that that detract from [Study] 302.**²⁴

Michael Gold, M.S., M.D., Vice-President, Neurosciences Development, AbbVie, North Chicago, Illinois; PCNS Drugs Advisory Committee Non-Voting Industry Representative:

Yeah, **so I have a particular issue with [viewing] Study 302 independently and without regard for [Study] 301 since those studies are identical in design, identical in inclusion-exclusion criteria, identical presumably in biomarker analysis**...I have real serious issues with how you can divorce the two studies from each other...²⁵

I think what I’m struggling with is the notion that it was almost **tacitly accepted that [Study] 302 represented truth and that [Study] 301 did not.** And so, **a lot of efforts are trying to discredit or to minimize the [Study] 301 data...** So, I just didn’t understand why there seemed to be this kind of...**unilateral effort to discredit one study.** It would have been interesting...to take the opposite position: To say [Study] 301 represents truth and what in [Study] 302 could have accounted for a false positive signal...²⁶

I think it’s important to sort of be respectful of the fact that **[Study] 301 was well-designed, well-conducted, well-executed. There’s no evidence that that it’s somehow...defective in any way, shape, or form.**²⁷

Aaron S. Kesselheim, M.D., J.D., M.P.H., Associate Professor of Medicine, Harvard Medical School; Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women’s Hospital, Boston, Massachusetts:

²³ *Ibid.* Approximately 03:35:22 to 03:36:30.

²⁴ *Ibid.* Approximately 04:33:45 to 04:34:15.

²⁵ *Ibid.* Approximately 03:43:07 to 03:43:28.

²⁶ *Ibid.* Approximately 05:47:04 to 05:47:48.

²⁷ *Ibid.* Approximately 05:49:37 to 05:49:50.

It's strange to rely on...half or two-thirds of a study [i.e., Study 302] as your evidence of effectiveness for a drug...²⁸

I think that's another issue to discuss just in terms of the way that the results are framed. Much of these results are framed in the context of percentage changes from placebo. **The actual real effect size is on the order of change in .4 on an 18-point CDR-SB scale**, and so I think that that's...also a relevant issue to think about.²⁹

Joel S. Perlmutter, M.D., Elliot Stein Family Professor of Neurology and Professor of Radiology, Neuroscience, Physical Therapy & Occupational Therapy, Washington University School of Medicine, St. Louis, Missouri:

First of all, I do think that **having this discussion point is being foisted upon us and is artificial**. The second then is about specific points about [Study] 302, as I'm concerned about describing the benefits in multiple endpoints when I do believe we saw data that they are correlated, multiple endpoints are correlated. I think we see a lack of correlation between the [amyloid] beta change and the clinical endpoint CDR-SB. I think that's of concern. I think the retrospective application of the definition of rapid progressors is...a concern for me, **and the differential unblinding in people getting the high dose**. And I think these are all raise questions. And even if we don't see statistical difference on the unblinding, when you add these things up, they can together cumulatively be an issue, and we saw that with just small groups of...rapidly progressors removed in other places. So, this analysis is very sensitive to small changes in the numbers in which people are being included and excluded...³⁰

If we approve something where the data is not strong, that **we have a risk of delaying good treatment and effective treatment for more than a couple of years — for many years**. And I think there's a huge danger in approving something that turns out not to be effective. I think that danger is much, much greater.³¹

John Duda, M.D., BLR&D Senior Clinical Research Scientist, National Director, Parkinson's Disease Research Education and Clinical Centers; Chairperson, National VA Parkinson's Disease Consortium; Director, Parkinson's Disease Research, Education and Clinical Center and Co-Director, Center for Neurotrauma, Neurodegeneration and Restoration at the Michael J. Crescenz VA Medical Center in Philadelphia; and Associate Professor of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania:

²⁸ *Ibid.* Approximately 04:03:16 to 04:03:26.

²⁹ *Ibid.* Approximately 04:03:47 to 04:04:14.

³⁰ *Ibid.* Approximately 04:09:48 to 04:10:59.

³¹ *Ibid.* Approximately 04:26:37 to 04:27:11.

I think all in all, the main — I think several others have said it already — **Dr. Massie's criticisms just were never addressed in the clinical overview** and there seemed to be a disconnect between different aspects of the FDA reporting that are very difficult for us to draw conclusions from.³²

Madhav Thambisetty, M.D., Ph.D., Senior Investigator and Chief, Clinical and Translational Neuroscience Section, Laboratory of Behavioral Neuroscience, National Institute on Aging, National Institutes of Health; Adjunct Professor of Neurology, Johns Hopkins University School of Medicine, Baltimore, Maryland (during discussion of question 7):

I think both [Study] 301 and [Study] 302 were well-designed phase 3 clinical trials to test clinical effectiveness of aducanumab, and they provided discordant results. **I don't think the post hoc exploratory analyses presented provide justification for discounting or overriding [Study] 301 and considering [Study] 302 independently.**³³

At the conclusion of its deliberations, the advisory committee vote on the key final question — ***In light of the understanding provided by the exploratory analyses of Study 301 and Study 302, along with the results of Study 103 and evidence of pharmacodynamic effect on Alzheimer's disease pathophysiology, is it reasonable to consider Study 302 as primary evidence of effectiveness of aducanumab for the treatment of Alzheimer's disease?*** — was 0 YES, 10 NO, 1 UNCERTAIN,³⁴ formally indicating near-unanimous opposition to FDA approval of aducanumab based on the available clinical trial data — opposition that was readily apparent throughout the meeting.

Accelerated Approval based on an unvalidated surrogate endpoint

The FDA approved aducanumab for treatment of Alzheimer's disease under its Accelerated Approval pathway based on an unvalidated surrogate endpoint — reduction of amyloid-beta plaques in the brain — and the erroneous conclusion that this reduction “is reasonably likely to predict clinical benefit to patients.”³⁵ In her June 7, 2021, statement announcing the FDA's approval decision, CDER Director Patrizia Cavazzoni asserted that “reduction in [brain amyloid-beta] plaques — a hallmark finding in the brain of patients with Alzheimer's — is expected to lead to a reduction in the clinical decline of this devastating form of dementia.”³⁶ However, the currently available evidence — including evidence from the clinical trials of aducanumab itself

³² *Ibid.* Approximately 05:51:26 to 05:51:49.

³³ *Ibid.* Approximately 05:52:08 to 05:52:28.

³⁴ Food and Drug Administration. Final summary minutes of the Peripheral and Central Nervous System Drugs Advisory Committee meeting. November 6, 2020. <https://www.fda.gov/media/145690/download>. Accessed August 10, 2021.

³⁵ Food and Drug Administration. FDA's decision to approve new treatment for Alzheimer's disease. June 7, 2021. <https://www.fda.gov/drugs/news-events-human-drugs/fdas-decision-approve-new-treatment-alzheimers-disease>. Accessed August 10, 2021.

³⁶ *Ibid.*

— fails to show a compelling, meaningful correlation between changes in this surrogate endpoint and changes in clinical measures of cognitive function.

Like the prior 22 unsuccessful experimental drugs targeting amyloid-beta that were pursued as potential treatments for Alzheimer's disease over the past two decades, use of aducanumab is predicated on the still-unproven "amyloid hypothesis," which was introduced in the early 1990s and posits that deposition of amyloid-beta plaques in the brain causes the neuronal degeneration seen in Alzheimer's disease.³⁷ None of these drugs was shown to be efficacious for treating Alzheimer's disease and many caused serious harm, including worsening of cognition in some cases (see Table below).³⁸

Table. List of Failed Experimental Drugs for Alzheimer's Disease Targeting Amyloid-Beta Accumulation

Drug Class	Drug Name (Publication Year of Final Phase Trial(s); Last Trial Phase Conducted; Reason(s) for Failure)
Aβ antigens	<ul style="list-style-type: none"> • AN-1792 (2002; phase 2; toxicity and lack of efficacy) • Vanutide (2013; phase 2; lack of efficacy) • Affitope AD02 (2014; phase 2; lack of efficacy, worsened cognition) • CAD106 (2014; phase 2; lack of efficacy, worsened cognition)
Aβ aggregation inhibitors	<ul style="list-style-type: none"> • Tramiprosate (2007; phase 3; lack of efficacy) • Scyllo-inositol (2009; phase 2; toxicity and lack of efficacy, increased mortality) • PBT2 (2014; phase 2; lack of efficacy)
γ-Secretase modulator	<ul style="list-style-type: none"> • Tarenflurbil (2009; phase 3; lack of efficacy, worsened global status)
γ-Secretase inhibitors	<ul style="list-style-type: none"> • Begacestat (2010; phase 2; toxicity and lack of efficacy) • Semagacestat (2011; phase 3; toxicity and lack of efficacy, worsened cognition) • Avagacestat (2012; phase 2 (2 trials); toxicity and lack of efficacy, worsened cognition)
Anti-Aβ monoclonal antibodies	<ul style="list-style-type: none"> • Ponezumab (2011; phase 2; lack of efficacy) • Bapineuzumab (2012; phase 3; lack of efficacy) • Crenezumab (2014; phase 2; lack of efficacy) • Gantenerumab (2014; phase 2 (2 trials); lack of efficacy) • Solanezumab (2013 (1 trial), 2016 (2 trials); phase 3 (3 trials); lack of efficacy)
Anti- Aβ polyclonal antibody	<ul style="list-style-type: none"> • Immunoglobulin (2013; phase 3; lack of efficacy)

³⁷ Alavi A, Barrio JR, Werner TJ, et al. Suboptimal validity of amyloid imaging-based diagnosis and management of Alzheimer's disease: why it is time to abandon the approach. *Eur J Nucl Med Mol Imaging*. 2020;47(2):225–230.

³⁸ Panza F, Lozupone M, Logroscino G, Imbimbo BP. A critical appraisal of amyloid- β -targeting therapies for Alzheimer disease. *Nat Rev Neurol*. 2019;15(2):73-88.

β-Secretase inhibitor	<ul style="list-style-type: none"> • LY2886721 (2013; phase 2; toxicity) • AZD3839 (2013; phase 1; toxicity) • Verubecestat (2016 (1 trial), 2018 (1 trial); phase 3 (2 trials); lack of efficacy, increased mortality, worsened cognition) • Atabecestat (2018; phase 3; toxicity, worsened cognition) • Lanabecestat (2018; phase 3 (2 trials); lack of efficacy, worsened cognition)
------------------------------	--

No test to measure amyloid-beta plaques in the brains of living patients was available until 2012, when the FDA approved florbetapir F18 injection for positron-emission tomographic (PET) imaging.³⁹ (Note that utility and validity of florbetapir F18 PET imaging for measuring amyloid-beta plaques in the brain has been challenged by some experts in nuclear medicine imaging.⁴⁰) Thus, for drugs targeting amyloid-beta that were being developed as potential treatments for Alzheimer's disease prior to 2012, there was no in vivo test that allowed researchers to assess whether these drugs reduced the accumulation of amyloid-beta plaques in the brains of clinical trial subjects.

Since 2012, clinical testing of several drugs that targeted amyloid-beta in the brain — including the anti-amyloid-beta monoclonal antibodies bapineuzumab⁴¹ and gantenerumab⁴² and the amyloid-beta antigen CAD 106,⁴³ — found that the drugs either prevented further accumulation of amyloid-beta in the brain or reduced the amount of amyloid-beta plaques in the brain based on florbetapir F18 PET imaging. But for each drug, there was no evidence that the drugs were effective in preventing cognitive decline.

Most germane, although the clinical trials of aducanumab demonstrated amyloid-beta reductions that were greater than those observed for bapineuzumab, gantenerumab, and CAD 106, data from the one phase 2 and two phase 3 clinical trials of aducanumab also found no compelling, meaningful correlation between the reductions in amyloid-beta plaques in the brain and changes in CDR-SB scores, which served as the primary clinical measure for cognitive function in the subjects with Alzheimer's disease who were enrolled in these trials.

In his statistical review, the FDA's Dr. Massie provided the following detailed explanation for why data from the clinical trials of aducanumab failed to demonstrate a compelling, meaningful correlation between changes in amyloid-beta deposits in the brain and changes in CDR-SB scores at both the individual-patient and study-group levels:

³⁹ Yang L, Rieves D, Ganley C. Brain amyloid imaging — FDA approval of florbetapir F18 injection. *N Engl J Med.* 367(10):885-887.

⁴⁰ Alavi A, Barrio JR, Werner TJ, et al. Suboptimal validity of amyloid imaging-based diagnosis and management of Alzheimer's disease: why it is time to abandon the approach. *Eur J Nucl Med Mol Imaging.* 2020;47(2):225–230.

⁴¹ Salloway S, Sperling R, Fox NC, et al. Two phase 3 trials of bapineuzumab in mild-to-moderate Alzheimer's disease. *N Engl J Med.* 2014;370(4):322-333.

⁴² Ostrowitzki, S, Lasser RA, Dorflinger E, et al. A phase III randomized trial of gantenerumab in prodromal Alzheimer's disease. *Alzheimers Res Ther.* 2017;9(1):95.

⁴³ Vandenberghe R, Riviere ME, Caputo A, et al. Active Aβ immunotherapy CAD106 in Alzheimer's disease: A phase 2b study. *Alzheimers Dement.* 2016;3(1):10–22.

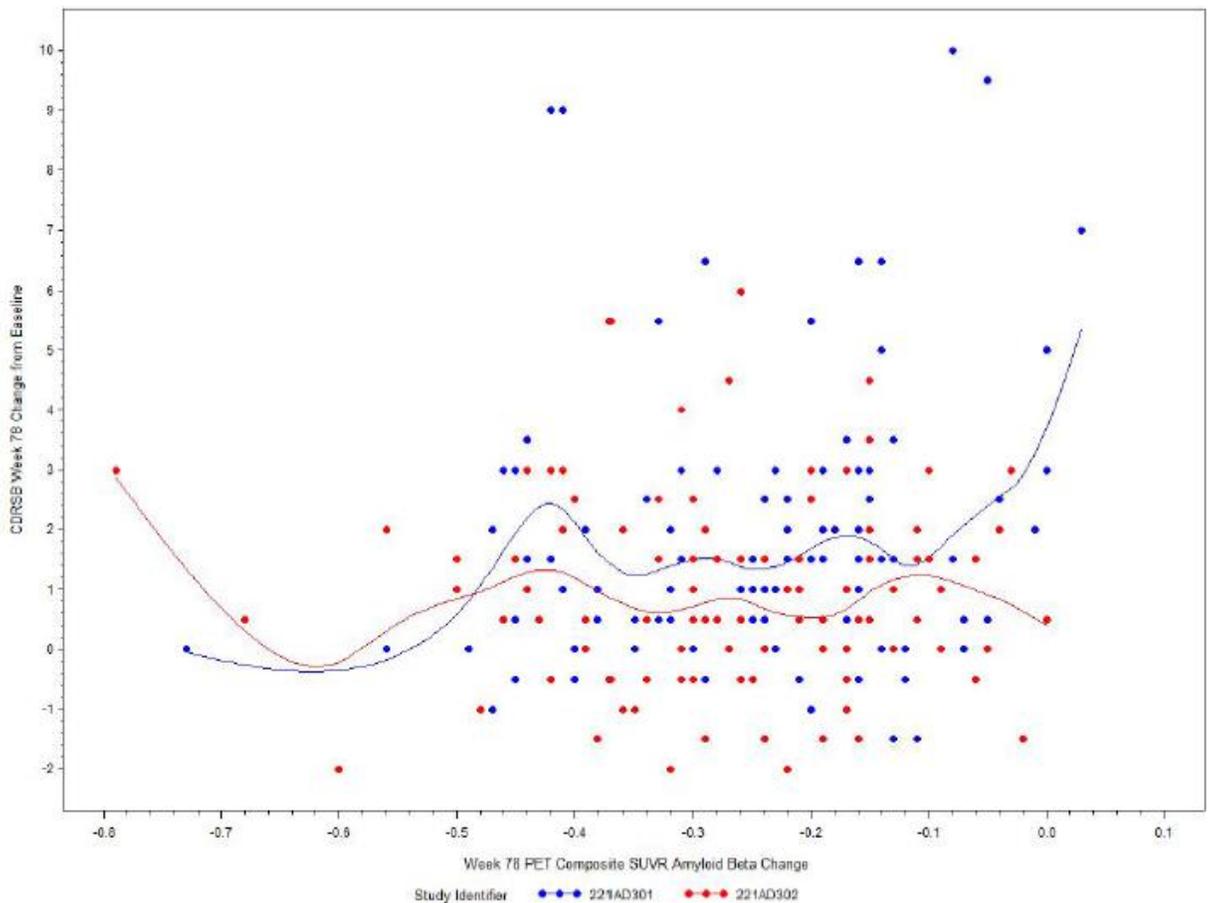
The final SAP [statistical analysis plan] for biomarker analysis is dated 2020; well after unblinding. The SAP specified before unblinding (12 Sep 2018) stated that correlations based on the pooled-study data would be considered primary. It planned Pearson and Spearman correlations[,] both unadjusted[,] and correlations adjusted for baseline CDRSB and baseline biomarker and also to examine correlations with the biomarker at Week 78, as well as Week 26. **The prespecified SAP states that the correlations were to be done by treatment group, which differs from the post-unblinding biomarker SAP which suggested pooling the low and high dose groups.**

In the PET biomarker substudy (N=442 subjects from [study] 302) overall 29% of placebo and 22% of the high dose group were missing Week 78 and of those who had the opportunity to complete Week 78 12% of the 99/(140) placebo and 12% of the 113/(145) high dose were missing the Week 78 assessment of composite SUVR [standard uptake value ratio, a measure of amyloid-beta in the brain on PET scans]. The composite SUVR change from baseline was significant at Week 78 (-0.28 [S.E.= 0.01]) as well as Week 26 (-0.09 [S.E.=0.01]). The low dose was also significant (-0.08 at Week 26 and at Week 78 -0.18 [S.E.=0.01]). Placebo LSMeans [least-squares mean] composite SUVR at Week 78 was 0.0003 [S.E.=0.008] in [study] 301 and +0.0158 [S.E.= 0.01] in [study] 302.

The PET substudy is not directly randomized or balanced. In completers of the PET substudy there were 12% more aged 71-80 in the high dose than placebo. 9-10% more age 71+ in the high dose and a 10% imbalance in APOE. There are 6% more aged 71-80 in the high dose than placebo in the PET population overall (this age group was the group with the highest apparent estimated effect in study 302). Mean CDRSB at baseline in the PET subgroup was 2.39 in study 302 as compared to 2.51 in the high dose overall; in [Study] 301 it was also 2.39 in the PET subgroup but 2.40 in the high dose overall.

If these PET SUVR amyloid changes are meaningful[,] why is the biomarker change positive in [study] 301 (N=544 total: High vs. Placebo = -0.07 [S.E.=0.008] at Week 26 and -0.24 [S.E.=0.01] at Week 78) and the biomarker shows dose dependence (low dose -0.07 [S.E.=0.008] at Week 26 and Week 78 -0.17 [S.E.=0.01]) but the clinical change in CDRSB is not significant and the low dose is numerically better than the high dose on CDRSB change from baseline at Week 78 in study 301? Unadjusted Pearson correlations of Week 78 changes from baseline between the biomarker and CDRSB within the high dose group are 0.135 for study 301 and -0.036 for study 302 (see Figure 18 which includes a local regression curve to help the eye assimilate the data points). The unadjusted Pearson correlation with the biomarker for the on-face positive study 302 is in the wrong direction. For example, the high dose patient with the largest decrease in SUVR Amyloid beta Composite with reference in the Cerebellum - 0.79 had a Week 78 CDRSB that was a 3.0 point increase (worsening) from baseline. The high dose patient with the second largest Week 78 decrease in

Amyloid beta (-0.68) via PET Composite SUVR with reference in the Cerebellum had a week 78 CDRSB, representing an increase of 0.5 from baseline. Pooling studies 301 and 302, the correlations adjusted for baseline CDRSB and baseline cerebellum SUVR for high dose were 0.145 (p=0.0375) and 0.145 (p=0.0376). However, for study 302 only (N=99), the baseline adjusted Pearson correlation was 0.104 (p=0.3090) and the adjusted Spearman correlation was 0.130 (p=0.2058). For 301 (N=110) the corresponding adjusted correlations (and nominal p-values) were 0.135 (0.1650) for Pearson and 0.074 (0.4455) for Spearman. **In summary, for high dose patients the Week 78 CDRSB change from baseline and the Week 78 PET Composite SUVR (with Cerebellum reference region) changes in A β from baseline are essentially uncorrelated[,] which raises doubts about disease modification claims and substantial evidence of effectiveness in light of the mixed results of 301 and 302.** In fact, while the correlation is nominally significant for [the] pooled high dose group[,] the corresponding regression model with the baseline scores and the biomarker change as predictors of the week 78 CDRSB change has a p-value for the significance of the variance explained by the full model of p=0.188 (not significant). With the high and low dose combined[,] the Pearson correlation was nominally significant in study 302 .168, p=.02 (and Spearman .194), but the Pearson was not significant in study 301, 0.003[,] p=0.96, and the Spearman correlation of the combined doses was in the wrong direction, i.e., a slightly negative correlation, -0.04, in study 301. Furthermore, in study 302 the low dose had a numerically bigger correlation than the high dose (adjusted Spearman correlation: 0.21 for low dose vs. 0.13 for high dose and 0.16 vs. 0.10 for adjusted Pearson correlation)[,] further complicating the interpretation if one tried to argue that these correlation magnitudes were meaningful. **Note that the p-values presented for correlation should not be overinterpreted, they are not very meaningful because they are for exploratory tertiary analyses and are not adjusted for multiplicity.**

Figure 18. Assessing Correlation of Amyloid Pet and CDRSB in High Dose at Week 78

When adjusted for all explanatory variables in the primary analysis model[,] the within patient correlation for the combined doses between week 78 CDRSB change and week 78 composite SUVR biomarker change is estimated at 0.087 $p=0.24$, not significant.

If the Week 78 SUVR biomarker change is added to the primary analysis model for CDRSB[,] it is found not to add anything significant to the model $p=0.8840$ (week 26 biomarker change $p=.7875$). **Therefore, adjusting for the other prespecified model covariates[,] there is essentially no correlation between CDRSB change and change in the biomarker composite SUVR with cerebellum reference.**

The sponsor's mediation analysis also showed Week 78 SUVR biomarker change explained a numerically higher proportion of clinical endpoint treatment effect for low dose than high: 36% low vs 33% high and the 95% bootstrap confidence intervals did not exclude 0% (p 125 of ise-appendix g6). **Thus, there is no evidence that the SUVR change is a surrogate for clinical change.** Examination of the mediation analysis model suggested that neither baseline

biomarker alone is a significant predictor of CDRSB change at Week 78 $p=.5809$; nor are both baseline and week 78 biomarker when included in the model together (baseline biomarker $p=0.3470$, Week 78 biomarker $p=0.3955$).

Table 11 presents both baseline-adjusted and baseline-unadjusted Pearson and Spearman correlations between Week 78 CDRSB and Week 78 Composite SUVR by the various possible combinations of Study and Treatment groups. **Of the 36 estimated correlations, 21 of them were either less than 0.1 or negative. The largest point estimate of the correlation is for the low dose[,] and this is only 0.22. Thus, while a few of these correlations are nominally significant, the magnitude of the correlation is still rather small and questionable for its meaningfulness.**

Table 11 Baseline Adjusted and Unadjusted Pearson and Spearman correlations between Week 78 CDRSB and Week 78 Composite SUVR by Study and Treatment groups

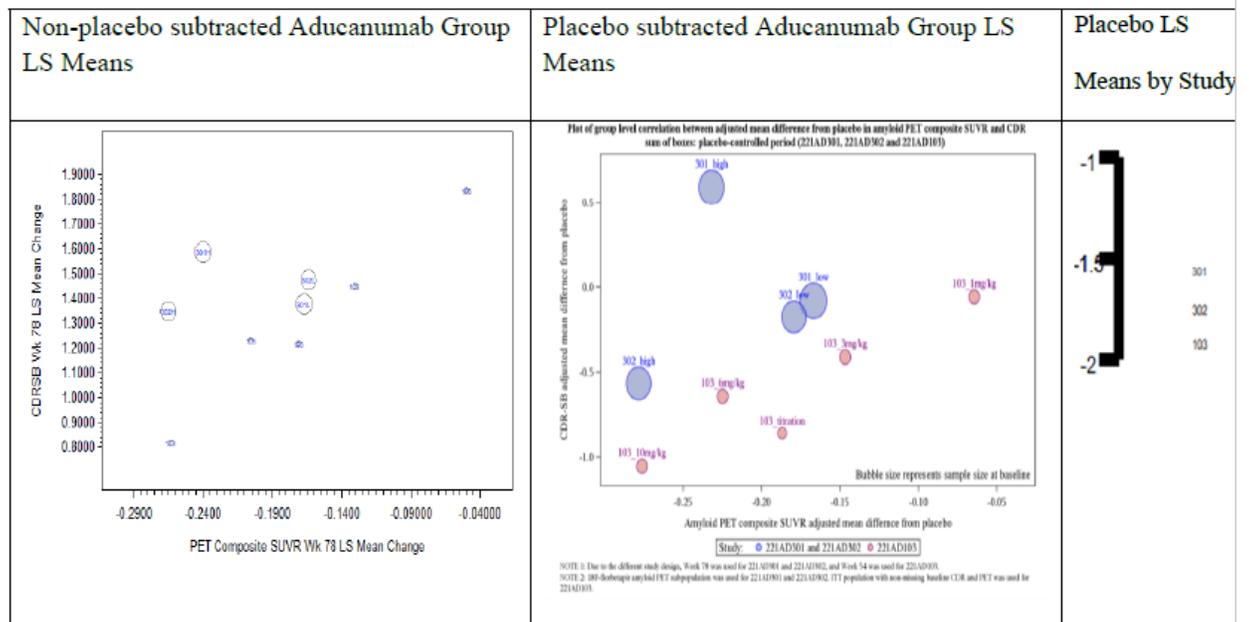
	301 High (N=110)	302 High (N=99)	301 Low (N=136)	302 Low (N=90)	301 Pooled doses	302 Pooled doses	High Pooled study	Low Pooled Study	Pooled studies
Pearson correlation	0.135 (0.160)	-0.036 (0.726)	0.009 (0.921)	0.165 (0.120)	0.026 (0.681)	0.105 (0.150)	0.084 (0.225)	0.083 (0.213)	0.066 (0.167)
Adjusted Pearson	0.135 (0.165)	0.104 (0.309)	-0.027 (0.754)	0.158 (0.142)	0.003 (0.960)	0.168 (0.021)	0.145 (0.038)	0.063 (0.350)	0.079 (0.102)
Spearman correlation	0.107 (0.265)	-0.005 (0.960)	-0.004 (0.966)	0.223 (0.034)	0.003 (0.957)	0.129 (0.078)	0.069 (0.319)	0.085 (0.203)	0.061 (0.201)
Adjusted Spearman	0.074 (0.446)	0.130 (0.206)	-0.049 (0.574)	0.211 (0.048)	0.042 (0.512)	0.194 (0.008)	0.145 (0.038)	0.054 (0.426)	0.075 (0.121)

3.2.1.4.2.2.1 Study Group Level Correlations

Figure 19 shows study group level correlations between CDRSB changes and Composite SUVR changes both at Week 78 non-placebo subtracted on the left (i.e., actual Aducanumab LSMeans) and placebo subtracted on the right (i.e., differences between Aducanumab and placebo LSMeans). Note that the differences within the same study are correlated due to sharing the same placebo and secondly, the placebo LS means are variable between studies but assumed equal[,] which is obscured by the placebo differencing. We can see that while the 301 high dose appears to be an outlier in the placebo differenced version, it appears less of an outlier and the 10 mg/kg group (high dose 301/302 equivalent) in study 103 appears more of an outlier on the left in the non-placebo differenced

version. The right figure should be interpreted cautiously when the placebo means are not very consistent between studies[,] as is the case here, and at least part of the reason why the 301 high dose appears to be an outlier is because among the 3 studies, the placebo LSMean change in CDRSB was least favorable for drug comparison in study 301.

Figure 19 Placebo Differenced and non-Placebo-differenced Aducanumab LS Means Correlations between CDRSB and SUVR at Week 78



The estimated group level Pearson correlation weighted by study size are .18 for non-placebo subtracted and .19 for placebo subtracted, respectively. Weighting by study size in the correlation analysis is necessary since otherwise the different sample sizes by study result in nonconstant variances across the 9 group level means. The unweighted analysis would also give excessive weight to study 103 estimates relative to those of [studies] 301 and 302 when study 103 estimates are much less reliable due to their much smaller sample size.

To summarize, at the group or study level, the correlation between CDRSB change and SUVR change LSMeans at Week 78 (not placebo subtracted) is only $r=0.18$ when adjusted for study sizes. Furthermore, as mentioned earlier, the proportion of the Week 78 clinical treatment effect in CDRSB explained by Week 78 change in SUVR was only 33% for the high dose and the corresponding 95% confidence interval did not exclude 0% explained. **Thus, it is not clear that Week 78 change in SUVR predicts change in Week 78 CDRSB in a meaningful way and there is no compelling evidence that Week 78 change in SUVR is a surrogate. We note that the SUVR biomarker endpoint was an exploratory endpoint with data only collected in a convenience subset of patients (33% with only 18% having Week 78 SUVR data in 302 [in 301,**

36% participation and 21% complete SUVR; in 103, 85% participation and 74% complete SUVR]]). Any formal discussion on patient-level or trial-level correlation based on incomplete data should be discouraged and the findings should be viewed as exploratory at best.⁴⁴

[Emphasis added]

Finally, previously issued draft FDA guidance on developing drugs for the treatment of early-stage Alzheimer's disease signaled an apparent agency position that should have precluded approval of aducanumab based on a surrogate endpoint. The agency's February 2013 draft guidance on this topic stated the following:

The approval of a drug for the treatment of AD [Alzheimer's disease] based on the use of a biomarker as a single primary surrogate efficacy measure can be considered under accelerated approval. However, no reliable evidence exists at the present time that any observed treatment effect on such a measure is reasonably likely to predict ultimate clinical benefit (the standard for accelerated approval), despite a great deal of research interest in understanding the role of biomarkers in AD. **Until there is widespread evidence-based agreement in the research community that an effect on a particular biomarker is reasonably likely to predict clinical benefit, we will not be in a position to consider an approval based on the use of a biomarker as a surrogate outcome measure in AD (at any stage of the illness).**⁴⁵ [Emphasis added]

There currently is not widespread evidence-based agreement in the research community that an effect on a particular biomarker is reasonably likely to predict clinical benefit in Alzheimer's disease.

In revised draft guidance issued in February 2018, although the FDA signaled willingness to entertain possible accelerated approval of a drug for treatment of early Alzheimer's disease, the agency nevertheless stated the following:

Although the issues and approaches discussed above for Stage 2 patients [patients with characteristic pathophysiologic changes of AD and subtle detectable abnormalities on sensitive neuropsychological measures, but no functional impairment] are relevant for Stage 1 patients [patients with characteristic pathophysiologic changes of AD but no evidence of clinical impact], **there is unfortunately at present no sufficiently reliable evidence that any observed treatment effect on such biomarker measures would be reasonably likely to predict clinical benefit (the standard for accelerated approval), despite a**

⁴⁴ Food and Drug Administration. Statistical review(s) for application number 761178Orig1s000, aducanumab. May 11, 2021. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2021/761178Orig1s000StatR_Redacted.pdf. Accessed August 10, 2021. PDF pages 54-58.

⁴⁵ Food and Drug Administration. Guidance for industry; Alzheimer's disease: Developing drugs for the treatment of early stage disease (draft guidance). February 2013. <https://www.regulations.gov/document/FDA-2013-D-0077-0002>. Accessed August 10, 2021.

great deal of research interest in understanding the role of biomarkers in AD. FDA strongly supports and encourages continued research in this area and stresses its potential importance in the successful development of effective treatments appropriate for use in the earliest stages of AD.⁴⁶

As discussed above, there still is not sufficient reliable evidence that decreases in brain amyloid-beta are predictive of any clinically meaningful benefit in patients with Alzheimer's disease.

In summary, the FDA's decision to approve aducanumab under the Accelerated Approval pathway based on an unvalidated surrogate endpoint — reduction of amyloid-beta plaques in the brain — was deeply flawed and not supported by the available scientific evidence.

Conclusions

Based on the currently available scientific evidence, reduction of amyloid-beta plaques in the brain is not a clinically meaningful outcome for Alzheimer's disease patients and is not sufficient for establishing a drug as being reasonable and necessary for treatment of such patients.

Health outcomes that would be important for patients receiving a drug treatment for Alzheimer's disease would be substantial slowing of declines on measures of cognitive function. Such clinically meaningful health outcomes must be demonstrated in large, randomized, placebo-controlled trials that have been completed and analyzed in accordance with prespecified statistical analysis plans described in the trial protocols. The phase 3 clinical trials for aducanumab utterly failed to meet these criteria.

In closing, Public Citizen urges CMS to issue a national coverage determination that excludes aducanumab from coverage under the Medicare program because there is a lack of scientific evidence that aducanumab provides any meaningful clinical benefit in terms of cognitive function outcomes in Alzheimer's disease patients and the drug thus is not reasonable and necessary for treatment of such patients. The lack of evidence that aducanumab provides meaningful clinical benefit and the fact that the reduction in amyloid-beta plaques following administration of the drug is not an established, validated surrogate endpoint for such clinical benefit, combined with the known risks of the drug, clearly justify its exclusion from Medicare coverage.

Thank you for the opportunity to comment on this important public health issue.



Michael A. Carome, M.D.
Director
Public Citizen's Health Research Group

⁴⁶ Food and Drug Administration. Early Alzheimer's disease: Developing drugs for the treatment; guidance for industry (draft guidance). February 2018. <https://www.regulations.gov/document/FDA-2013-D-0077-0043>. Accessed August 10, 2021.