



1600 20th Street, NW • Washington, D.C. 20009 • 202/588-1000 • www.citizen.org

December 9, 2020

The Honorable Christi A. Grimm
Principal Deputy Inspector General
Office of Inspector General
U.S. Department of Health and Human Services
330 Independence Avenue SW
Washington, DC 20201

RE: Request for an Office of Inspector General investigation of the Food and Drug Administration's inappropriate close collaboration with Biogen before and after the submission of the biologics license application for aducanumab for treatment of Alzheimer's disease

Dear Principal Deputy Inspector General Grimm:

Public Citizen, a consumer advocacy organization with more than 500,000 members and supporters nationwide, respectfully requests that your office immediately launch a formal investigation to scrutinize the unprecedented close collaboration between the Food and Drug Administration (FDA) and Biogen that occurred before and after the submission of Biogen's biologics license application (BLA) for the new biologic drug aducanumab for treatment of Alzheimer's disease.

Such close collaboration — which was made fully 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 — has 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 two identical pivotal phase 3 clinical trials of the drug. ON Director Billy Dunn, M.D., 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 Dunn's presentation at the November 6, 2020, advisory committee meeting was not supported by an objective review of data from the pivotal phase 3 clinical trials, which had been terminated early after enrollment had reached only 50% of the planned target enrollment because a planned prespecified interim analysis showed the trials were unlikely to yield evidence that the drug was effective for treating Alzheimer's disease.

The FDA's close collaboration with Biogen before and during the agency's review of the BLA for aducanumab is evidence of regulatory capture at the agency and is reminiscent of the regulatory capture on the part of the Federal Aviation Administration that resulted in grossly

insufficient regulatory oversight of the Boeing 737 MAX aircraft and contributed to the chain of errors that led to the crashes of Lion Air flight 610 in 2018 and Ethiopian Airlines flight 302 in 2019, causing the preventable deaths of 346 people.¹

The following is a detailed discussion of the evidence of the close collaboration between the FDA and Biogen and the ensuing regulatory capture on the part of the agency.

Background on aducanumab

Aducanumab, which Biogen is developing as a potential treatment for Alzheimer's disease, is an experimental recombinant human monoclonal antibody targeting amyloid- β multimers.²

Like the prior 22 unsuccessful experimental drugs targeting amyloid- β 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 plaques in the brain *causes* the neuronal degeneration seen in Alzheimer's disease.³

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 and Study 302).⁴ 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.⁵ That action should have marked the end

¹ The House Committee on Transportation & Infrastructure. Final Committee Report: The Design, Development & Certification of the Boeing 737 MAX. September 2020. <https://transportation.house.gov/imo/media/doc/2020.09.15%20FINAL%20737%20MAX%20Report%20for%20Public%20Release.pdf>. Accessed November 28, 2020.

² Panza F, Seripa D, Solfrizzi V, et al. Emerging drugs to reduce abnormal β -amyloid protein in Alzheimer's disease patients. *Expert Opin Emerg Drugs*. 2016;21(4):377-391.

³ 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.

⁴ 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 November 28, 2020. PDF pages 22-23.

⁵ 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 November 28, 2020.

of aducanumab as a potential treatment for Alzheimer's disease, at least as it pertains to the studies thus far completed.

Subsequent unprecedented close collaboration between the FDA and Biogen

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.⁶ 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]⁷

On December 5, 2019, Biogen presented topline results of Studies 301 and 302 at the Clinical Trials on Alzheimer's Disease (CTAD) 2019 conference. The post hoc analyses conducted by Biogen showed that in Study 301 aducanumab at both the low and high dosing regimens did not show improvement in the trial's primary efficacy endpoint, whereas in Study 302 the drug at only the high dosing regimen resulted in statistically significant but small improvements in the primary and several secondary efficacy endpoints.⁸

In a July 8, 2020, press release announcing the completion of its BLA submission for aducanumab to the FDA, Biogen reported that the "completed submission followed ongoing collaboration with the FDA."⁹

On November 4, 2020, the FDA posted on its website the briefing documents for the FDA's PCNS Drugs Advisory Committee meeting on November 6, 2020. Disturbingly, the primary briefing document was a 139-page document written jointly by the FDA and Biogen.¹⁰ Most of the primary briefing document content appears to have been written by Biogen (with most sections of the document beginning with the heading "The Applicant's Position"). Interspersed within the sponsor's content were generally brief text boxes (ranging from a single sentence to

⁶ 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 November 17, 2020.

⁷ Biogen. Aducanumab update. October 22, 2019. <https://investors.biogen.com/static-files/5a31a1e3-4fbb-4165-921a-f0ccb1d64b65>. Accessed November 28, 2020.

⁸ Biogen. EMERGE and ENGAGE topline results: Two phase 3 studies to evaluate aducanumab in patients with early Alzheimer's disease. Presented at CTAD 2019 on December 5, 2019. <https://investors.biogen.com/static-files/ddd45672-9c7e-4c99-8a06-3b557697c06f>. Accessed November 28, 2020.

⁹ Biogen. Biogen completes submission of biologics licensing application to FDA for aducanumab as a treatment for Alzheimer's disease. July 8, 2020. <https://investors.biogen.com/news-releases/news-release-details/biogen-completes-submission-biologics-license-application-fda>. Accessed November 29, 2020.

¹⁰ 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 November 29, 2020. PDF pages 22-23.

one page and containing the heading “The FDA’s Position”) written by the FDA, most of which endorsed or concurred with Biogen’s presentation of information, analyses, and conclusions. Public Citizen, which has had staff experts attend and testify before hundreds of FDA advisory committee meetings over the past five decades, does not recall ever seeing an advisory committee meeting briefing document that was explicitly written jointly by the FDA and the sponsor of the medical product being considered by the committee. A separate briefing document written solely by the FDA is an obvious prerequisite for an independent and objective agency assessment of the clinical trial data presented by a sponsor in support of a new drug application (NDA) or BLA.

The joint briefing document for the PCNS Drugs Advisory Committee meeting revealed further details of the close collaboration that occurred between the FDA and Biogen over more than a year following the company’s March 2019 decision to terminate the pivotal phase 3 trials testing aducanumab because of the interim futility analysis. For example, according to Biogen, the company had a June 2019 meeting with the FDA that included a discussion of post hoc analyses of data from Study 302 conducted after termination of the study showing apparently positive results.¹¹ According to Biogen, the FDA stated the following at this meeting:

It is imperative that extensive resources be brought to bear on achieving a maximum understanding of the existing data. Given the wholly unique situation that is the current state of the aducanumab development program ..., those further analyses would best be conducted as part of a **bilateral effort involving the Agency and sponsor, i.e., through a ‘workstream’ or a ‘working group’ collaboration.**¹² [Emphasis added]

The FDA confirmed the accuracy of Biogen’s above statement about the June 2019 meeting describing what appears to have been an extraordinary plan for a collaborative working group involving Biogen and the FDA staff, noting that “[g]iven the unmet medical need and unique nature of the data, FDA and the Applicant agreed upon a plan for further analyses and FDA formally engaged with the Applicant through Type C Meetings on four occasions between June 2019 and June 2020, the last of which included a discussion of pre-BLA questions submitted by the Applicant.”¹³ The FDA likewise stated in its introduction to the joint briefing document that “the FDA recognized that additional work was necessary to achieve a maximum understanding of the results and established a collaborative plan for further rigorous analyses of the data.”¹⁴

A joint review document tilted heavily in favor of Biogen’s position

Although we understand that it is not unusual for the FDA to meet with sponsors and provide advice regarding the development of drugs, the design of clinical trials, and the statistical analyses of trial data, among other things, the close collaboration that occurred between the agency and Biogen in conducting post hoc analyses of data from the aducanumab clinical trials is, to our knowledge, unprecedented. Typically, sponsors conduct their own detailed statistical analyses of clinical trial data supporting NDAs and BLAs, and the FDA then conducts its own independent analyses of the data following submission of these applications for approval. Such

¹¹ *Ibid.* PDF page 19.

¹² *Ibid.* PDF pages 19-20.

¹³ *Ibid.* PDF page 20.

¹⁴ *Ibid.* PDF page 10.

appropriate separation between the clinical trial data analyses conducted by the sponsor and those conducted by the FDA is critical to maintaining the independence and integrity of the FDA's review of the data.

In the case of aducanumab, the close collaboration between the FDA and Biogen in the post hoc analyses of clinical trial data and the subsequent joint authorship of the primary briefing document for the November 6, 2020, PCNS Drugs Advisory Committee meeting resulted in a one-sided consensus briefing document. That document overwhelmingly emphasized the post hoc analyses that yielded positive results suggesting that high-dose aducanumab was an effective treatment for Alzheimer's disease (primarily the analyses of data from Study 302) but significantly downplayed the results of post hoc analyses showing that aducanumab was not effective for treating Alzheimer's disease (analyses of Study 301, which was essentially identical to Study 302).

Relying on dubious statistical gymnastics, Biogen and the FDA in their joint review document sought to discount the discordance between the negative results of Study 301 and the positive results of Study 302 and portray the post hoc analyses of Study 302 data (with supporting data from the small phase 1b Study 103 that was not even designed to assess efficacy) as representing the true picture of aducanumab's effectiveness in treating Alzheimer's disease. This "cherry-picking" approach is neither statistically nor scientifically appropriate.

For example, Biogen presented an 18-page summary of the analyses of Study 302 that concluded thusly: "Study 302 is a positive study, providing the primary contribution to the substantial evidence of the effectiveness of aducanumab."¹⁵ The FDA's subsequent inserted position text box included the following conclusions:

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

Study 302 is a **strongly positive study** on multiple distinct and important clinical measures, robust to numerous sensitivity analyses, and supported by well-characterized biomarker data.¹⁶

[Emphasis added]

Biogen then provided only a four-page summary of the analyses of Study 301, which was followed by the following conclusions from the FDA:

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

Upon initial review, the one positive study (Study 302) and one negative study (Study 301) were given equal weight and consideration. Despite divergent outcomes in the

¹⁵ *Ibid.* PDF pages 39-57.

¹⁶ *Ibid.* PDF page 57.

primary endpoint, there were some key similarities between the two studies. The low-dose aducanumab treatment arms, while not statistically significant, demonstrated consistent numerical effects favoring aducanumab in the studies. Also, aducanumab produced a time- and dose-dependent reduction in brain amyloid burden. Upon closer review of the individual studies, Study 302 appeared to be a strongly positive study on many distinct clinical measures, robust to numerous sensitivity analyses, and supported by well-characterized biomarker data. In the context of a positive Study 302, the suggestion of a dose-response relationship observed in Study 103, and the numerically favorable results of similar magnitude in the low-dose groups in both studies, the high-dose group in Study 301 tends to stand apart not only for the negative outcome on the primary endpoint, but also the difference in biomarker profiles compared to Study 302.¹⁷

Biogen then presented a series of post hoc analyses intended to assess the effect of the negative Study 301 on the purported persuasiveness of Study 302 given the divergent results between the two trials.¹⁸ According to Biogen, the “overarching hypothesis [for these analyses] was that if aducanumab were an effective treatment for Alzheimer’s disease, then there should be patients in Study 302, the positive study, that drive the overall effect. Additionally, in such a situation, a set of patients in Study 301, the failed study, with those same characteristics should show a response similar to that seen in Study 302.”¹⁹

Commenting on this approach, the FDA stated the following:

In general, the FDA agrees with the Applicant’s characterization of the hypotheses tested. **If aducanumab is effective**, it follows that Study 302 is a positive study and that there would be patients in Study 301 who, based on certain characteristics, should show response of similar character to patients in Study 302...²⁰

The purpose of these analyses is to provide maximum understanding of the partially discordant results and to determine if this understanding precludes independent consideration of Study 302...²¹

A guiding principle of the hypothesis was that **if aducanumab is effective and the effect is dose-related as in Study 302**, it follows that patients in Study 301 with adequate and consistent dosing should also demonstrate an effect on clinical endpoints. An absence of an effect in this subgroup of patients in Study 301 would diminish the persuasiveness of Study 302. Although it is impossible to fully account for all factors that may contribute to findings in subgroups formed by post-randomization factors (i.e., dosing), a variety of approaches, each with strengths and limitations, **appears to show that consistent exposure to high doses of aducanumab does lead to similar treatment effects in the two studies.**²²

[Emphasis added]

¹⁷ *Ibid.* PDF page 61.

¹⁸ *Ibid.* PDF pages 62-82.

¹⁹ *Ibid.* PDF page 62.

²⁰ *Ibid.* PDF page 64.

²¹ *Ibid.* PDF page 64

²² *Ibid.* PDF page 83.

The following summary conclusions regarding the overall assessment of the efficacy of aducanumab for the treatment of Alzheimer's disease was provided by the FDA:

Study 302 provides the primary evidence of effectiveness of aducanumab. The effect of aducanumab in Study 302 is **robust and exceptionally persuasive** on several of the instruments used to evaluate efficacy...

The results of Study 103 [the small phase 1b trial] **are appropriately viewed as supportive evidence of the effectiveness of aducanumab**. Despite the limitations of a trial designed to assess safety and tolerability rather than effectiveness, the 10 mg/kg dose arm was able to achieve statistical significance according to the prespecified analysis plan...

Study 301 is a negative study and does not contribute to the evidence of effectiveness of aducanumab... The rapid progressor analysis indicated that a small imbalance in the number of rapid progressing patients in the high-dose arm in Study 301 had a disproportionate impact on the estimate of the treatment effect using the primary analysis method. An examination of dosing in Study 301 indicates that patients with higher exposure to the 10 mg/kg dose in Study 301 had similar responses to patients in Study 302. **These two factors contribute to the overall understanding of Study 301 and together do not meaningfully detract from the persuasiveness of Study 302.** There were no findings from the exploration that represented evidence that aducanumab is not effective.²³

[Emphasis added]

The FDA's remarkable comments above are indicative of the agency going to extraordinary lengths to slant its assessment to be in agreement with the sponsor's position by giving much greater weight to the positive results of Study 302 and discounting the overall negative results of Study 301. Importantly, from a scientific and regulatory standpoint, the type of post hoc analyses undertaken by Biogen in close collaboration with the FDA were highly susceptible to bias and should only be used to generate hypotheses for future clinical trials that would be needed to establish whether aducanumab is effective for treating Alzheimer's disease, not as a basis for aducanumab approval.

Most troubling, the post hoc analytic approach undertaken collaboratively by Biogen and the FDA essentially started with the assumption that the null hypothesis for these trials (i.e., that aducanumab is *not* effective for the treatment of Alzheimer's disease) was *false* and then tried to understand why Study 301 yielded negative results. The more scientifically valid, unbiased approach would have been to start with the assumption that the null hypothesis for the pivotal phase 3 trials was *true* and that the Study 302 efficacy data represented a *false positive*.

Indeed, viewed in the context of the two-decade history of 22 failed drugs targeting amyloid- β accumulation, including five other anti-amyloid- β monoclonal antibodies (see the Table in Appendix 1 at the conclusion of our letter), there was a significant probability a priori that the Study 302 efficacy results represented a false positive.

²³ *Ibid.* PDF page 94.

The FDA statistical review

Appended to the joint briefing document for the PCNS Drugs Advisory Committee meeting were two FDA review documents. The first was a clinical review written by Kevin Krudys, Ph.D., Clinical Efficacy Reviewer, Division of Neurology 1, ON, CDER, that provided analyses of the data from Studies 301, 302, and 103 and conclusions that were fully concordant with the primary joint briefing document.²⁴

The second appended review document was the draft 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.²⁵ Notably, Dr. Massie is not under the supervisory chain of Dr. Dunn. Dr. Massie's review clearly indicates that he is 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 were stopped early for futility... when both studies had reached 50% completion (thus, in a sense, together equivalent in information to a completed study) since **it was estimated based on the interim study-pooled estimate of the treatment effects that both studies had <20% chance of success for either dose if completed.** Following...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 [BLA] 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 CDR-SB [Clinical Dementia Rating-Sum of Boxes, the primary efficacy outcome measure] which is the only timepoint that shows any significance and that is only significant in one of the two studies (the first study high dose is numerically worse than placebo at Week 78 on the primary endpoint).** A worse placebo response in study 302 than was observed in study 301 could explain the significance of study 302 (p=0.01).

There is a reason why two positive studies has been the standard in Alzheimer's, e.g., the need for reproducibility and adequate strength of evidence, in a disease with soft (more subjective and variable than mortality) endpoints. **This BLA submission does not have a situation such as just one study in existence and for which that study is strong.** We have a second large adequate well controlled study that directly contradicts the first and is not even close to significance p=.8252. **Under the null hypothesis (no drug effect),**

²⁴ *Ibid.* PDF pages 141-245.

²⁵ *Ibid.* PDF pages 247-343.

there is a .0975 chance of at least one type I error across 2 studies. If one has two studies and takes the best and pretends like it's the only study, one's estimate is most likely biased and misleading... 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... Any selection of patients would need a proper placebo control, that is the regulatory standard in Alzheimer's. The overall 301 primary result is the only valid well controlled, multiplicity adjusted, randomization validated analysis of 301 (and it had a substantial sample size).

The sponsor tries to discount study 301 due to post-hoc defined 'rapid progressors'. Rapid progressors are part of the reality of Alzheimer's and after the fact it is too late to address them in a completed large randomized study. A highly effective drug would not be likely to fail because of rapid progressors especially in the early stages of a disease. **Study 302 could just as well be the outlier relative to the true proportion of outliers in the natural progression...**

The sponsor's analysis of Study 103, 10 mg/kg vs. pooled placebo arms, is not supported by the randomization (3 of the placebo arms had no chance of receiving 10 mg/kg and one was entirely APOE carriers, which 10 mg/kg was not). Outside of rare diseases[,] there is no justification for an analysis involving the pooling of staggered arms that is not supported by the study's overall randomization scheme.²⁶

[Emphasis added]

Dr. Massie made the following conclusions and recommendations at the end of his review:

The totality of the data does not seem to provide sufficient evidence to support the efficacy of the high dose. There is much inconsistency and no replication. There is only one positive study at best and a second study which directly conflicts with the positive study. Both studies were not fully completed as they were terminated early for futility and had sporadic unblinding for dose management of ARIA [amyloid-related imaging abnormalities] cases which was much higher in the drug group... Therefore, there is no convincing evidence of delaying clinical progression cognitive or functional: only a single positive timepoint (unreplicated and conflicted by a second study) and no delayed start design (termination for futility does not help with completeness or interpretability of long term follow up)... In addition, the low dose in study 301 was numerically better than the high dose despite having no 10 mg/kg doses and this comparison is supported by randomization. **For these reasons, a study fully completed according to protocol without major non-prespecified amendments while the study is ongoing is needed to confirm or deny the positive study or the negative phase 3 study.**²⁷ [Emphasis added]

Strikingly, although the FDA portions of the primary joint briefing document for the PCNS Drugs Advisory Committee meeting significantly echoed the content of Dr. Krudys' clinical review document, they offered no hint of the content of the careful and detailed critiques

²⁶ *Ibid.* PDF pages 253-256.

²⁷ *Ibid.* PDF page 342.

provided by the FDA's own lead statistical reviewer. That ignored statistical review offered a sweeping and cogent refutation of the post hoc analyses conducted and emphasized collaboratively by Biogen and other FDA staff.

The November 6, 2020, PCNS Drugs Advisory Committee meeting regarding aducanumab

On November 4, 2020, the FDA posted on its website for the November 6, 2020, PCNS Drugs Advisory Committee meeting prerecorded presentations made by Biogen and by certain FDA staff who were involved in the review of Biogen's BLA for aducanumab²⁸ — including Dr. Krudys, who presented the clinical overview of efficacy; Natalie Branagan, M.D., Medical Officer/Safety Team, Division of Neurology 1, ON, CDER, and Brian Trummer, M.D., Ph.D., Medical Officer, Division of Neurology 1, ON, CDER, who presented the clinical overview of safety; and Dr. Massie, who presented the statistical review — as well as the slide sets^{29,30} and transcripts of these presentations.^{31,32}

Unsurprisingly, the presentations by Biogen and all FDA reviewers, *except Dr. Massie*, were completely concordant with the content of the aforementioned primary joint review document written by Biogen and the FDA and overall reflected enthusiastic support for aducanumab as an effective treatment for Alzheimer's disease, primarily based on post hoc analyses of incomplete Study 302 and the phase 1b Study 103. Dr. Massie's presentation was consistent with his strongly worded critical written statistical review of Biogen's BLA for aducanumab.

During the November 6, 2020, PCNS Drugs Advisory Committee meeting, Samantha Budd Haeblerlain, Ph.D., Senior Vice President, Head Neurodegeneration Unit, Biogen, gave the sponsor's summary presentation.³³ Her presentation reiterated the numerous post hoc analyses of Studies 301, 302, and 103 conducted collaboratively by Biogen and the FDA that were detailed in the joint briefing document for the meeting. She concluded her presentation with the following comments:

In closing, as you've seen, after an extensive review by Biogen and the FDA, it's clear that Study 302, with support from Study 103 and compelling mechanistic evidence provided by the biomarkers, provides substantial evidence of effectiveness, and Study

²⁸ Food and Drug Administration. FDA and Biogen pre-recorded presentation links for the November 6, 2020 meeting of the Peripheral and Central Nervous System Drugs Advisory Committee.

<https://www.fda.gov/media/143508/download>. Accessed November 29, 2020.

²⁹ Food and Drug Administration. Biogen 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/143506/download>. Accessed November 29, 2020.

³⁰ 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>. November 20, 2020.

³¹ Food and Drug Administration. Biogen 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/143507/download>. Accessed November 29, 2020.

³² 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 November 29, 2020.

³³ 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 00:16:44 to 0:56:14.

301 does not detract from this understanding. Across the three studies, in patients with Alzheimer’s disease who had consistent exposure to 10 mg/kg, aducanumab demonstrated a reduction in clinical decline... So, given the totality of the evidence, we can conclude that the benefit-risk profile for aducanumab is favorable and potentially prolongs patients’ independence by several months, even a few years...³⁴

Following Dr. Haeberlain’s presentation, several advisory committee members — clearly not subject to the regulatory capture that appears to have compromised the FDA’s independence and objectivity during its review of Biogen’s BLA for aducanumab — 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 (see Appendix B for additional examples):

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? ...³⁵

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? ...**³⁶

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.³⁷

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,

³⁴ *Ibid.* Approximately 00:53:32 to 00:54:41.

³⁵ *Ibid.* Approximately 01:00:27 to 01:01:02.

³⁶ *Ibid.* Approximately 01:02:21 to 01:02:38.

³⁷ *Ibid.* Approximately 01:06:08 to 01:06:31.

301 and 302. **What you haven't discussed is the possibility that the placebo groups differed.**³⁸

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.³⁹

Following committee members' questions to Biogen, Dr. Dunn, Director, Office of Neuroscience, Office of New Drugs, CDER, gave the FDA's summary presentation.⁴⁰ The language he used during his presentation at the meeting, as in the following excerpts, made him sound more like a consultant hired by Biogen to endorse the company's BLA for aducanumab, than like an independent and objective federal regulator paid by American taxpayers:

I'm going to spend the next few minutes discussing some of the issues involved in the consideration of the aducanumab marketing application and **why the evidence supporting its approval appears strong...**⁴¹

It was apparent that if the results presented at that meeting [between Biogen and the FDA] did, in fact, represent the true effect of aducanumab, it was imperative that all efforts would be made to understand how reliable the results were and to achieve a maximum understanding of the data giving rise to these results so as to determine both the reliability and the impact of Study 301's results on the interpretation of Study 302. Taken on face, even on initial viewing of the data in May of 2019, it was apparent that the results of Study 302, again taken on face, had the potential to represent **exceptionally persuasive evidence of effectiveness**. Therefore, the FDA proposed a collaborative

³⁸ *Ibid.* Approximately 01:10:36 to 01:11:07.

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

⁴⁰ *Ibid.* Approximately 01:34:01 to 02:22:22.

⁴¹ *Ibid.* Approximately 01:34:08 to 01:34:17.

effort that would be conducted with the applicant in order to achieve a maximum understanding of the data to inform appropriate advice regarding the future development of aducanumab...⁴²

The FDA advised the applicant that the development of aducanumab should not be abandoned as the available clinical data suggest the drug may be clinically active and...**the data do not provide convincing evidence that the drug is ineffective...**⁴³

Taken together, multiple lines of evidence regarding both similarities and differences between Studies 301 and 302 suggest the partially discrepant results between Studies 301 and 302 are qualitatively sufficiently well understood to allow for **independent consideration of the persuasiveness of study 302...**⁴⁴

Study 302 provides the primary evidence of effectiveness of aducanumab. **The effective of aducanumab in Study 302 is robust and exceptionally persuasive** on several of the instruments used to evaluate efficacy...⁴⁵

When considered on its own, **Study 302 would appear to be a home run...**⁴⁶

Speaking about the impact that the COVID-19 pandemic has had on clinical trials — which seemed immaterial because the pandemic occurred well after Biogen’s trials for aducanumab were completed — Dr. Dunn talked at length about the need for the agency to be “flexible and sensible” in applying the substantial evidence standard when reviewing pivotal clinical trials used to establish the efficacy of drugs. In this context, he made the following troubling statement:

Dr. [Peter] Stein [Director of CDER’s Office of New Drugs] noted that we are working on advising sponsors on what kinds of sensitivity analyses sponsors should consider in managing the situation. He says that all I can say is at the end of the day, we certainly recognize the impacts. We can’t change our substantial evidence standard, but we can be sensible and flexible about how it is applied, the kinds of information we’re looking at, and try to be as sensible as possible. We don’t want to see drugs that are potentially very effective delayed. Dr. Stein pointed out that the Office of New Drugs is committed to helping sponsors overcome the challenges caused by COVID-19. We will be working very cooperatively with sponsors, assessing what kind of analysis can be done to make sure that the data can be put together in a way that is convincing and persuasive and lets us get to an approval decision, where appropriate.

We rigorously assessed the impact of the early termination [of Studies 301 and 302] and determined that it is not an issue. The data represent accurately the effects of aducanumab in the two trials. **With that established, it appears obvious that [Study] 302 is independently extremely persuasive.**⁴⁷

⁴² *Ibid.* Approximately 01:40:31 to 01:41:20.

⁴³ *Ibid.* Approximately 01:43:08 to 01:43:21.

⁴⁴ *Ibid.* Approximately 02:07:10 to 02:07:29.

⁴⁵ *Ibid.* Approximately 02:12:30 to 02:12:42.

⁴⁶ *Ibid.* Approximately 02:17:01 to 02:17:06.

⁴⁷ *Ibid.* Approximately 02:18:41 to 02:19:39.

As was the case with the main joint briefing document written by Biogen and the FDA, Dr. Dunn's presentation made no mention of the careful, detailed critique provided by the FDA's own lead statistical reviewer, Dr. Massie, that substantially rebutted the post hoc analyses of clinical trial data conducted collaboratively by Biogen and the FDA.

During the discussion of the eight questions (including four voting questions) that the FDA asked the advisory committee members to address, committee members unleashed a torrent of appropriately harsh criticism of the post hoc analyses of Studies 301, 302, and 103; the nature and organization of the questions posted by the FDA; the FDA's collaborative review process; and the one-sided joint briefing document. The following are some illustrative examples of the committee members' comments (presented in the order in which they were made during the meeting; see Appendix B for numerous additional examples):

Dr. Caleb Alexander (during discussion of question 1 [*The primary evidence of effectiveness presented in support of aducanumab for the treatment of Alzheimer's disease is provided by Study 302. Discuss the evidence of effectiveness provided by Study 302, viewed independently and without regard to Study 301, with particular consideration of the size of the study, design of the study, analysis of results to assess the effects of the drug, and consistency of results among various subgroups in the study.*]):

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 [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.⁴⁸

Dr. Emerson (during discussion of question 1):

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-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

⁴⁸ *Ibid.* Approximately 03:35:22 to 03:36:30.

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.**⁴⁹

Michael Gold, M.S., M.D., Vice-President, Neurosciences Development, AbbVie, North Chicago, Illinois; PCNS Drugs Advisory Committee Non-Voting Industry Representative (during discussion of question 1):

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.⁵⁰

Dr. Emerson (during discussion of question 1):

Well, I would then like to just register that **I have not been super-impressed with how the briefing book and presentations have gone from the FDA for this study.** I feel that ...to a certain extent the clock has been run out, and we haven't been able to ask questions, that **mainly the FDA just gave us just conclusions and not results.** And so now, ...we have trouble discussing this because it's all being supplanted by saying look over here and answer this irrelevant question, and we aren't really going to give you the opportunity to say how this study should be interpreted if we want to ignore the numbers from [Study] 301. **But we may never, ever, ever, ever ignore the fact that [Study] 301 was done.**⁵¹

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 (during discussion of question 1):

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...⁵²

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.⁵³

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 (during discussion of question 1):

⁴⁹ *Ibid.* Approximately 03:40:08 to 03:41:24.

⁵⁰ *Ibid.* Approximately 03:43:07 to 03:43:28.

⁵¹ *Ibid.* Approximately 03:48:17 to 03:49:02.

⁵² *Ibid.* Approximately 04:03:16 to 04:03:26.

⁵³ *Ibid.* Approximately 04:03:47 to 04:04:14.

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...⁵⁴

Dr. Perlmutter (during the discussion after the vote on question 2 [*Does Study 302, viewed independently and without regard to Study 301, provide strong evidence that supports the effectiveness of aducanumab for the treatment of Alzheimer's disease? Committee vote: 1 YES, 8 NO, 2 UNCERTAIN*])

I voted no... 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.⁵⁵

Dr. Emerson (during discussion after the vote on question 2):

This is the **first time I've heard an FDA person say that statistical significance automatically was clinical importance**.⁵⁶

Dr. Alexander (during discussion of question 3 [*The primary evidence of effectiveness presented in support of aducanumab for the treatment of Alzheimer's disease is provided by Study 302. Study 103 is presented as supportive evidence of aducanumab's effectiveness. Discuss the evidence of effectiveness provided by study 103.*]):

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**.⁵⁷

Dr. Kesselheim (during discussion of question 3):

I also wanted to...raise the point that **it is challenging to ask us to...identify supportive evidence...for a trial [i.e., Study 302] that's already of questionable strength. In a**

⁵⁴ *Ibid.* Approximately 04:09:48 to 04:10:59.

⁵⁵ *Ibid.* Approximately 04:26:37 to 04:27:11.

⁵⁶ *Ibid.* Approximately 04:31:39 to 04:31:46.

⁵⁷ *Ibid.* Approximately 04:33:45 to 04:34:15.

trial like [Study] 103, that was not designed to provide supportive evidence but was designed for evaluation of other things, of which the efficacy...measurements were...kind of a supplementary or secondary component of that analysis. And as a result, I think that's why you're getting in [Study] 103 efficacy results that seem very discordant from the efficacy results that you see in [Study] 302, in addition to the fact that the efficacy results are observed over the course of a 54-week study, whereas in figure 5 of the FDA documents, there doesn't appear to be any effect of the high-dose group [in Study 302] at 50 weeks of analysis. **And so, their discord is not only in the level of the effect size but in the timing of the effect...**⁵⁸

Dr. Emerson (during discussion of question 3):

This is exactly the point that I wanted to make. [Study] 103 was a preliminary screening trial. Had it been completely negative, [Studies] 301 [and] 302 would have never been done. So...**phase 2 studies are always positive in some way**, and what's nice about [Study] 103 in this particular case is, I viewed the modifications to the eligibility criteria and what else they were doing as relatively slight, so, you know, there's other times where you're chasing after subgroups and you're just saying it's there. **But every phase 2 study is so impossibly biased in its treatment effect that you should never be surprised when...you get less results in the confirmatory study...**

The thing that bothered me the most is — **again due to pressures of time and again this is a direct complaint, so much time spent telling me things were excessively understood and very persuasive and not enough time looking at the data** — I never got to really delve into what the problems were with the randomization schemes and direct comparisons and particularly direct comparisons by randomization comparisons superimposed on the [Study] 302 results before I believe it was very supportive. So, there is something to be gained if you told me you had [Study] 302 with no phase 2 study, and I'd say, 'great, give me two more confirmatory studies.' **But in no sense would I regard that [Study] 103 is going to [take] the place of another confirmatory study.** That doesn't make me relax criteria for what would regard [Study] 302 as pivotal. And just note that an underpowered study decreases the positive predictive value of a positive result. Lots of people go, well yeah, it was a small study, but the effect was huge. Well they've got cause and effect wrong. In order for a small study to be statistically significant, it has to have a huge effect — it has to...**But that doesn't mean it's correct, and by the time you say we're not taking all results, were only taking it when it's positive, it's a very, very biased result.** So, the positive predictive value, we don't just want to worry about the type 1 error, which says make certain we don't approve distilled water and the sponsor wants to say if we have an effective drug it really works. That's the power, but we are concerned with the Bayesian positive predictive value, and that in an underpowered study and one in which you let the type 1 error creep up, it's very low. This is the reason why confirmatory studies, depending upon the disease area and depending upon how much we know about it,...why do anywhere between 20% and 70% of the phase 3 studies confirm the phase 2 results, and it has to do with that positive predictive value.⁵⁹

⁵⁸ *Ibid.* Approximately 04:44:46 to 04:45:58.

⁵⁹ *Ibid.* Approximately 04:52:42 to 04:55:58.

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 (during the discussion of question 5 [*The application presents evidence in support of effects on the pathological hallmarks of Alzheimer's disease, including effects on amyloid beta, tau, and downstream markers of neurodegeneration using multiple assessment modalities. Discuss the impact of these results.*]):

I think the evidence is fairly compelling that there is an effect on [amyloid] beta in the brain... A number of us are having difficulty with the lack of an association between the CDR-SB and the PET [positron emission tomography] imaging. **I think it would be very helpful if the statistician and the other members of the FDA team had come together and tried to understand the discrepancy between the two analyses.**⁶⁰

Dr. Alexander (during discussion after the vote on question 6 [*Has the applicant presented strong evidence of a pharmacodynamic effect on Alzheimer's disease pathophysiology? Committee vote: 5 YES, 0 NO, 6 UNCERTAIN*]):

I also want to just note **the briefing packet was unique in that it was co-produced [by Biogen and the FDA]**, and I do think there's some merit in having separate packets produced by both parties, or at a minimum having the FDA provide the briefing materials and having the sponsor add their commentary to the FDA's review rather than vice versa, **given the FDA's role as regulator here.**⁶¹

Dr. Emerson (during discussion of question 7 [*Study 301 was a negative study. Post hoc exploratory analyses were conducted in order to achieve maximum understanding of the partially discordant results of Study 301 and Study 302, and to determine if this understanding precludes independent consideration of Study 302. Additional contribution to the understanding of aducanumab's pharmacological activity and clinical effects is provided by the results of study 103. In light of the exploratory analyses that were conducted and the results of Study 103, discuss the impact of the results of Study 301 on the consideration of the results of study 302.*]):

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

⁶⁰ *Ibid.* Approximately 05:22:17 to 05:22:48.

⁶¹ *Ibid.* Approximately 05:40:04 to 05:40:26.

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.⁶²

Dr. Gold (during discussion of question 7):

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...⁶³

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.**⁶⁴

Dr. Duda (during discussion of question 7):

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.⁶⁵

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.**⁶⁶

Presumably embarrassed by the advisory committees' withering criticism of the FDA's approach to the review of the clinical trial data presented in Biogen's BLA for aducanumab, Dr. Dunn interjected near the end of the discussion of question 7 with the following remarkable, defensive comments in an apparent attempt to rewrite the history of the agency's close collaboration with Biogen:

⁶² *Ibid.* Approximately 05:44:46 to 05:46:21.

⁶³ *Ibid.* Approximately 05:47:04 to 05:47:48.

⁶⁴ *Ibid.* Approximately 05:49:37 to 05:49:50.

⁶⁵ *Ibid.* Approximately 05:51:26 to 05:51:49.

⁶⁶ *Ibid.* Approximately 05:52:08 to 05:52:28.

Actually, Dr. Fountain [the chair of the advisory committee]... Can I just ask some of those folks who are commenting about how they feel about **what we thought was clear** on page 226 [of the briefing documents]. I mean **there's only so many pages we can write of the history of this, but it's a short sentence...** Well I could just read it: 'Upon initial review, the one positive study (Study 302) and the one negative study (Study 301) were given equal weight and consideration.' And I suspect...if you ask the applicant to weigh in, I think they will relate to you probably the degree to which [Study] 301 was given credence for a very long time, and it was quite clear that either study could represent, in the abstract, truth. And I'm just kind of curious about the comments, because we wouldn't want to have conveyed that, and I'm wondering if that was missed or if it wasn't understood in the way that we intended it. That's kind of what I'm getting at, and I wouldn't mind asking the applicant actually to weigh in on that aspect because I don't think there was any sense of the people [who] were working on this that it was entered into with a belief in [Study] 302 a priori and a desire to throw [Study] 301 away. I remember taking great pains to make sure that wasn't the case and maybe I could ask the applicant to weigh in on that. And also, if people could just clarify if we didn't communicate well our stance there. Maybe the applicant can...⁶⁷

Consistent with the months of collaboration that had taken place between the FDA and Biogen, Dr. Haeberlain from Biogen immediately chimed in after Dr. Dunn's tag-team prompt:

Yes, thank you. **That was absolutely the case through our investigations that we treated both studies equally**, and the resulting output of those investigations are indeed that Study 302 is robust and that Study 301 is a negative study. So that's not lending different weight to truth, but those outcomes are different. The nature of our investigations [was] to understand why study 301 was a negative study.⁶⁸

But the PCNS Drugs Advisory Committee members were having none of it, and several rebuked Dr. Dunn with the following statements:

Dr. Alexander:

But Dr. Dunn, if you review the briefing materials...but **the framing of the briefing materials [was] very much, at least as I interpreted them, as very much emphasizing an interest in identifying whether or not [Study] 301 could still provide sufficient evidence for [Study] 302 as a stand-alone pivotal study**. And the conclusion that was stated by the FDA used the words that the applicant has provided 'substantial evidence of effectiveness' and referred to [Study] 302 is a 'robust and exceptionally persuasive' study. And I believe what you've heard today and as well as what's been communicated through the vote is that the — well I don't want to speak or presume to speak on behalf of the entire committee — **but certainly I do not feel that the evidence has been presented to support that view from the FDA. So, throughout the briefing materials in many, many places the emphasis is not on using [Study] 302 to understand why [Study] 301 was negative and raising the question that perhaps [Study] 302 is really**

⁶⁷ *Ibid.* Approximately 05:52:38 to 05:54:24.

⁶⁸ *Ibid.* Approximately 05:54:24 to 05:54:51.

a negative study too. It's all framed in one direction, which is using [Study] 301 to support [Study] 302.⁶⁹

Dr. Perlmutter:

I would say...just to make a quick comment about impression of how the data...was presented to us. **Just go to first discussion point. The first discussion point seemed very biased in the sense that, okay, now consider [Study] 302 in light of and ignore everything in [Study] 301.** That just seems that we were being pushed in one direction or there was a bias in that one direction. So that really sums up kind of how I perceived the presentations.⁷⁰

Dr. Emerson:

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 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.⁷¹

Dr. Thambisetty:

I just wanted to note that the discordant ways in which we have perceived [question 7]. I think is also very aptly summarized in **the discordance between the FDA clinical reviewer and the FDA statistical reviewer,** and I think you know to paraphrase Dr. Tristan Massie, **if you have two and you take the best and pretend like it's the only one, your estimate is likely** biased. But I think that discordance is captured in the way the FDA's clinical review and statistical review differ as well.⁷²

Following this heated discussion, the advisory committee vote on question 8 — ***In light of the understanding provided by the exploratory analyses of Study 301 and Study 302, along with***

⁶⁹ *Ibid.* Approximately 05:54:57 to 05:56:29.

⁷⁰ *Ibid.* Approximately 05:56:32 to 05:57:09.

⁷¹ *Ibid.* Approximately 05:57:12 to 05:58:40.

⁷² *Ibid.* Approximately 05:59:17 to 05:59:47.

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,⁷³ 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.

Conclusions and requested actions

A decision by the FDA to approve aducanumab now would have several wide-ranging adverse consequences. First, approving a drug for Alzheimer's disease that has not been shown to be effective — and that in the end may turn out to be ineffective, assuming another pivotal phase 3 trial is conducted appropriately and completed — would provide false hope to millions of desperate patients with the disease and their families. Second, because the drug would be exorbitantly expensive (therapy would be priced at about \$50,000 per year,⁷⁴ and that does not include the cost of the serial brain imaging tests, such as magnetic resonance imaging, that patients would need to undergo) and used by potentially millions of patients for years, it would have a massive impact on health-care economics and potentially bankrupt the Medicare program, as well as many patients and their families. Such economic costs would only be justifiable for a drug that has definitive evidence of significant, clinically meaningful benefit. Finally, as alluded to by at least one member of the PCNS Advisory Committee,⁷⁵ the premature approval of aducanumab could impede the development of other experimental treatments for Alzheimer's disease for many years, potentially delaying progress on drugs that actually may turn out to be beneficial.

It seems likely that, but for the statistical review provided by Dr. Massie and the intervention of the FDA's PCNS Drugs Advisory Committee — whose members had not been subject to the apparent regulatory capture that compromised the independence and objectivity of the senior staff and clinical reviewers in CDER's Office of Neuroscience — the FDA was prepared to rush to the U.S. market a drug for Alzheimer's disease that lacks substantial evidence of effectiveness, despite these potentially catastrophic impacts.

As the HHS Principal Deputy Inspector General, you yourself must recognize the critical importance of ensuring that a regulatory agency like the FDA maintains its independence and objectivity when overseeing regulated industries. Breaches of the FDA's independence and objectivity undoubtedly could lead to agency approval of drugs and medical devices that are unsafe or ineffective, which could result in substantial harm to public health and to the private and public institutions and individuals who pay for health care.

In conducting an investigation of the FDA's review of aducanumab, we would encourage your staff to interview all FDA staff who were involved in the close collaboration between the agency

⁷³ *Ibid.* 06:04:38, see displayed slide.

⁷⁴ Belluck P. F.D.A. panel declines to endorse controversial Alzheimer's drug. *The New York Times*. November 6, 2020. <https://www.nytimes.com/2020/11/06/health/aducanumab-alzheimers-drug-fda-panel.html>. Accessed November 29, 2020.

⁷⁵ 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 30, 2020. Approximately 04:26:37 to 04:27:11.

and Biogen and in the review of Biogen's BLA for aducanumab, as well as Dr. Massie, whose critical review and analyses of the clinical trial data was not tainted by this FDA–Biogen collaboration. It is particularly important that the OIG explore the nature of “the ‘workstream’ or...‘working group’ collaboration” described by Biogen and confirmed by the FDA⁷⁶ that occurred during the post hoc analyses of the aducanumab clinical trial data. We also encourage you to examine whether there have been any similar close collaborations with the FDA and other pharmaceutical companies that likewise may have compromised the integrity of the agency's regulatory review and decision-making.

Public Citizen hopes that you share our concern regarding this troubling matter, and we look forward to an appropriate, favorable response to our urgent request. Please contact me at 202-588-7781 if you have any questions or need additional information.

Sincerely,



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

cc: Stephen M. Hahn, M.D., Commissioner of Food and Drugs, FDA
Patrizia Cavazzoni, M.D., Acting Director, CDER, FDA
Alex Azar, Secretary of Health and Human Services

⁷⁶ 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 November 30, 2020. PDF pages 19-20.

Appendix A

Table. List of Failed Experimental Drugs for Alzheimer’s Disease Targeting Amyloid- β (A β) Accumulation⁷⁷

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)
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)

⁷⁷ 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.

Appendix B

Additional Key Excerpts of Questions Raised and Comments Made by Members of the FDA's PCNS Drugs Advisory Committee During Its November 6, 2020, Meeting

Questions and comments following Biogen's summary presentation

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

Okay, so just doing a simple Bonferroni correction [for multiple comparisons], that P-value that you're quoting of .012 [for the primary outcome in Study 302], I don't know how to correct for the idea that you are looking at something different than the futility analysis dataset. I don't know how to correct for a lot of the other decisions that you might have considered, but certainly I can correct for you looking for the minimum of two P-values, in which case, ...well assuming that they're both independent, that .012 is not a true P-value. A true P-value would be closer to .0233, just adjusting for that aspect, no other multiplicity... I have looked at conditional upon deciding that we're [going to go] forward with this. The probability that the other study [i.e., Study 301] would have...a one-sided P-value... [of] .59 or higher — which I believe corresponds to your two-sided [P-value of] .833 — there's a 40% chance under the null that the...other independent study would have a P-value that large or larger, conditional on the fact that you've gone through and selected the results after you already knew them and decided what to present. Do you have an alternative calculation to this idea of this 40% chance...this idea that it's this discordant given the way that you sampled, which results you were going to present to us, that this P-value would be wrong?⁷⁸

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:

So, the incidence of ARIA-E [amyloid-related imaging abnormalities – edema] is 35% in the treatment group compared to 2.7% in the placebo group. I'd like to know how the diagnosis of ARIA is communicated to the patient and their caregivers. Are they told that they have brain swelling or microbleeds in the brain that requires them to come in for a previously unscheduled MRI scan? And if that is the case, are they also told that they would have to keep coming back for an MRI scan until these abnormalities improved and until such abnormalities improve that their dose of medication or placebo would have to be held? How is this information communicated to patients and caregivers, and do you have a sense for what their understanding is about the nature of ARIA and how it affects them scheduling previously unscheduled MRI visits?⁷⁹

⁷⁸ 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 30, 2020. Approximately 01:03:01 to 01:04:45.

⁷⁹ *Ibid.* Approximately 01:21:51 to 01:22:50.

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:

If I could interrupt one minute... I mean my question was about the placebo, the separation of the placebo curves. But I think if you're bringing up the consistency across multiple endpoints, it's also worth calling out another point raised by the FDA biostatistical review, which...correct me if I'm wrong here, but I believe that they indicated that because the low-dose primary endpoint was not met, technically the secondary high-dose endpoints can't be formally evaluated. And in fact, that the correlation between the primary and secondary endpoints was moderate, with correlation coefficients of...0.4 to 0.64, regardless of what principle components analysis may have suggested.⁸⁰

Questions and comments following the FDA's summary presentation

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

Let me ask about the 103 study... The amount of actual amyloid reduction in the 103 study was apparently much larger than in [Study] 302, and the difference between amyloid reduction at the top dose between [Studies] 301 and 302 appears to be really, really small. So it would be helpful to try to understand how with that pattern of data, one can view [Study] 103 as being supportive of [Study] 302 and how one can actually argue that a miniscule difference in the difference between [Studies] 301 and 302 can explain such a whopping difference in efficacy.⁸¹

Dr. Emerson:

I just want to make one comment, of course, is that the advisory committee is meant to see whether the FDA opinions are advisable. And so, of course, we're not just to be a rubber stamp for the FDA at all and that futility rules, in general, help public health immensely, although I will concede that the particular futility rule specified for this study was ill-advised, more because how it was so liberal in futility. But you remarked that the assumption of a common treatment effect was violated in the futility rule. If that's the case, how will you distinguish between the population that was in [Study] 302 and therefore has a treatment effect and the population in [Study] 301 that apparently does not have a treatment effect. I'll note that the futility analysis presumed that there would be differences in the estimated treatment effect and that is why, presumably, they chose to use the combined groups to try to get a better estimate. So, your statement that the treatment effect common between the two groups is violated argues that we should not write an indication that encompasses both study populations...⁸²

⁸⁰ *Ibid.* Approximately 01:28:00 to 01:28:53.

⁸¹ *Ibid.* Approximately 02:22:47 to 02:23:29.

⁸² *Ibid.* Approximately 02:30:17 to 02:31:37.

If you are saying that...the futility analysis was wrong because there was an assumption of a common treatment effect for both studies and that that was violated, that must mean that your belief that the treatment works in the [Study] 302 population but doesn't necessarily work in the [Study] 301 population must be somehow taken into account as you write an indication for this drug. How will you do that?...⁸³

So the FDA statistician, who we haven't heard from in this meeting but who did write a very nice report, might also bring to bear on this about what the distinction is between a treatment effect common between the two studies and similar estimates of treatment effect between the two studies and the difference between those. Which of those were assumed in the futility rule and which of those need to be assumed for issuing the general indication for all patients?⁸⁴

Comments during the discussion of question 1:

Question 1: The primary evidence of effectiveness presented in support of aducanumab for the treatment of Alzheimer's disease is provide by Study 302. Discuss the evidence of effectiveness provided by Study 302, viewed independently and without regard to Study 301, with particular consideration of the size of the study, design of the study, analysis of results to assess the effects of the drug, and consistency of results among various subgroups in the study.

Dr. Emerson:

They asked us to talk about [Study] 302 by itself, which I can do, but it has to take into account that this is the most exciting of results of two studies. And I need to make certain that the FDA is aware of that as they ask this question and, in part of this I guess, we could ask Dr. Dunn to tell us what he thinks the P-value is from the primary analysis in [Study] 302, and this will tell me a lot about...whether he's...thinking that [Study] 302 independent of [Study] 301 means pretend that 301 was never done or whether it means adjust the inference to allow for the fact that this is the best of two studies. So, Dr. Dunn if you wouldn't mind telling me what you think the P-value is for the primary endpoint. That would help a lot...⁸⁵

No, this is important, this is a very important question because if we are to pretend that [Study] 301 never existed, well we can talk about the scientific rigor and departures from that that such would be, but for instance, you could say that we adjusted for that statistic by saying that the P-value is not .012, but it's closer to .024, still ignoring some multiple comparison issues, but at least adjusting for the major aspect. In which case, I can answer this question all based on that...⁸⁶

If I can clarify here...it's possible that if you tell me you are analyzing the best of two studies, I cannot know anything about the other study except for the fact that it wasn't the best, and I can talk about what the results are, okay. And in that case the P-value of the

⁸³ *Ibid.* Approximately 02:31:47 to 02:32:17.

⁸⁴ *Ibid.* Approximately 02:32:55 to 02:33:28.

⁸⁵ *Ibid.* Approximately 03:37:50 to 03:38:47.

⁸⁶ *Ibid.* Approximately 03:39:21 to 03:39:56.

primary endpoint is something [like] .024 or higher. Okay. On the other hand, you could say pretend that Study 301 never existed, imagine a world in which [only Study 302] was there. Now I think that's a silly questions to ask given the history, but the idea of just wanting to stress that we can view independently the results of [Study] 302, recognizing that it's the best of two studies without ever knowing what the results are of Study 301 were...⁸⁷

But again..., I don't think it is the same. I mean, I think, yeah OK, I'm going to choose two numbers and only tell you what the highest number I chose was. That has a different sampling distribution than if I just choose one number and tell you what it is. So again,...I just want it clear — because I have a problem with this entire question — unless it's recognized that the P-value that was being reported throughout this thing of .012 is not a true P-value. But if you wanted us to do this,...it would be possible in a prespecified manner to tell the FDA I'm going to do two studies and I'm only going to give you the results of the best one, we can we can compute P-values. We can compute confidence intervals. We can do all sorts of things there.⁸⁸

Dr. Alexander:

I agree with [Dr. Emerson's] assessment, and I very much would like to get into some details here about the totality of evidence and about the conclusions that the FDA seems to be reaching. And about, as I said, the incongruous materials that have been provided and the dozens of questions that we really haven't had a chance to ask the FDA about. And in particular, I'd like to query the FDA about any number of concerns that their own statistical reviewer has identified and that I have not yet heard either an adequate response from the sponsor, but more importantly from the FDA regarding their own interpretation of those reviews...⁸⁹

So, I will speak only to [Study] 302 and resist the inclination to do otherwise. So, I think even with Study 302, there are some reasons for [questioning it]. One is that there's no correlation between plaque reduction and week 78 outcomes, and I think this is a good example where it feels a little bit like people want to have it both ways. In other words, there's an argument that molecular mechanisms provided strong support of the body of evidence to back up [Study] 302 is a robust exceptionally persuasive study. But there's also a disclaimer that no formal claim of biomarker is being made and no ability to explain why there's no correlation between plaque reduction and outcomes. A second source of concern about [Study] 302 is that the major stratum driving the findings, up to one-third of patients had a mid-study dose increase and more unblinding or potential unblinding. A third is that the placebo responses before and after amendment 4 are completely separate among at least some subgroups, suggesting that these dose increases are entangled with placebo worsening. In a sentence that as was pointed out by the FDA's own [statistical] reviewer, the failure of the low-dose arm in 302 means that technically the secondary endpoints for the high-dose arm are not interpretable, and even if they are, they're moderately correlated. And then the last thing that I'd say is that once again, as

⁸⁷ *Ibid.* Approximately 03:45:40 to 03:46:39.

⁸⁸ *Ibid.* Approximately 03:46:53 to 03:47:46.

⁸⁹ *Ibid.* Approximately 03:49:04 to 03:49:52.

pointed out by the FDA's own [statistical] reviewer, there's no consistent effect across subgroups in [Study] 302, yet one would hope to see this with a strong efficacy signal. So, these are five concerns about Study 302, even ignoring the fact that other studies have been performed.⁹⁰

Dr. Gold:

So, I think part of my concerns about [Study] 302, if I talked about it in isolation, is what happened pre- and post-amendment. In the actual numbers of subjects..., and again, I'm sorry, [Study] 301 comes into it because it's a question of ...who was being enrolled and what happened... So, I will stipulate... the studies were well-designed. I think that...there's no question with the design of the study. Part of the question I have is on the execution in the study. This is some of the materials that Biogen presented at the CTAD meeting a year ago where they actually showed that when they made the amendment, these amendments did not get implemented overnight. They took a long time, and in fact, there was a lot of heterogeneity in how the amendments got implemented... I just would like to get more clarity on exactly how the amendment really impacted [Study] 302 because if you think about it,... but there's evidence from looking at the ITT [intention-to-treat] analysis and for the post-amendment 4 population that there was an effect on the low-dose in the 302 study, which makes absolutely no sense to me, and again it's material Biogen presented at CTAD. So, if you start to see changes in the low-dose on the primary outcome measure from the ITT population versus the post...amendment fork, you have to wonder whether what you're seeing in the high-dose is noise... The Study [302] was declared futile...subjects were brought to closeout visits. There's a huge amount of missing information, which again has been referred to...by both...the sponsor and the FDA and the [FDA's] statistical reviewer. But I've heard no discussion about whether the pattern of missingness actually has any bearing here. So...it would be helpful to understand whether the analyses and the effects — really, do we actually believe that these data are missing at random, because that seems to be the assumption that was made in the analyses where the FDA reviewer was clearly saying...there are red flags here that these data are...missing not at random.⁹¹

Dr. Thambisetty:

My main concerns are with regards to the potential impact of unblinding of patients and caregivers. I think that's a huge concern with...the studies. 35% of patients exposed to the drug developed ARIA, and so it's inconceivable that patients and caregivers who are given a diagnosis of ARIA and who are then subjected to very intense serial MRI surveillance, which happens every month until the abnormalities are resolved, are going to be unblinded to the treatment and what makes this especially concerning is that the primary endpoint, which is the CDR-Sum of Box scores, is entirely dependent upon subjective information that is provided to the rater by patients and their caregivers. And the same goes for the secondary endpoints... These scales are very, very sensitive to biases due to expectations on the part of patients and caregivers when they realize that they are on the treatment arm, which is very likely to have occurred when you're being

⁹⁰ *Ibid.* Approximately 03:55:56 to 03:57:47.

⁹¹ *Ibid.* Approximately 03:58:26 to 04:00:58.

called in...for additional MRI scans because you have a drug-related adverse event. And I really think the fact that these potentially unbiased patients and caregivers are then providing subjective information about behavior and function that determined their scores on the primary endpoint as well as key secondary endpoints is a big concern that I don't think has been adequately addressed. The fact that the raters were blinded is really immaterial to this particular question because the information that the rater uses comes entirely from patients and caregivers for some of these scales...⁹²

I have one additional point about minimal clinically important difference, which I think is relevant in terms of the magnitude of the effects that are being reported. So I think the concept of minimally clinically important difference is very relevant to dementia clinical trials, and the fact that several of these outcomes are reported as relative differences in terms of percentage points in comparison to placebo make this slightly difficult to interpret because the strongest result is a relative difference of negative 0.39 points from placebo in the CDR-Sum of Box scores. This is also presented as a relative difference of 22% from placebo, but what do these changes mean in terms of their functional significance; do they to represent tangible real world benefits? — are they clinically important? — so this is what is captured by the concept of minimal clinically important difference. And that's defined as the smallest difference in score in the domain of interest which patients perceive as beneficial and which would mandate, in the absence of any troublesome side effects and cost, a change in patients' management, and there is empirically derived evidence for what constitutes minimal clinically important differences in dementia clinical trials. There's one paper that was just published by Andrews et al in *Alzheimer's and Dementia*, which suggests that...patients with MCI [mild cognitive impairment] is a change in CDR-Sum of Box scores or of MMSE [minimal status examination] of one point and for patients with mild Alzheimer's disease of two points. And if you use that as a yardstick these changes [in Study 302] are extremely small.⁹³

Vote on question 2

Question 2: Does Study 302, viewed independently and without regard to Study 301, provide strong evidence that supports the effectiveness of aducanumab for the treatment of Alzheimer's disease? Committee vote: 1 YES, 8 NO, 2 UNCERTAIN⁹⁴

Comments during the discussion of question 3

Question 3: The primary evidence of effectiveness presented in support of aducanumab for the treatment of Alzheimer's disease is provided by Study 302. Study 103 is presented as supportive evidence of aducanumab's effectiveness. Discuss the evidence of effectiveness provided by study 103.

⁹² *Ibid.* Approximately 04:12:11 to 04:14:05.

⁹³ *Ibid.* Approximately 04:14:17 to 04:16:06.

⁹⁴ *Ibid.* 04:24:26, see displayed slide.

Dr. Alexander:

So, I understand that [Study] 103 was not designed to allow for prespecified efficacy analysis... I was interested that the FDA's own biostatistical reviewer noted that the efficacy analyses that were performed lose statistical significance after excluding individuals who were initiating concomitant medications for treatment of Alzheimer's disease.⁹⁵

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:

And all that stuff makes it very hard to try to bolster something that...already... needs...real bolstering. And then, also, by the way, to skip over [Study] 301 because you know again the way that you try to bolster a study like [Study] 302 is by looking at another...well-designed similar study, but that other study, which is [Study] 301 — which again, we're not supposed to be talking about in this context — ...is a negative study. So...for me, I think that for those reasons, Study 103 provides...minimal support.⁹⁶

Dr. Thambisetty:

So, there are a couple of things that set Study 103 apart. So unlike Studies 301 and 302, the applicant actually has made data and results from Study 103 available for independent peer review, and these findings were published in 2016 in *Nature*, and I think it's really important to quote directly from the *Nature* paper about the appropriateness of using these data to make decisions about clinical efficacy. So, let me quote directly from the *Nature* paper: 'The trial was not powered for exploratory clinical endpoints. Thus, the clinical cognitive results should be interpreted with caution. Primary analyses were based on observed data with no imputation for missing values. Nominal P-values were presented with no adjustments for multiple comparisons.' So, I think it's worth remembering yet again that this was a safety and tolerability study.⁹⁷

Comments and vote on question 4

Question 4: Does Study 103 provide supportive evidence of the effectiveness of aducanumab for the treatment of Alzheimer's disease? Committee vote: 0 YES, 7 NO, 4 UNCERTAIN⁹⁸

Dr. Alexander:

The reasons why I would have concerns about using [Study] 103 to support [Study] 302 [include] that [Study]103 wasn't designed to allow for prespecified efficacy analysis, the efficacy lost statistical significance after excluding those with concomitant Alzheimer's

⁹⁵ *Ibid.* Approximately 04:34:13 to 04:34:40.

⁹⁶ *Ibid.* Approximately 04:46:00 to 04:46:38.

⁹⁷ *Ibid.* Approximately 04:57:04 to 04:58:00.

⁹⁸ *Ibid.* 05:10:57, see displayed slide.

medicines, the effect was 20 times larger, if I understood correctly than that of Study 302. I know that there were small sample sizes, but contrary to [Study] 302, the effect was larger in non-[ApoE ε4 (apolipoprotein E ε4)] carriers than carriers. Then the last two points that have been pointed out:... some highly sensitive measures did not reach statistical significance, and then finally, as was recently mentioned, I think by Dr Thambisetty, there was a lack of a strong dose-response relationship.⁹⁹

Dr. Emerson:

I voted no. Study 103, the positivity or any evidence it has is sort of a prerequisite for the other clinical trials. But just for added emphasis, in no way would I be accepting of regarding this as an adequate and well-controlled trial to make it two... We need a confirmatory study. And so again, [Study] 302 as a pivotal study or [Studies] 302 and 301 as two confirmatory studies are there, but [Study] 103 cannot take the place of another confirmatory study.¹⁰⁰

Comments during the discussion of question 5

Question 5: The application presents evidence in support of effects on the pathological hallmarks of Alzheimer's disease, including effects on amyloid beta, tau, and downstream markers of neurodegeneration using multiple assessment modalities. Discuss the impact of these results.

Dr. Thambisetty:

So, I think that from the results published from [Study] 103, as well as with [Studies] 301 and 302, there's clear evidence from brain amyloid PET imaging that aducanumab dose dependently clears amyloid plaque from the brain. So, I think that's pretty compelling. So, the drug appears to generate precisely the neuroimaging biomarker that you would expect by virtue of target engagement. I don't think there's any doubt about that in my mind. But in the larger context of the discussion today, particularly with relevance to the impact of aducanumab in slowing Alzheimer's disease progression, the question is whether lowering of brain amyloid burden as evidenced by PET imaging results in a clinical benefit. I think those are very distinct questions, but I think one follows the other very logically. And with regards to this question, I think the data are far less compelling, and I would point to slide 20 of the FDA statistical reviewer's presentation where you examine the relationship between change in global brain amyloid burden at week 78 in individuals exposed to high-dose aducanumab and change in the CDR-Sum of Box scores, there really appears to be no relationship either in Study 302 or [Study] 301. And this appears to be the case even when the analysis is restricted to only individuals exposed to the 10 mg/kg dose. I think there are some larger implications of these findings, which we're not tasked with discussing today. And so, one of the larger questions relevant to these observations are whether lowering brain amyloid burden is in fact the correct target in Alzheimer's disease.¹⁰¹

⁹⁹ *Ibid.* Approximately 05:13:11 to 05:14:06.

¹⁰⁰ *Ibid.* Approximately 05:15:49 to 05:16:25.

¹⁰¹ *Ibid.* Approximately 05:20:07 to 05:21:49.

Vote on question 6

Question 6: Has the applicant presented strong evidence of a pharmacodynamic effect on Alzheimer's disease pathophysiology? Committee vote: 5 YES, 0 NO, 6 UNCERTAIN¹⁰²

Comments during the discussion of question 7

Question 7: Study 301 was a negative study. Post hoc exploratory analyses were conducted in order to achieve maximum understanding of the partially discordant results of Study 301 and Study 302, and to determine if this understanding precludes independent consideration of Study 302. Additional contribution to the understanding of aducanumab's pharmacological activity and clinical effects is provided by the results of study 103. In light of the exploratory analyses that were conducted and the results of Study 103, discuss the impact of the results of Study 301 on the consideration of the results of study 302.

Dr. Emerson:

I'll just make three points... The linear dose-response — which much was made of that being one of Cook's postulates and things you want to do — none of this removed the fact that we did not have a linear dose-response in Study 301, with no explanation of why it was [not] there...¹⁰³

And then some of the other times are the missing data, I believe it was doctor Alexander who said earlier...that the missing data analyses were not very comprehensive to the possibility of missing not at random...So I was bothered by that.¹⁰⁴

Vote on question #8

Question 8: 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? Committee vote: 0 YES, 10 NO, 1 UNCERTAIN¹⁰⁵

¹⁰² *Ibid.* 05:36:10, see displayed slide.

¹⁰³ *Ibid.* Approximately 05:44:26 to 05:44:45.

¹⁰⁴ *Ibid.* Approximately 05:46:34 to 05:46:53.

¹⁰⁵ *Ibid.* 06:04:38, see displayed slide.