Received: 8 May 2021 | Revised: 10 January 2022 | Accepted: 11 January 2022

DOI: 10.1002/alz.12615

Alzheimer's & Dementia®
THE JOURNAL OF THE ALZHEIMER'S ASSOCIATION

PERSPECTIVE

A Bayesian perspective on Biogen's aducanumab trial

Anna G. M. Temp[1,2] | **Alexander Ly**[3,4] | **Johnny van Doorn**[3] | **Eric-Jan Wagenmakers**[3] | **Yi Tang**[5] | **Michael W. Lutz**[6] | **Stefan Teipel**[1,7]

[1] German Center for Neurodegenerative Diseases (DZNE), Rostock, Germany

[2] Translational Neurodegeneration Section "Albrecht Kossel," Department of Neurology, University of Rostock, Rostock, Germany

[3] Department of Psychological Methods, University of Amsterdam, Amsterdam, the Netherlands

[4] Machine Learning Group, Centrum Wiskunde & Informatica, Amsterdam, the Netherlands

[5] Department of Neurology, Xuanwu Hospital, National Center for Neurological Disorders, Capital Medical University, Beijing, China

[6] Department of Neurology, Duke University School of Medicine, Durham, North Carolina, USA

[7] Department of Psychosomatic Medicine and Psychotherapy, University Medicine Rostock, Rostock, Germany

Correspondence
Stefan Teipel, DZNE Rostock and University Medicine Rostock, Gehlsheimer Str. 20, 18147 Rostock, Germany.
E-mail: stefan.teipel@med.uni-rostock.de

Abstract

This perspective is a companion to a recent editorial on the use of Bayesian analysis in clinical research. We aim to introduce and highlight the relevance and advantages that Bayesian inference offers to clinical trials using the data on the amyloid antibody aducanumab presented at a Food and Drug Administration hearing in November 2020 as an applied example. We apply Bayesian analysis of model plausibility and effect sizes based on simulated data of the two phase 3 trials of aducanumab in prodromal and mild dementia stages of Alzheimer's disease (AD). Bayesian analysis can quantify evidence in favor of, or against, the presence of an effect (i.e., provide evidence of absence), as well as assess the strength of the effect. This is in contrast to the binary conclusions provided by frequentist tests.

KEYWORDS
aducanumab, Alzheimer's disease, Bayesian statistics, clinical trials

1 | INTRODUCTION

Therapeutic developments in Alzheimer's disease (AD) are in a critical phase. Biogen's presentation of its aducanumab trial results in October 2019 has provided new hope for patients and clinicians. However, interpreting these results is difficult, as shown by the debate after aducanumab's approval by the Food and Drug Administration in June 2021. Some of these problems are rooted in the conclusions facilitated by null hypothesis significance testing (NHST) as used in the analysis of the respective trials, and countless trials before. NHST is a statistical approach that facilitates decisions between two alternative conclusions: either a rejected null hypothesis at $P < .05$ (or any other chosen significance threshold), or a failure to reject the null hypothesis at $P \geq .05$. The *P*-value is continuous but commonly interpreted with a cut-off that provides its meaning ("statistically significant"). These all-or-nothing binary conclusions promote a false sense of assurance, as they ignore the fact that inference is based on noisy data. This type of uncertainty is acknowledged within a Bayesian framework. For testing problems, Bayes factors are used to provide a continuous measure of evidence in favor of one hypothesis over another; more evidence leads to more certain test conclusions and likewise, less evidence leads to more uncertain conclusions about the hypotheses. Similarly, for estimation problems the inference is based not only on a point estimate, but on the posterior distribution over the parameter of interest: a more peaked posterior allows for more certain inferences about the magnitude of an effect, whereas a wide posterior conveys more uncertainty if the effect exists (examples are demonstrated in Figure 1A and 1B). Using a posterior distribution means that we can quantify how likely it is that the parameter of interest falls within a certain interval (so-called credible intervals, CI).[1] By comparison, frequentist confidence intervals

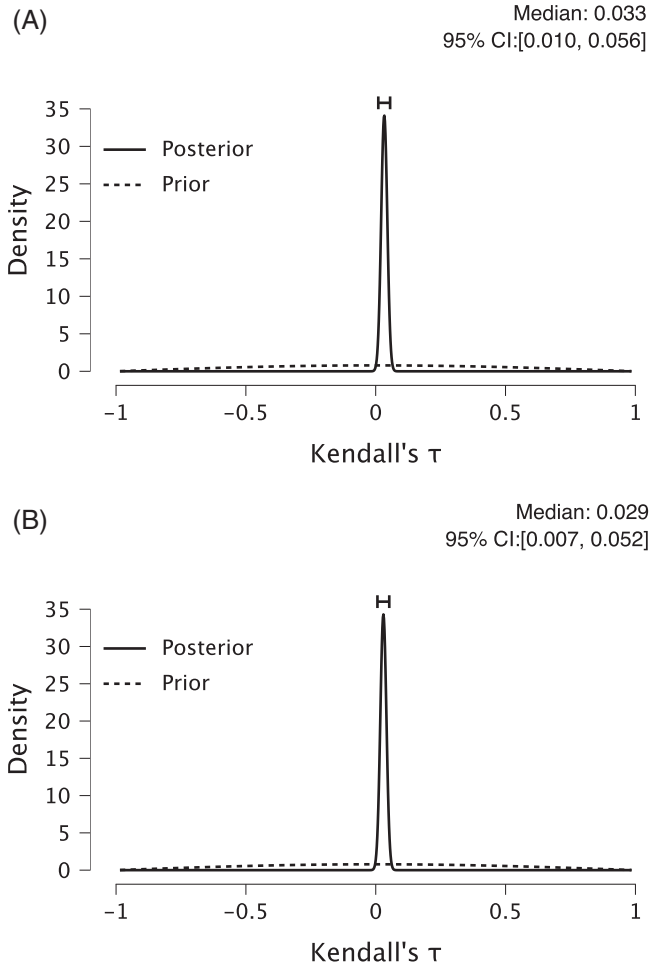

FIGURE 1 A, The prior and posterior plot shows that given that the effect of aducanumab exists, it is likely to be small. Effect size estimate based on data without the rapid progressors. Figure from JASP. B, The prior and posterior plot shows that given that the effect of aducanumab exists, it is likely to be small. Effect size estimate based on data with the rapid progressors. Figure from JASP. CI, credible interval; JASP, Jeffreys's Amazing Statistics Program

will—at best—contain the true value in a fixed proportion of repeated samples.[2]

Issues surrounding NHST are manifold and affect trial results in three crucial areas: first, theoretical limitations of this approach to probability as highlighted by the editorial and discussed at length elsewhere;[1,3–8] second, the limited and perhaps overconfident conclusions, which do not acknowledge uncertainty, as discussed above; and third, frequent misinterpretation of these conclusions. Concerns about the prevalence of such misinterpretations have been raised by statisticians since at least the 1980s.[4,9–12]

The current perspective aims to draw readers' attention to another statistical framework: the more intuitive Bayesian understanding of probability that is being advocated for in the editorial. We believe it is time to promote Bayesian inference techniques to a wider audience within the community of AD researchers, as they are still somewhat niche: of the current literature under PubMed's MeSH (Medical Subject Headings) term "Alzheimer disease," approximately 0.24% apply Bayesian modeling techniques.

We simulated Biogen's aducanumab data and re-analyzed these data using Bayesian techniques. We simulated our data as the original data are not available for public use.

We used Bayesian equivalents of the frequentist analyses performed by Biogen to reveal how a Bayesian perspective can be more informative. First, we used Bayes factors to quantify the evidence supporting the hypothesis of presence over the hypothesis of absence of aducanumab's treatment effect, on the primary cognitive outcome as measured by the Clinical Dementia Rating Sum of Boxes (CDR-SB[13]). Hence, Bayes factors were used to infer whether the effect exists. Second, provided that the effect exists, posterior distributions were used to (1) estimate the size of the effect, and (2) provide a measure of uncertainty about this estimate. We believe that Bayes factors and posteriors can provide informative statements that are useful for patients, physicians, and the research community as a whole.

2 | MATERIAL AND METHODS

2.1 | Data generation

The presented data were simulated based on summary statistics of Biogen's aducanumab trial,[14] available here: https://www.fda.gov/media/143502/download. Data used include study membership (301, "ENGAGE" [NCT02477800] vs. 302, "EMERGE" [NCT02484547]) and treatment group membership (placebo vs. low-dose aducanumab vs. high-dose aducanumab). The outcome we simulated was mean change in CDR-SB. If aducanumab was effective, then we would expect a numerically smaller mean change in CDR-SB in both treatment groups than in the placebo group in our data. A smaller mean change in CDR-SB would indicate that participants treated with aducanumab decline more slowly than participants in the placebo group. We simulated the data based on the means of Biogen's original sample, reported in Tables 35 and 36 on pp. 229—230 of their report.

Biogen claimed that the presence of rapid progressors may have distorted their conclusions. Lacking a "precedent in clinical trials for defining rapid progressors" (p. 69[14]) in their data analysis "[a] cut-off of change > 8 points on CDR-SB was chosen as the primary definition of rapid progression. This is four-fold the upper limit of the anticipated mean decline in this study population and, therefore, is an extreme change compared to typical progression" (p. 69[14]). The summary statistics reported in Table 35 and 36 in the briefing are based on the data with the rapid progressors filtered out. To adequately simulate the number and CDR-SB of the rapid progressors, the histogram on p. 228 of the briefing was consulted. The original report did not contain standard deviations; we derived those from the standard errors in Table 9 (p. 177[14]) for study 302, and in Table 15 (p. 190[14]) for study 301. The R code used to simulate these data can be found in the supporting information, "Biogen Sim Code.R". The simulation to produce our data was run once.

TABLE 1 Descriptive statistics of the simulated change in CDR-SB over 78 weeks, excluding rapid progressors

						95% CI	
Study number	Treatment group	Biogen's mean	Mean	SD	N	Lower	Upper
Study 301	High aducanumab	1.38	1.380	2.020	546	1.210	1.550
	Low aducanumab	1.28	1.280	1.990	542	1.112	1.448
	Placebo	1.47	1.470	1.990	541	1.302	1.638
Study 302	High aducanumab	1.24	1.240	1.940	542	1.076	1.404
	Low aducanumab	1.37	1.370	1.980	539	1.202	1.538
	Placebo	1.66	1.660	2.030	544	1.489	1.831

Abbreviations: CDR-SB, Clinical Dementia Rating Sum of Boxes; CI, credible interval; SD, standard deviation.

TABLE 2 Descriptive statistics of the simulated change in CDR-SB over 78 weeks, including rapid progressors

					95% CI	
Study number	Treatment group	Mean	SD	N	Lower	Upper
Study 301	High aducanumab	1.516	2.273	555	1.327	1.706
	Low aducanumab	1.359	2.149	547	1.179	1.540
	Placebo	1.528	2.097	545	1.352	1.705
Study 302	High aducanumab	1.319	2.100	547	1.143	1.495
	Low aducanumab	1.433	2.106	543	1.255	1.610
	Placebo	1.725	2.161	548	1.544	1.906

Abbreviations: CDR-SB, Clinical Dementia Rating Sum of Boxes; CI, credible interval; SD, standard deviation.

2.1.1 | Variables

Variables in the data set include "studynumber," which distinguishes between study 301 and 302; "group," which distinguishes the treatment groups of high-dose aducanumab, low-dose aducanumab, and placebo; and "progressionspeed," which distinguishes between rapid progressors and non-rapid progressors as dichotomized by Biogen. The latter variable is used to exclude, that is, filter out, rapid progressors from the analyses where necessary. The outcome variable represents the change in CDR-SB. The simulation code, resultant data set, and Jeffreys's Amazing Statistics Program (JASP)-derived HTML results files can be found on the Open Science Framework: https://osf.io/scwaj/

2.1.2 | Descriptive statistics

The simulated means and standard deviation of CDR-SB change excluding the rapid progressors can be found in Table 1.

The descriptive statistics in Table 1 correspond to those reported for Biogen's original non-rapid progressors (Tables 35-36, pp. 229-230[14]). When including the rapid progressors, we obtained the descriptive statistics of CDR-SB change shown in Table 2.

2.2 | Statistical analysis

Analyses based on these simulated data were conducted in JASP.15 For the simulation and the subsequent analyses, the seed was set to 8735. We opted for JASP's default JZS priors on the parameters, and the a priori assumption that each hypothesis/model under investigation is equally likely. Instead of *P*-values, the analysis provides Bayes factors. When denoted as BF_{10}, the Bayes factor quantifies the evidence in the observed data in favor of model "1" against model "0." For instance, $BF_{10} = 7$ implies that the data are 7 times more likely under model "1" compared to model "0." Analogously, $BF_{10} = 1/3$, which is equivalent to $BF_{01} = 1/BF_{10} = 3$, implies that the data are three times more likely under model "0" compared to model "1." Bayes factors are best interpreted as continuous measures of evidence, but the following qualitative descriptions of strength of evidence can help with the communication and interpretation of BF_{10}: 1–3 for anecdotal evidence, 3–10 for moderate evidence, 10–30 for strong evidence, 30–100 for very strong evidence, and $BF_{10} > 100$ for extreme evidence favoring model "1" over model "0"; for example, the alternative hypothesis against the null hypothesis. Conversely, 0.33–1.00, 0.10–0.33, 0.03–0.10, 0.01–0.03, and $BF_{01} < 0.01$ indicate evidence for model "0" over model "1"; for example, the null against the alternative hypothesis. $BF_{10} = 1$ indicates that there is no evidence favoring either of the competing hypotheses.[16]

2.2.1 | Analytical strategy

First, we conducted frequentist analyses of variance, in analogue with Biogen's provided strategy: testing the relevance of the treatment grouping factor (high dose, low dose, and placebo treatment), separately for each study. These confirmed that our simulated data

yield conclusions resembling the ones based on the original data. For the Bayesian perspective we replicated these analyses of variance (ANOVAs) in the Bayesian framework to quantify the amount of support for the hypothesis that aducanumab's treatment effect is present over its absence indicated by the simulated data. To investigate claims of Biogen's post hoc analysis, we ran another Bayesian ANOVA to investigate whether for rapid progressors inference about aducanumab's treatment was different across both studies and all treatment arms. Last, provided aducanumab's treatment effect on CDR-SB exists, we conducted Kendall's tau correlation analysis to estimate the size of the treatment effect. Kendall's tau is robust to outliers and violations of normality, making it suitable for our simulated data.[17,18]

2.3 | Quantifying the evidence in favor of the presence over the absence of an effect of aducanumab on CDR-SB

2.3.1 | Replicating Biogen's original (frequentist) results

Biogen found no treatment effect in study 301, only in study 302, even when excluding rapid progressors ($P = .012$ in the high-dose group, p. 41). Here, their null hypothesis was that there was no effect of aducanumab on CDR-SB in either study. A frequentist re-analysis of our simulated data confirmed that the null hypothesis may be rejected for study 302, including rapid progressors ($F_{(2,1635)} = 5.33$, $P = .005$, $\eta^2 = 0.006$). In study 301, the simulated data failed to reject the null hypothesis when including rapid progressors ($F_{(2,1644)} = 1.03$, $P = .357$, $\eta^2 = 0.001$) and when excluding them ($F_{(2,1626)} = 1.22$, $P = .295$, $\eta^2 = 0.002$). This shows that our simulated data behaved comparably to the original data, providing the same conclusions: the null hypothesis of no effect would be rejected in study 302, but not in study 301. Notably, $P = .295$ indicates the failure to reject the null hypothesis, which is not the same as supporting the null hypothesis. From the post hoc observation of rapid progressors, Biogen generated the hypothesis that the disproportionate amount of rapid progressors in the high aducanumab treatment arm of study 301 led to this study to not reject the null hypothesis of aducanumab being ineffective.

2.3.2 | How strong is the evidence in favor of the presence over the absence of the effect of aducanumab on CDR-SB?

Following Biogen's original pre-registered analysis plan, we performed a Bayesian ANOVA on the full data set of each individual study. The Bayesian ANOVA compared the null model, representing the hypothesis that aducanumab has no treatment effect, to the "group" model, representing the hypothesis that there is a difference between treatment groups. We repeated the analysis for study 301 and study 302 separately, and for each study with and without rapid progressors. By focusing on the Bayes factor scores we can quantify the additional evidence for the presence over the absence of the aducanumab on CDR-SB, when the rapid progressors were excluded from the data.

For the ANOVA-based model comparisons, we report: the "prior model probabilities" (P[M]), which represent the prior belief of each hypothesis/model before data observation; the "posterior model probabilities" (P[M|data]), which represent the belief of each hypothesis after data observation; the Bayes factors BF_{10}, which represent the evidence in favor of a generic model "1" over a base model "0" brought about by the simulated data; and the error% as an indicator of the numerical methods used to compute the Bayes factors. A lower error% indicated lower fluctuations of the BF over the Markov chain Monte Carlo (MCMC) sampling; an error% below 20 can be considered acceptable.[15]

As before, we followed Biogen's original analysis plan by first conducting the ANOVA with the rapid progressor data included, and then repeating the ANOVA with the rapid progressors filtered out. The "progressionspeed" variable was not used as a predictor in these analyses.

With rapid progressors included in study 302, our Bayesian re-analysis yielded inconclusive results, P(M) = 0.50, P(M|data) = 0.568, $BF_{10} = 1.32$. This can be interpreted as an absence of evidence for or against aducanumab's treatment effect in study 302 ($BF_{10} = 1.32$). A Bayesian re-analysis of the data of study 302 without rapid progressors showed moderate evidence in favor of the treatment effect hypothesis (i.e., model 1) compared to the absence (i.e., model 0) of the effect hypothesis, P(M) = 0.50, P(M|data) = 0.787, $BF_{10} = 3.69$, error% = 0.02. This means that for the non-rapid progressors of study 302 we had an a priori belief of 50% that the treatment is effective, which after data observation was increased to 78.7%. The change from prior to posterior model probabilities brought about by the data can be computed using the Bayes factor. In particular, the Bayes factor indicates that the (simulated) data of study 302's non-rapid progressors was almost four times more likely under the hypothesis that aducanumab does have a treatment effect compared to not having an effect. Applying the aforementioned cut-offs, this can be described as moderate evidence favoring aducanumab's treatment effect over its absence ($BF_{10} \geq 3$). The evidence cut-offs should be interpreted with care. Focusing on the raw Bayes factor scores we see that the post hoc criterion of removing the rapid progressors from the data set led to only a small increase in evidence for the presence over the absence of aducanumab on CDR-SB.

A non-significant result $P \geq .05$ implies absence of evidence for the effect, but it does not imply evidence of absence of the effect.[4,19–21] On the other hand, Bayes factors can be used to quantify evidence for the absence over the presence of an effect. Following Biogen's approach directly by retaining the rapid progressors in the ANOVA for study 301 yielded very strong support for the null hypothesis, which explained our data 51 times better than the treatment effect hypothesis, P(M) = 0.50, P(M|data) = 0.019, $BF_{10} = 0.02$, equivalently, $BF_{01} = 51.06$, error% = 0.03. Hence, our prior belief in aducanumab's effect on CDR-SB has been reduced: from 50% to 1.9%. Excluding the rapid progressors from study 301, our Bayesian re-analysis provided still very strong support for the null model compared to the treatment effects P(M) = 0.50, P(M|data) = 0.023, $BF_{10} = 0.024$, equivalently,

$BF_{01} = 1/BF_{10} = 41.87$, error% = 0.03. This shows that our simulated data of trial 301 were nearly 42 times more likely to occur under the null hypothesis that aducanumab does not have a treatment effect on non-rapid progressors compared to alternative model ($BF_{01} = 41.87$). Applying the cut-offs, this can be interpreted as very strong evidence against aducanumab's treatment effect on study 301's non-rapid progressors ($BF_{01} \geq 30$). Hence, excluding the rapid progressors from the analysis led to a decrease in evidence for the absence of an effect of aducanumab on CDR-SB. Regardless of including or excluding the rapid progressors, however, the evidence for the null over the alternative remained very strong in study 301.

These results provide more informative conclusions put forth by Biogen's original analyses: when including rapid progressors, we found inconclusive evidence for aducanumab's treatment effect in study 302 ($BF_{10} = 1.32$) and strong evidence for its absence over presence in study 301 ($BF_{01} = 51.06$). After excluding rapid progressors, study 302 provided some evidence favoring aducanumab's treatment effect ($BF_{10} = 3.69$) but study 301 still provided strong evidence for the absence of the treatment effect ($BF_{01} = 41.87$). This differs from the binary conclusions in the frequentist framework: Study 302 led to the rejection of the null hypothesis, whereas the Bayes factor quantifies the support for the presence over the absence of the treatment effect. At the same time, the non-significant *P*-value in study 301 cannot be interpreted as evidence for the null,[4,19–21] whereas the Bayesian analysis allows quantifying the evidence in favor of the null model.

2.3.3 | Is there evidence that rapid progressors distorted the treatment effect?

We performed a Bayesian ANOVA to study Biogen's justification of the post hoc exclusion of rapid progressors in the data set. Table 3 contains a Bayesian ANOVA with three independent variables: study type (study 301 vs. study 302), group (placebo vs. low dose vs. high dose) and progression speed (non-rapid vs. rapid progressors). Note that explicit testing of the main effect of rapid progressors on rates of progression reveals some collinearity as rapid progressors were actually defined according to their rate of progression—which was the case in our simulated data. The Bayesian ANOVA results are based on 10,000 MCMC samples and simultaneously address three questions: (1) How strong is the evidence for the factor "group"; that is, for aducanumab's treatment effect (or its absence) across the two studies 301 and 302? (2) How relevant is the study indicator; that is, is "studynumber" a good predictor for the outcome? (3) Is study 301's high dosage arm disproportionately affected by rapid progressors, as Biogen had hypothesized (pp. 69–72[14]); that is, is there a three-way interaction among studynumber, group, and progression speed? To answer these questions, we compared several statistical models (Table 3).

With the main effect of progression speed, and a possible three-way interaction there are a total of 19 models, which are reported in Table 3. Uniform prior model probabilities imply that each model gets a prior model probability of 1/19 = 0.053 before data observation. The following models were most interesting to us:

1. The "null model," representing the null hypothesis that the main effects of treatment, the studynumber, and progressionspeed were absent;
2. The treatment main effect-only model, that is, "group" only model, that placebo, low-dose, and high-dose groups showed differences in CDR-SB progression;
3. The study number main effect-only model that study 301 and 302 provided conflicting evidence, by producing differences in CDR-SB progression;
4. The interaction effect model that aducanumab's effect on CDR-SB progression affected treatment groups distinctly, depending on the study and number of rapid progressors ("studynumber*group*progressionspeed").

The results are in Table 3, with models 1–3 highlighted in bold. Here, the null hypothesis has been fixed to the top of the table, and the Bayes factors BF_{10} are reported in favor of models listed in the left column. Hence, the larger BF_{10}, the more evidence there is for the alternative model specified in the left column.

Table 3 shows that the treatment main effect ("group"-only model) explained our data much more poorly than the null model (i.e., $BF_{10} = 0.161$, conversely, $BF_{01} = 6.21$). This signifies moderate evidence that the treatment effect is absent. Furthermore, our simulated data were 167 times more likely to occur under the null model compared to the "studynumber" main effect only model (i.e., $BF_{10} = 0.006$, conversely, $BF_{01} = 167$) indicating extremely strong evidence that the studies do not differ in the absence of evidence for aducanumab's treatment effect. This is contrary to Biogen's proposal that study 301 and study 302 provide conflicting evidence with 302 providing evidence favoring aducanumab's treatment effect, and 301 providing evidence supporting its absence. Finally, the three-way interaction model of "studynumber*group*progressionspeed" appeared congruent with our data, as our data were much more likely under the three-way interaction than under the null model ($BF_{10} = 1.457e+105$). However, observe that model (4) featuring only the progressionspeed main-effect was $BF_{10} = 1.187e+110$ times more likely than under the null. This large amount of evidence reflects that progressionspeed was defined by the outcome. Furthermore, the large evidence for model (4) suggests that the large evidence for the full three-way interaction model (3) was mainly driven by the inclusion of progressionspeed.

In the supporting information we present an alternative approach to test model fit in the Bayesian framework that is independent of the use of Bayes factor. Congruent with model (4) of Table 3, the comparison of model fit in the supporting information shows that the retrospectively defined rapid progressor variable indeed improves the posterior fit of the outcome.

To infer the overall importance of each factor (main effects, two-way, and three-way interactions) across all 19 models in Table 3, we reported the "analysis of effects" table; see Table 4. This table provides the Bayes factor of inclusion or exclusion of each factor model-averaged across all 19 models shown in Table 3. The larger the inclusion Bayes factor, the more likely that factor predicts the outcome variable CDR-SB. Table 4 also contains the prior inclusion and exclusion

TABLE 3 Model comparison investigating the distortion effect of the rapid progressors on CDR-SB across the high, low, and placebo treatment groups in studies 301 and 302

Model content	P(M)	P(M|data)	BF_M	BF_{10}	Error%
(0) Null model	**0.053**	**3.848e-111**	**6.927e-110**	**1.000**	
(4) progressionspeed	0.053	0.457	15.136	1.187e+110	9.122e-116
(5) group + progressionspeed	0.053	0.420	13.061	1.093e+110	1.323
(6) group + progressionspeed + group * progressionspeed	0.053	0.060	1.158	1.571e+109	18.593
(7) studynumber + group + progressionspeed	0.053	0.023	0.425	5.999e+108	11.392
(8) studynumber + progressionspeed	0.053	0.022	0.409	5.770e+108	1.552
(9) studynumber + progressionspeed + studynumber * progressionspeed	0.053	0.007	0.126	1.801e+108	18.646
(10) studynumber + group + progressionspeed + studynumber * progressionspeed	0.053	0.005	0.098	1.402e+108	3.144
(11) studynumber + group + progressionspeed + group * progressionspeed	0.053	0.002	0.045	6.463e+107	2.601
(12) studynumber + group + progressionspeed + studynumber * group	0.053	0.001	0.019	2.794e+107	2.776
(13) studynumber + group + progressionspeed + studynumber * progressionspeed + group * progressionspeed	0.053	6.475e −4	0.012	1.682e+107	4.095
(14) studynumber + group + progressionspeed + studynumber * group + studynumber * progressionspeed	0.053	2.804e −4	0.005	7.286e+106	3.754
(15) studynumber + group + progressionspeed + studynumber * group + group * progressionspeed	0.053	1.289e −4	0.002	3.349e+106	3.938
(16) studynumber + group + progressionspeed + studynumber * group + studynumber*progressionspeed + group * progressionspeed	0.053	3.521e −5	6.337e −4	9.148e+105	6.174
(3) studynumber + group + progressionspeed + studynumber * group + studynumber * progressionspeed + group * progressionspeed + studynumber * group * progressionspeed	**0.053**	**5.606e −6**	**1.009e −4**	**1.457e+105**	**2.545**
(1) group	**0.053**	**6.214e-112**	**1.119e-110**	**0.161**	**0.029**
(2) studynumber	**0.053**	**1.593e-112**	**2.867e-111**	**0.041**	**3.260e −4**
(17) studynumber + group	0.053	2.489e-113	4.479e-112	0.006	1.431
(18) studynumber + group + studynumber * group	0.053	2.092e-114	3.765e-113	5.435e -4	11.594

Abbreviations: BF_{10}, Bayes factor in favor of the model in this row against the null model on top; BF_M, degree of change from prior to posterior model odds; CDR-SB, Clinical Dementia Rating Sum of Boxes; error%, numerical stability of the result; P(M), assumed prior model probability; P(M|data), posterior model probability after having analyzed the simulated data.

probabilities, which sum to one, and the posterior probability of inclusion and exclusion (P[incl|data] and P[excl|data]), respectively, which also sum to one. Because the main effects also occur in the interaction models, they occur more often in Table 3 and therefore have a higher prior inclusion probability. For instance, the group factor occurs in 14 out of the 19 models reported in Table 3 and the prior inclusion probability is therefore 14/19 = 0.787, and the prior exclusion probability is 1-0.787 = 0.213.

Table 4 suggests that our simulated data support the inclusion of the retrospectively defined progressionspeed predictor ($BF_{incl} = \infty$). This is not surprising, due to this factor being defined by the outcome itself. All other effects on CDR-SB were recommended for exclusion: there was anecdotal evidence against including treatment group ($BF_{incl} = 0.378$, conversely $BF_{excl} = 2.647$), moderate evidence against the interaction between group and progression speed ($BF_{excl} = 6.778$), very strong evidence against studynumber ($BF_{excl} = 42.162$) and its interaction with progression speed ($BF_{excl} = 34.256$), and extremely strong evidence against studynumber's interaction with group ($BF_{excl} = 302.116$), and even more evidence against the inclusion of the three-way interaction ($BF_{excl} = 9910.738$). Note that the raw Bayes factor values here provide a more informative picture of the evidence than the labels "moderate," "strong," and so on.

Taken together, the analyses in Tables 3 and 4 inform us that the presence of between-study effects and disproportional distortions by

TABLE 4 Analysis of effects on CDR-SB mean change over 78 weeks

Effect	P(incl)	P(excl)	P(incl|data)	P(excl|data)	BF_{incl}	BF_{excl}
Studynumber	0.737	0.263	0.062	0.938	0.024	42.162
Group	0.737	0.263	0.514	0.486	0.378	2.647
Progressionspeed	0.737	0.263	1.000	0.000	∞	0.000
Studynumber * group	0.316	0.684	0.002	0.998	0.003	302.116
Studynumber * progressionspeed	0.316	0.684	0.013	0.987	0.029	34.256
Group * progressionspeed	0.316	0.684	0.064	0.936	0.148	6.778
Studynumber * group * progressionspeed	0.053	0.947	5.606e -6	1.000	1.009e –4	9910.738

Abbreviations: BF_{excl}, exclusion Bayes factor; BF_{incl}, inclusion Bayes factor; P(excl), prior exclusion probability of this effect; P(incl), prior inclusion probability; P(excl|data), posterior exclusion probability; P(incl|data), posterior inclusion probability.

rapid progressors in the high aducanumab group are highly unlikely compared to their absence. The rapid progressor effect was defined by the outcome retrospectively. Table 4 tells us that only the rapid progression speed indicator is relevant for predicting CDR-SB; there is little evidence to include the group and study indicators, or higher order interactions. Each effect is recommended for exclusion, suggesting that they are poor predictors of CDR-SB. The presence of an effect of a rapid progression speed indicator in the post hoc analysis generates but does not confirm the hypothesis that rapid progressors may have influenced the outcome of the aducanumab trials. Confirmation of this newly generated hypothesis requires prospective testing in an independent study.

The low dose group may be considered a "fall-back" dose in case the high dose is not safe enough to administer. We re-ran our analyses using data from both study 301 and 302 combined in a one-way Bayesian ANOVA, but without the low-dose treatment arm. Following this approach, the posterior probability of the null model increased from 0.50 to 0.62, with inconclusive evidence regarding the high-dose treatment arm (P(M) = 0.50, P(M|data) = 0.38, BF_{01} = 1.65, error% = 2.105e-6). This implies that our data were 1.6 times more likely under the hypothesis that aducanumab's treatment effect is absent than under the hypothesis that it is present in the high treatment group alone. Without rapid progressors, the simulated high-dose data were four times more likely under the treatment effect compared to the null hypothesis (P(M) = 0.50, P(M|data) = 0.796, BF_{10} = 3.91, error% = 3.184e-7).

2.3.4 | If there is a treatment effect, how strong is it?

To estimate the magnitude of the effect, we inferred it using Kendall's tau and assumed that it was not zero. We calculated Kendall's tau correlation coefficient with dummy-coded group membership (3 = placebo, 2 = low dose, 1 = high dose). This coding means that we would expect a positive correlation: the higher the numerical group membership, the numerically greater the change in CDR-SB. This would mean that the greatest change in CDR-SB would occur in the placebo group and the smallest change in the high aducanumab group. To estimate the magnitude of the effect in the population based on the (simulated) data, we report the so-called posterior median for Kendall's tau and the 95% credible intervals (CIs) as a measure of uncertainty regarding this estimate. Data from studies 301 and 302 were combined for this purpose.

Including the rapid progressors, the two-sided tau = 0.029, 95% CI (0.007, 0.052) was very small. Excluding the rapid progressors as per Biogen's argument, the median tau was slightly larger at 0.033 with a two-sided 95% CI of 0.010 to 0.056—very small effects. This informs us that if aducanumab's treatment effect exists, we can be 95% confident that its true size lies between 0.010 and 0.056. This means that if the effect exists, we would be quite assured that it would be very small. Figure 1 plots the prior and posterior distributions without and with the rapid progressors (A and B, respectively).

These so-called "prior and posterior plots" (Figure 1A and 1B) provide us with the following insights. The dashed line represents the density of the size of the effect if it exists prior to observing data (see van Doorn et al.[17]). Note that the prior is quite spread out conveying a relatively large amount of uncertainty regarding the size of the effect. This prior was updated to a posterior distribution for Kendall's tau (solid line) and quantifies our uncertainty about the magnitude of tau in the population *after* observing the (simulated) data. In Figure 1A and 1B, the posterior distributions are relatively peaked compared to the prior distribution, indicating that the data substantially decreased our prior uncertainty about the size of Kendall's tau in the population.

Figure 2 paints the same picture: The size of CDR-SB varied similarly between all three groups.

3 | SUMMARY OF THE RESULTS

Of note, the subsequent statements are based on simulated data. They are not based on the analysis of the trials' original data, so conclusive inference on the effects in the trials themselves cannot be drawn. Rather, our analyses illustrate which kind of conclusions a Bayesian approach would offer for analyzing the trials' data, which frequentist analysis cannot provide. When following Biogen's approach by analyzing the simulated data of study 302 separately, it provided some evidence favoring aducanumab's effect on CDR-SB; in study 301, there

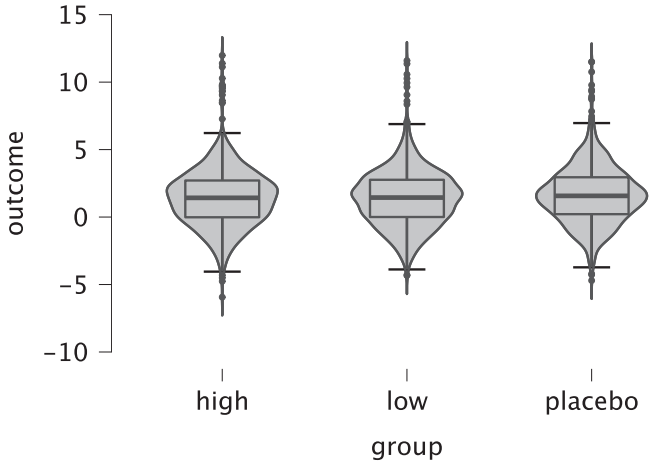

FIGURE 2 Distributions and densities of aducanumab's treatment effect across groups (rapid progressors included)

was stronger evidence against a treatment effect. Excluding the rapid progressors from our simulated data did not produce decisive evidence supporting aducanumab's treatment effect. When investigating whether study indicator (study 301 or study 302) provided conflicting evidence of aducanumab's treatment effect on CDR-SB change, our simulated data provided strong evidence for irrelevance of "studynumber" as a predictor (i.e., $BF_{exc} = 42.16$, Table 4): CDR-SB appeared homogeneous across study 301 and study 302. Consequently, these simulated data do not support the dismissal of study 301 selectively, as proposed by Biogen based on their original data. Our simulated data further offered overwhelming evidence against Biogen's assumption that rapid progressors affected any treatment arm of either study disproportionately (i.e., $BF_{exc} = 9910.74$, Table 4). This indicates that even though the three-way interaction model is much better than the null model in Table 3, it is still a poor-performing model whose effect is purely driven by the retrospectively defined rapid progressors. We also found anecdotal evidence against the treatment arms (i.e., $BF_{exc} = 2.65$, Table 4) as a predictor of CDR-SB change. The Bayesian ANOVA both based on Bayes factor and using model fit as an alternative approach in the supplement showed that incorporating the predictor of "progressionspeed" introduced substantial bias into any prediction of CDR-SB change because it is the driving force: There is only evidence for rapid progression indicator as a predictor, whereas there is evidence to exclude group, study assignment, and the higher order interactions. Without rapid progressors, there is only modest support for aducanumab's effect in study 302 ($BF_{10} = 3.69$), and very strong evidence against it in study 301 ($BF_{10} = 41.87$). The prior and posterior plots show that if the effect existed, it would be very small, likely ranging from Kendall's tau 0.007 to 0.052, with a median of 0.029.

4 | DISCUSSION

Bayesian inference offers demonstrable advantages to clinical trial evaluation. With regard to Biogen's recent trials of aducanumab, this analysis of simulated data showed that the Bayesian framework would have permitted the data analysts to directly quantify the plausibility of the absence over the presence of aducanumab's treatment effect in study 301. Applying Bayes factors as a continuous measure of evidence also allowed us to quantify the effect of excluding rapid progressors. The added value of Bayes factors is that they can be regarded as a measure of uncertainty regarding the conclusion that the effect is absent. Similarly, the Bayes factor *quantifies the amount of evidence* for the presence over the absence of a treatment effect in study 302. Table 5 summarizes what our Bayesian analyses have revealed from the simulated aducanumab data that was not available from Biogen's original frequentist analysis.

In addition, we performed a Bayesian three-way ANOVA to quantify the evidence for the explanation regarding the diverging effects in study 301 and study 302 and whether rapid progressors across study arms affected CDR-SB, specifically that a disproportionally higher proportion of rapid progressors in the high-dose arm of study 301 reduced aducanumab's treatment effect in this study. The Bayesian ANOVA using both Bayes factors and model fit comparison showed that once the rapid progression speed indicator was used to predict CDR-SB, there was evidence against including the other factors. In particular, the outcome CDR-SB did not differ between study 301 and 302.

TABLE 5 Summary and comparison of conclusions regarding aducanumab's treatment effect based on frequentist analysis of the original data and Bayesian analysis of simulated data

Frequentist analysis of the original data			Bayesian analysis of simulated data	
Analysis	***P*-value**	**Conclusion**	**Bayes factor**	**Conclusion**
ANOVA in study 301 (excluding rapid progressors)	>.05	Study 301 failed to reject the hypothesis that aducanumab has no treatment effect.	$BF_{01} = 41.87$	The data of study 301 were 42 times more likely to occur under the null hypothesis of absence compared to the hypothesis of presence of a treatment effect.
ANOVA in study 302 (excluding rapid progressors)	.012	Study 302 rejected the hypothesis that aducanumab has no treatment effect.	$BF_{10} = 3.69$	The data of study 302 were four times more likely under the hypothesis of presence compared to the null hypothesis of absence of the treatment effect.

Abbreviation: ANOVA, analysis of variance.

Similarly, we quantified the evidence against the interaction between rapid progressors, study type, and treatment arm, suggesting that the simulated data did not indicate that rapid progressors in the high-dose treatment group of study 301 distort aducanumab's effect. Evidence from our simulated data was inconclusive regarding aducanumab's treatment effect across studies: The effect was neither demonstrably present nor absent. Our simulated data suggested that the effect—if considered present—was very small.

Overall, our Bayesian analyses do not contradict those of Biogen's frequentist analyses when the studies were considered separately; instead, they enrich them: In the Bayesian framework we can quantify the amount of support favoring a treatment effect in study 302, and contrast this with the absence of evidence in study 301. The continuous measures of evidence paint a more complete picture than significant in study 302 and non-significant in study 301. Furthermore, the Bayes factors could be used to assess replicability; that is, the flow of the evidence across the phases of trials, see for instance Ly et al.,[22] and Verhagen and Wagenmakers.[23] This is a key advantage of Bayesian analysis that deserves to be exploited in future trials.

Our Bayesian three-way ANOVA was conceptualized based on Biogen's proposition of the rapid progressor distortion effect affecting the high-dose group of study 301, but not originally run by Biogen. It provides additional information by quantifying the evidence against two possible explanations of the inconclusive evidence regarding aducanumab, which Biogen had proposed; namely, between-study effects and rapid progressor distortions. Both the between-study main effect and the rapid progressors distortion effect were recommended for exclusion (Table 4). This evidence of absence of a between-study effect in our analysis occurred despite the finding of a significant effect in study 302 and a non-significant effect in study 301 based both on our frequentist analysis of the simulated data and Biogen's analysis of the studies separately. This underscores Goodman's argument against the misconception that "studies with *P* values on opposite sides of .05 are conflicting."[4] Consequently, the simulated data indicate that neither the between-study effect nor the rapid progressor distortions seem plausible compared to the absence of any effect, that is, the null model. In essence, the approach of explaining differences between studies 301 and 302 by retrospectively defining and excluding a subgroup of patients based on their disease course can be considered a legitimate post hoc analysis that serves to generate a hypothesis of a rapid progressor effect to be tested in an independent trial. It must not be considered, however, as a confirmatory analysis.[24]

Unlike NHST using *P* values, Bayesian probability is not conditional on the likely existence of the collected data or hypothetically existing but never collected data under the null hypothesis.[1,6] Instead, it is conditioned on observed data only.[1,4,25] This might provide desirable clarity regarding estimates of treatment efficacy for regulatory decisions. Of note, as ours are simulated data, they provide no definite conclusion on the presence or absence of treatment effects of aducanumab. These data show, however, how Bayesian analysis could provide a worthwhile, informative perspective that is complementary to the originally frequentist approach to answer the burning questions the aducanumab trials have posed to the field.

It must be noted that Bayesian inference is not a one-size-fits-all solution to all our statistical problems; it is still susceptible to violated assumptions (e.g., normality, heterogeneity[16]), and it is equally vulnerable to unintentional (or intentional) misuse as frequentist techniques. Naturally, our conclusions here are limited by the fact that they are based upon a single simulation. We are not commenting on the actual efficacy of aducanumab as a treatment for AD as this is not within the scope of our simulated data. Though our synthetic CDR-SB outcome data behaves comparably to the original data, we did not simulate secondary study endpoints. Instead, we used the synthetic CDR-SB data as an exemplar for Bayesian analysis of a multi-study clinical trial. Our data analysis mirrored Biogen's as closely as possible, but there are numerous Bayesian modeling approaches that can be taken, for example, by choosing different priors or modeling strategies altogether. Interested readers may obtain an alternative Bayesian framework modeling strategy—programmed in R, using the "brms" package[26]—from the supporting information and our OSF folder. This strategy relies on assumptions about the outcome's distribution and compares models using leave-one-out cross-validation instead of the Bayes factor. Of course, a large variety of hypothesis testing and model comparison approaches are available in the Bayesian framework. The focus here on Bayes factor testing reflects the original design of the analysis of the aducanumab data and the fact that our simulation was restricted to publicly available data, so models using individual trajectories of change were not feasible for us. Another limitation worth noting is the absence of secondary endpoints in our synthetic data. The published information on these was insufficient for simulation as the briefing book states that they were non-normally distributed without providing histograms, skew, or curtosis. This precluded the creation of any synthetic secondary endpoints suitable for Bayesian re-analysis, and hence any reliable exploration. Future briefing books should contain equal amounts of information on all endpoints. Still, even with this restricted set of analyses, we believe that we could demonstrate the advantages of Bayesian over frequentist hypothesis testing for clinical trials.

Importantly, Bayesian modeling techniques are a powerful tool in our statistical toolbox. It is our view that so far, they have been seriously underused. In the past, this was explicable by the lack of sufficient computing power[25] and by the necessity of programming skills. However, with the advent of sufficient computing power, the inclusion of Bayesian approaches in widely used commercial statistical software packages, such as SPSS, STATA, or SAS, and the provision of free software packages, such as JASP,[15] and a multitude of packages in R,[27] Bayesian modeling techniques have become accessible to all. This development has been accompanied by a variety of recent open access publications to guide Bayesian analyses, providing details on the theoretical background,[1,6,28] *t*-tests,[29,30] ANOVA,[7,16,19,31] correlations,[7,17,19] and clinical trials.[32–38] Similarly, there has been a surge in textbooks.[39–41] We believe that acquiring

Bayesian thinking and analytical skills is a worthwhile investment for statistical and biomedical researchers.

In 1975, Lindley prophesied that the 21st century would have become Bayesian by 2020.[25] This is not yet the case, but Bayesian thinking and modeling will enrich our statistical conclusions, and in turn, produce more informative research.

ACKNOWLEDGMENTS

AGMT was supported by donations in favor of ALS research by André Greipel to DZNE, and by the Boris Canessa Foundation.

CONFLICTS OF INTEREST

Stefan Teipel participated on scientific advisory boards of Roche Pharma AG, Biogen, Grifols SA, and MSD, and received lecture fees from Roche Pharma AG and MSD. Alexander Ly, Johnny van Doorn, and Eric-Jan Wagenmakers declare their involvement in the open-source software package JASP (https://jasp-stats.org/), a non-commercial, publicly funded effort to make Bayesian statistics accessible to a broader group of researchers and students. Anna Temp, Yi Tang, and Michael W. Lutz have no conflicts of interest.

ORCID

Anna G. M. Temp https://orcid.org/0000-0003-0671-121X

REFERENCES

1. Wagenmakers E-J, Marsman M, Jamil T, et al. Bayesian inference for psychology. Part I: theoretical advantages and practical ramifications. *Psychon Bull Rev*. 2018;25:35-57.
2. Morey RD, Hoekstra R, Rouder JN, Lee MD, Wagenmakers E-J. The fallacy of placing confidence in confidence intervals. *Psychonomic Bulletin & Review*. 2015;23:103-123.
3. Benjamin DJ, Berger JO, Johannesson M, et al. Redefine statistical significance. *Nat Hum Behav*. 2018;2:6-10.
4. Goodman S. A dirty dozen: twelve p-value misconceptions. *Semin Hematol*. 2008;45:135-140.
5. Kim J, Bang H. Three common misuses of P values. *Dent Hypotheses*. 2016;7:73-80.
6. Wagenmakers E-J. A practical solution to the pervasive problems of p values. *Psychon Bull Rev*. 2007;14:779-804.
7. Wagenmakers E-J, Love J, Marsman M, et al. Bayesian inference for psychology. Part II: example applications with JASP. *Psychonomic Bulletin & Review*. 2018;25:58-76.
8. Wasserstein RL, Lazar NA. The ASA statement on p-values: context, process, and purpose. *The American Statistician*. 2016;70(2):129–133. https://doi.org/10.1080/00031305.2016.1154108
9. Araoye I, He JK, Gilchrist S, et al. A national survey of orthopaedic residents identifies deficiencies in the understanding of medical statistics. *J Bone Joint Surg Am*. 2020;102:e19.
10. Borak J, Veilleux S. Errors of intuitive logic among physicians. *Soc Sci Med*. 1982;16:1939-1943.
11. Wegwarth O, Schwartz LM, Woloshin S, Gaissmaier W, Gigerenzer G. Do physicians understand cancer screening statistics? A national survey of primary care physicians in the United States. *Ann Intern Med*. 2012;156:340-349.
12. Windish DM, Huot SJ, Green ML. Medicine residents' understanding of the biostatistics and results in the medical literature. *JAMA*. 2007;298:1010-1022.
13. Wessels AM, Dowsett SA, Sims JR. Detecting Treatment Group Differences in Alzheimer's Disease Clinical Trials: a Comparison of Alzheimer's Disease Assessment Scale - Cognitive Subscale (ADAS-Cog) and the Clinical Dementia Rating - Sum of Boxes (CDR-SB). *J Prev Alzheimers Dis*. 2018;5:15-20.
14. Peripheral and Central Nervous System (PCNS) Drugs Advisory Committee Meeting. Combined FDA and Applicant PCNS Drugs Advisory Committee Briefing Document on Aducanumab. 2020; USFDA.
15. The JASP Team. JASP (Version 0.14.1). 2020.
16. van Doorn J, van den Bergh D, Böhm U, et al. The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychon Bull Rev*. 2021;28(3):813–826. https://doi.org/10.3758/s13423-020-01798-5
17. vanDoorn J, Ly A, Marsman M, Wagenmakers E-J. Bayesian Inference for Kendall's Rank Correlation Coefficient. *The American Statistician*. 2018;72(4):303–308. https://doi.org/10.1080/00031305.2016.1264998
18. Field A, Miles J, Field Z. Discovering Statistics using R. *6.5.6 Kendall's Tau (Non-Parametric)*. London, England: SAGE Publications Ltd; 2012:225-229..
19. Ly A, Marsman M, Wagenmakers EJ. Analytic posteriors for Pearson's correlation coefficient. *Stat Neerl*. 2018;72:4-13.
20. Goodman SN. Toward evidence-based medical statistics. 1: the P value fallacy. *Ann Intern Med*. 1999;130:995-1004.
21. Goodman SN. Toward evidence-based medical statistics. 2: the Bayes factor. *Ann Intern Med*. 1999;130:1005-1013.
22. Ly A, Etz A, Marsman M, Wagenmakers EJ. Replication Bayes factors from evidence updating. *Behav Res Methods*. 2019;51:2498-2508.
23. Verhagen J, Wagenmakers EJ. Bayesian tests to quantify the result of a replication attempt. *J Exp Psychol Gen*. 2014;143:1457-1475.
24. Alexander GC, Emerson S, Kesselheim AS. Evaluation of aducanumab for Alzheimer disease: scientific evidence and regulatory review involving efficacy, safety, and futility. *JAMA*. 2021;325:1717-1718.
25. Lindley D, The future of statistics: A Bayesian 21st Century. Conference on Directions for Mathematical Statistics Applied Probability Trust; 1975.
26. Bürkner P-C. brms: an R package for Bayesian multilevel models using stan. *Journal of Statistical Software*. 2017;80(1):1–28. https://doi.org/10.18637/jss.v080.i01
27. Park JH, CRAN task view: Bayesian inference. 2021.
28. Masson ME. A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behav Res Methods*. 2011;43:679-690.
29. Demsar J, Repovs G, Strumbelj E. bayes4psy-An Open Source R package for Bayesian statistics in psychology. *Front Psychol*. 2020;11:947.
30. Quintana DS, Williams DR. Bayesian alternatives for common null-hypothesis significance tests in psychiatry: a non-technical guide using JASP. *BMC Psychiatry*. 2018;18:178.
31. van denBergh D, vanDoorn J, Marsman M, et al. A Tutorial on Conducting and Interpreting a Bayesian ANOVA in JASP. *L'Année psychologique*. 2020;120(1):73–96. https://doi.org/10.3917/anpsy1.201.0073
32. Cai C, Liu S, Yuan Y. A Bayesian design for phase II clinical trials with delayed responses based on multiple imputation. *Stat Med*. 2014;33:4017-4028.
33. Campbell G. Guidance for the use of Bayesian statistics in medical device clinical trials. In: *Biostatistics Do*, Food and Drug Administration; 2010. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/guidance-use-bayesian-statistics-medical-device-clinical-trials
34. Fayers PM, Ashby D, Parmar MKB. Tutorial in biostatistics: Bayesian data monitoring in clinical trials. *Stat Med*. 1997;16:1413-1430.
35. Ferreira D, Meyer N. Post Hoc Bayesian analyses. *JAMA*. 2019;321:1632.
36. Lewis RJ, Angus DC. Time for Clinicians to embrace their inner Bayesian?: reanalysis of results of a clinical trial of extracorporeal membrane oxygenation. *JAMA*. 2018;320:2208-2210.

37. Vail A, Hornbuckle J, Spiegelhalter DJ, Thornton JG. Prospective application of Bayesian monitoring and analysis in an "open" randomized clinical trial. *Stat Med*. 2001;20:3777-3787.
38. Yin G, Lam CK, Shi H. Bayesian randomized clinical trials: from fixed to adaptive design. *Contemp Clin Trials*. 2017;59:77-86.
39. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A. *Bayesian Data Analysis*. 3rd ed. CRC Press; 2013.
40. Kruschke J. *Doing Bayesian Data Analysis*. 2nd ed. Academic Press; 2015.
41. McElreath R. *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. CRC Press; 2015.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Temp AGM, Ly A, van Doorn J, et al. A Bayesian perspective on Biogen's aducanumab trial. *Alzheimer's Dement*. 2022;1-11. https://doi.org/10.1002/alz.12615