Please also respond to these additional comments by the committee:

*Our primary concern is that this is particularly technical in nature (compared to our typical original research articles), and felt that this should be heavily revised with our clinical readership in mind. This is of paramount importance in this case.

We have substantially revised the manuscript with this goal in mind. Given that the Fast Facts piece will describe our methods as well as the technical nature of our methods, we have moved the majority of the technical portions of the methods section to the appendix.

*Please expand your thoughts on the pathophysiology and why you hypothesized that the current understanding regarding the amyloid hypothesis is not correct.

The role of amyloid in the pathophysiology of AD has long been controversial, and the repeated failure of individual RCTs of anti-amyloid drugs to demonstrate cognitive benefits has magnified questions about whether amyloid beta is causal, or whether it is a result, i.e. marker, for other pathological processes such as blood-brain barrier dysfunction. Other pathologies, such as tau, blood-brain barrier dysfunction, mitochondrial dysfunction, microvascular disease, are also known or hypothesized to contribute to the clinical syndrome of AD. Our analysis was motivated by the premise that if amyloid does not play a causal role in AD, the focus on amyloid reducing therapies is misguided. On the other hand, we considered that when combining results from all trials, the small, individually statistically non-significant findings might add up to a more convincing picture in support of amyloid. In fact, our combined results indicate that reductions in amyloid were not associated with improvements in cognition, and our confidence intervals are fairly narrow. However, the results of our manuscript cannot be interpreted as conclusively falsifying the amyloid-cascade hypothesis. It is plausible that amyloid has a causal role, but that anti-amyloid interventions to date have not targeted the right individuals, during the physiologically relevant phase of disease progression, or for a long enough duration. We have added to the discussion section to further address this question in the manuscript.

*Do you think if you had had the benefit of a systematic review you might have come to a different conclusion? How do you foresee the current reliance on a limited selection of papers as influencing your results? In other words, do you think there could be a bias introduced with the literature selection?

As is the case with most systematic reviews, we had a comprehensive approach to selecting published research. With that in mind, we believe the potential bias due to missing data on some trials would be to overstate a potential benefit of amyloid reduction. This implies that our results -- indicating no benefit of amyloid reduction -- are a best-case story for amyloid. Trials that were 'successful' or hinted towards success would be more likely to publish results. Trials for which we could not obtain data were either completed but without published results, terminated early, or not yet completed. Completed trials for which results were never posted and never published are also unlikely to show an effect of decreasing amyloid on cognition. Terminated trials were likely terminated due to futility, and therefore were unlikely to show a
benefit of drug treatment on cognition and thus would be unlikely to show an effect of amyloid on cognition. Trials that are not yet completed may of course ultimately report relevant findings, which is why we included a web application to allow for convenient recalculation of our results to incorporate new findings. We have edited and added to the following to the text in the discussion section to help clarify this:

After review of publicly available reports and contacting pharmaceutical companies directly, we were also unable to obtain data for 20 trials that met our eligibility criteria. Most companies would complete and post trial results if the trial showed promising findings, so it is unlikely that drugs evaluated in trials with no available results showed evidence of benefit of amyloid reduction. Any bias from omitting data for these studies is likely to be in favor of a beneficial effect of amyloid. Our null results should thus be interpreted as an optimistic estimate for the cognitive benefit of amyloid reduction based on RCTs to date. Data for the "active, not recruiting" trials may become available at a future date and our estimates can easily be updated via the web application.

*We felt that the critical issue in the work is that you are using summary data on both the amyloid and MMSE. Thus you cannot take into account the covariance structure of these two outcomes. It seems that ideally if you want to understand a causal impact you need that information. While what is presented is likely the best that can be done with the data, can you comment on the impact that not having individual data presents for a causal interpretation?

This is not a problem that will substantively change our results. Conceptually, our analysis combines instrumental variables analyses of the individual trials (a method commonly based on summary data [cite mendelian randomization studies]) with a meta-analysis (nearly always conducted with summary data). Combining these two is methodologically novel and we believe particularly pertinent for interpreting evidence from RCTs of medications with well-established biological targets.

While ideally information on the covariance would be available, this covariance is negligible if error in change in measured SUVr is large compared to the variance in true change in SUVr and predictors of cognition that do not also affect amyloid (e.g. education, vascular risk factors, and other non-amyloid pathologies such as TDP-43) account for the majority of the variance in cognition. It is likely these are the case: amyloid-PET produces noisy measurements in SUVr and variance in amyloid most plausibly accounts for only a minority of the variance in cognition across individuals. Furthermore, if estimated covariances between measured SUVr and measured cognition become available, such information could be easily incorporated into the proposed estimation procedure.

To determine the potential effect of not including the unmeasured covariance in our estimation procedure, we performed simulations, a description of which is now included in Appendix 3. We performed simulations under 4 scenarios:
1. Amyloid Mediates All Cognitive Change: This scenario reflects a parameterization of the amyloid cascade hypothesis. In this scenario, there are no common causes of change in amyloid and cognition. That is, amyloid mediates all cognitive change.

2. No Confounding, No Effect of Amyloid: This scenario is identical to scenario 1, except that there is no effect of amyloid on cognition.

3. Confounding, An Effect of Amyloid: This scenario is similar to scenario 1, except that in addition to an effect of amyloid on cognition, there are common causes of change in amyloid and change in cognition.

4. Confounding, No Effect of Amyloid: This scenario is identical to scenario 3, except that there is no effect of amyloid on cognition.

Given the technical nature of the discussion on this topic and the revision of our methods section, we have moved the majority of this discussion to the Appendix. The revised text reads as follows:

We did not account for the covariance between measured cognition and measured SUVr that might be induced if other factors influence amyloid and independently affect cognition, since the covariance was not reported. Simulation results indicate that this is unlikely to appreciably affect our estimates (see Appendix 3.)

In your response please provide, point by point, your replies to the comments made by the reviewers and the editors, explaining how you have dealt with them in the paper.
Reviewer: 1

Comments:

Dear editor,

I am supportive of this article, although I would suggest making it more accessible to all publics before publication. Involving patients/carers in re-writing some parts and defining terms could be highly beneficial.

Thank you for this suggestion. We reached out to two individuals directly affected by dementia but with no specialized training in science or health research. The feedback was extremely useful and has guided our edits. They embraced the importance of the research topic (“a bunch of money [is] being spent on these drugs and they don’t have much evidence that they work. There needs to be a back-track. We have reason to believe but no evidence that they do work?”) and the value of a meta-analysis (“all these drugs using the same idea, do we need more evidence on that? [is this] like a universal study?”). As the reviewer anticipated, they provided suggestions for simplifying statistical language or restating in more commonplace terms.

We have added the following to the Patient and Public Involvement statement:

 Two family members of people living with Alzheimer's disease provided feedback on the framing and interpretation of results.

Dear authors,

Thank you all the hard work that must have gone into this paper and for sending in the manuscript. I am reviewing the manuscript based on my own perspective and in application of "general" patient experiences in healthcare practice and policy.

Below are my comments:

The chosen topic is relevant and important for patients and carers and has a significant impact on clinical practice. However, to make sure all publics can benefit from it, I encourage the authors to consider defining some jargons and explaining more clearly their methods and interpretations of the results (more on this below). Authors could also re-organise the discussion to provide the reader with a clear summary of the results, a strengths and limitations section as well as a conclusion that includes implications of this study for various stakeholders and perspectives for future research.

We reorganized and revised the methods and results discussion to make these sections more accessible to a broader audience. We added a concluding section that includes implications of this research and suggested future directions for research. We also defined a number of terms and added additional citations.
• Are the questions the paper addresses relevant and important to patients and/or carers?
This paper explores whether amyloid-beta reduction improve cognitive outcomes by integrating existing evidence. As they explain in the introduction, no treatment based on the popular amyloid cascade hypothesis in Alzheimer’s disease (AD) have been approved to date. Consequently, revisiting the body of evidence on amyloid-beta is both relevant and important to patients and/or carers who are expecting new and better treatments of AD.

While anti-amyloid therapies have not yet been approved, questions around approaches to drug development and the biology of AD are relevant to patients. Particularly given a dearth of effective treatment options, time and money should be spent on the most effective avenues of research and testing. It is important to rigorously evaluate whether targeting amyloid on the timescales of most trials is a viable strategy.

• Are there topics or issues that are missing, or need to be highlighted more?
The focus of this study is clear and narrow enough so that the issues highlighted form a coherent and complete body of information. It does not feel there are topics or issues missing, although it would be beneficial to clarify some of the issues covered that all readers might not be familiar with (e.g. amyloid cascade hypothesis, different types of trials and methods to group evidence).

We have added definitions and additional citations to clarify our terminology.

• Is the treatment or intervention suggested or guidance given something which patients/carers can readily take up? or does it present challenges?
Although this paper does not support any treatment or intervention patients and carers can readily take up the findings of this study and understand some of the challenges in developing new treatments and interventions in this area. Patients and carers are also provided with a tool to keep updated with the evidence in this area and potentially take up any new treatment or intervention ensuing from it. Indeed, the web-based interface provided that allows for a re-calculation of results is a great idea and contribution to a fast-moving field! This tool will be very helpful to all stakeholders et certainly encourage the public to engage with this publication.

• Are the outcomes described/measured in the study important to patients/carers? Are there others that should have been considered?
The primary outcome being studied (effect of amyloid reduction on cognitive change) derives from a hypothesis that has been the focus of a number of AD research studies over the years, as well as a promise for treatment. Consequently, the outcomes measured are relevant for patients/ carers and clinical practice.
However, as I suggest in the last point of this review, involving patients/ carers in the consideration of this primary outcome and of other secondary outcomes might have strengthened the study.
Three of our co-authors are involved in patient recruitment into cohort studies of MCI, AD, and dementia patients, and one of our co-authors is involved in patient care. We have updated the Patient and Public Involvement statement to reflect this.

• Do you have any suggestions that might help the author(s) strengthen their paper and make it more useful for doctors to share and discuss with patients/ carers?
  I suggest re-writing some sections to make them clearer to a wider audience, as well as defining jargons and explaining key concepts covered in this study. In particular, the Methods section and some parts of the results need to be made more accessible. More details on this is included below, in “Other comments”.
  Also, including a Box with simple graphics explaining the amyloid cascade hypothesis would be helpful for readers and doctors sharing this study.

We have taken care to address readability for a clinical audience. See “Other Comments” below for specific changes to the manuscript text.

• Do you think the level of patient/carer involvement in the study could have been improved? If there was none do you have ideas on how they might have done so?
  As the authors declare, the research was conducted without patient involvement. I believe patients and/ or carers could have been invited to collaborate in a number of stages, such as study design, definition of relevant patient outcomes that could have been added to the primary outcome being studied (effect of amyloid reduction on cognitive change).

Patients/ carers were also not included in the writing stage of this study. The paper, which contains jargon and unexplained technical concepts and methods, could have highly benefit from the consideration of lay voices at this particular stage. The result could be an accessible readable paper by all.

We reached out to two individuals directly affected by dementia but with no specialized training in science or health research. In addition, three of our co-authors are involved in patient recruitment into studies of MCI, AD, and dementia patients, and several co-authors are involved in patient care. We have updated the Patient and Public Involvement statement to reflect this. The co-author involved in patient care has played a significant role in revising the manuscript to address these comments about readability for a clinical audience.

Other comments

Abstract
Since this manuscript does not include a lay summary, the abstract needs to be accessible to all publics. To do so, please consider the following comments:
- l.1: explain what amyloid beta is
- l.8: non-US lay audience may not be familiar with FDA. Please define
- l.18: define RCTs
- l.25: explain “maximum likelihood” and “fixed effects model”
- l.34: introduce what an R Shiny App is

Introduction
This section is very informative, concise and well-written. However, some parts might not be clear to all publics. Please consider the following suggestions:
- l.6 p.4: define amyloid plaques and oligomers
- l.15 p.4: Re-introduce what the FDA is for a non-US audience
- l.16 p.4: Highlight the difference between amyloid oligomers and soluble oligomers – e.g. why treatments to target those separately?
- l.40 p.4: explain what “statistical power” (and thus, “underpowered” is)
- l.50 p.4: explain what a “modification of intent-to-treat meta-analysis based on instrumental variable analyses” entails and define “intent-to-treat meta-analysis based on instrumental variable analyses”
- l.3 p.5: please define “randomized trials”
- l.5 p.5: explain “randomization”
- l.8 p.5: please explain “differential effects”

Methods
This section requires major re-writing so that the methods followed in this study can be clear to a wide audience. On the other hand, the methods described are rigorous and valid, so I suggest changes simply to make this section more accessible.
- l.18 p.6: define what a “placebo control” is
- l.18-20 p.6: please provide a rationale for excluding “trials that did not have measures… within randomization arms”
- l.48 p.6: define “maximum likelihood estimator” and explain how it allowed you to overcome the limitations of standard meta-analysis methods
- l.50 p.6: how did the “principles of instrumental variable analysis” led you to the assumptions listed?
- The paragraph in p.7, l.16-24 needs to be simplified
- could you introduce a figure illustrating your method to determine the likelihood as explained from l.26 p.7?
- l.43 p.7: explain what “sensitivity analyses” are

Results
This section is concise and clear but might still contain too much jargon and implicit technical knowledge. This makes it is too obscure for a lay audience. Some suggestions to overcome this:
- l.33 p.9: explain what a “crosswalk” is
- l.43-46 p.9: this sentence is key to understand the results of the study. Please reformulate in simpler terms to emphasise the significance of the findings: e.g. what “consistent with the null” means might not be clear to all readers
- l.46 p.6: I believe it should be “yields”
- l.6 p.10: please remind the reader what the value of the “annual effect of APOE-ε4 carriage” is
- Figure 2: please provide further explanation of what is presents in the text in simple language. At the moment, it might be hard to read and interpret the figure for a wide audience
We have defined these acronyms and terms and moved portions of the methods section to the appendix. We’ve removed “R Shiny App” and just said “Web Application” to keep our terminology consistent. We replaced “null” with no effect of amyloid on cognition.

Discussion
This section is clear and informative. I suggest adding a “strengths of this study” and a “limitations of this study” to help structure it, and support the reader. For example, the paragraph in l.17-22, p.12 could go into the “strengths of this study” sub-section, whilst l.34-50 p.12 and l.9-19, p.13 could be included into the “limitations of this study” one. Similarly, a conclusive paragraph would help the reader understand the implications of your findings for patients, policy and practice and consider perspectives for future research.

We added the following standard subheadings for improved readability: strengths & limitations, implications, and conclusions.

- l.33 p.13: explain the concept of “covariance” and how attempting to account for it might have changed the results of this study

Given the technical nature of the discussion on this topic and the revision of our methods section, we have moved the majority of this discussion to the Appendix. The revised text reads as follows:

We did not account for the covariance between measured cognition and measured SUVr that might be induced if other factors influence amyloid and independently affect cognition, since the covariance was not reported. Simulation results indicate that this is unlikely to appreciably affect our estimates (see Appendix 3.)

I would like to thank the BMJ for giving me the opportunity to review this manuscript.
Reviewer: 2

Comments:
Thank you for the opportunity to review this work. Please note that I am a statistician, and so my review mostly relates to the statistical aspects of the manuscript.

The work seems well motivated. While no individual study has provided strong evidence of a causal effect of amyloid beta reduction on cognition, it may be that there is a critical mass of evidence when combined together in a meta-analysis. However, this does not appear to be the case here. The work is well performed, and I have few questions:

1) Is there any evidence of potential publication bias (or reporting bias)?

As stated above, we believe the potential bias due to missing data on some trials would be to overstate a potential benefit of amyloid reduction. This implies that our results -- indicating no benefit of amyloid reduction -- are a best case story for amyloid.

We considered this question while formulating our analysis and whether a funnel plot would be feasible. In this case, such methods did not seem applicable since estimates for benefit of reduction of amyloid on cognitive decline would seem favorable if the drugs were indeed effective by this mechanism, and only drugs with believed to have a more favorable effect on amyloid or cognition (based on animal and phase 1 studies) would be pursued. So, we should not necessarily expect to see symmetry. Additionally, heterogeneity in effect estimates could be due to cognitive harms of the drugs by other mechanisms or benefits associated with specific amyloid reduction strategies.

We did evaluate for heterogeneity in effect by drug type (small molecule versus antibody) and by data publication status to specifically address concerns about publication bias.

2) It is possible that studies with longer follow-up will see greater effect. Is this the case here? Would be good to know about the length of follow-up (eg in the Appendix Table).

We have added length of follow-up to the table appendix. We note that there was little variability in follow up--most studies were 1-2 years. We agree this is an important possibility and emphasize this in the discussion.

3) I'd be interested to see a graph of the difference in amyloid levels and the change in MMSE (one point per study, with 95% confidence intervals). This would be more of a raw presentation of the study-level estimates.

We’ve added plots of the raw data and fits to Appendix 1.
4) I didn't fully understand what was being reported with respect to BAN2401. You state "BAN2401 has been reported to be effective (AAIC 2018)", but I don't see a significant estimate for that study. I was therefore slightly confused.

We rephrased this statement because it was confusing and tangential. We emailed Eisai Pharmaceuticals for precise data necessary for our analysis (rather than extracting from the presentations at the conference); the data they provided indicated a non-significant improvement in MMSE associated with reductions in amyloid. As noted elsewhere, this observation does not rule out the potential relevance of BAN2401 on cognition, if the medication influences cognition via other mechanisms besides amyloid. We plan to publish the data for BAN2401 and other trials with the manuscript.

5) While no blame can be assigned to the authors, it is a shame that so many studies failed to respond.

We agree with this sentiment but any potential bias from trials for which no data were released is unlikely to explain our largely null results. Our goal was to obtain data from as many trials as possible to mitigate any biases due to selection into this review. We attempted to obtain data from any source available and directly contacted pharmaceutical companies to obtain additional data, but we did not obtain additional data for the reasons outlined in the manuscript.

We believe the potential bias due to missing data on some trials would be to overstate a potential benefit of amyloid reduction. This implies that our results -- indicating no benefit of amyloid reduction -- are a best case story for amyloid. Trials that were 'successful' or hinted towards success would be more likely to publish results. Trials for which we could not obtain data were either completed but without published results, terminated early, or not yet completed. Completed trials for which results were never posted and never published are also unlikely to show an effect of decreasing amyloid on cognition. Terminated trials were likely terminated due to futility, and therefore were unlikely to show a benefit of drug treatment on cognition and thus would be unlikely to show an effect of amyloid on cognition. Trials that are not yet completed may of course ultimately report relevant findings, which is why we included a web application to allow for convenient recalculation of our results to incorporate new findings.

We have edited and added to the following to the text in the discussion section to help clarify this:

*After review of publicly available reports and contacting pharmaceutical companies directly, we were also unable to obtain data for 20 trials that met our eligibility criteria. Most companies would complete and post trial results if the trial showed promising findings, so it is unlikely that drugs evaluated in trials with no available results showed evidence of benefit of amyloid reduction. Any bias from omitting data for these studies is likely to be in favor of a beneficial effect of amyloid. Our null results should thus be interpreted as an optimistic estimate for the cognitive benefit of amyloid reduction based on RCTs to date. Data for the "active, not recruiting" trials may become available at a future date and our estimates can easily be updated via the web application.*
6) The authors speculate as to potential reasons why amyloid beta mechanisms may still be a valid treatment target: "However, it is plausible that reductions in amyloid-beta improve cognitive outcomes, but that the drugs evaluated harmed cognition via other mechanisms." and "When interpreting these results with respect to the amyloid cascade hypothesis, we note that amyloid reduction may have delayed effects on cognition that do not manifest until years later." However, they do not speculate as to the elephant in the room - that amyloid beta is not a relevant therapeutic target. While it would be overreach as a reviewer to insist on this, for me this is important to provide balance. Given the potential reasons cited for how amyloid-beta may still be a causal risk factor, it would seem worthwhile to also suggest reasons why amyloid-beta may not be a causal risk factor (ie it is a symptom of cognitive decline, not a cause).

We agree with the reviewer’s characterization of the elephant in the room and have added this to the discussion.

These results provide evidence that amyloid reduction alone is unlikely to slow cognitive decline substantially within the follow-up period of most current trials. In all analyses of amyloid, there is a concern that amyloid may not itself be a causal step in the pathologic process, but merely a biomarker that is modified by the causal pathology. These results may indicate anti-amyloid drugs are not a viable strategy for treatment or prevention of AD and other potential targets merit more attention.

7) Appendix 4 - is there uncertainty in these numbers? (Given that the are estimates, it seems there should be.)

For two studies we needed to convert another cognitive outcome to the MMSE, and, indeed, this conversion may be imprecise. However, alternative crosswalks may be tested using our web application. The following text addresses this limitation:

Without full reports, another limitation of this study is possible error in the input data. Other crosswalks between cognitive measures may be preferable. To address these concerns, we have published an interactive version of our analysis online at https://amyloidintegratingevidence.shinyapps.io/application/. The interface is autopopulated with the values we used in the analysis for each trial, and we provide the option of editing those values or additionally including a hypothetical trial. Thus, input values can be manually modified and new data can be added, to recalculate individual trial and pooled estimates under different assumptions.
Reviewer: 3

Comments:
This study addresses the issue of the putative value of reducing amyloid in individuals with Alzheimer’s disease (AD) with respect to their cognitive improvement. The authors reviewed the literature on published clinical trials involving amyloid reduction and concluded from the aggregated results of these trials that the pooled estimate of the effect of reducing amyloid by 0.1 standardized uptake ratio units (SUVr) would result in the improvement on the MMSE of 0.02. Hence, the implication is that these trials are futile.

This is an elegant demonstration of analyzing data across multiple clinical trials aimed at the same target, reducing amyloid. As such, it confirms what is known in the literature, that all of the amyloid lowering trials have been negative thus far and suggests that further work on targeting this protein might be ill advised. The analysis is well done and supports the views expressed by the authors. However, there are a few methodological concerns.

This analysis extends previous knowledge in two ways: we effectively increased the sample size for evaluating amyloid targeting mechanisms by meta-analyzing results, and we provide a systematic way to combine information on medication effects on amyloid and cognition to derive an estimate of how much reductions in amyloid improve cognition. The staged development of pharmaceutical therapies and the cost of clinical trials makes it likely that any one trial would be underpowered to detect a small benefit. Statistical power can be improved by combining results from multiple trials of different medications that all target amyloid. Evidence from multiple trials of amyloid targets has not been systematically combined previously and trial data have not been leveraged to evaluate whether changes in amyloid are likely to improve cognition. Our study provides the best available estimate of the effect of amyloid reduction on cognition, and with the meta-analyzed results we have fairly narrow, informative confidence intervals.

The authors discuss reducing amyloid by the SUVr metric but do not take into account the multiple tracers that could have been used to come to this conclusion. That is, there are four tracers available for measuring amyloid, and likely different tracers were used in the clinical trials. As such, a single claim of a SUVr metric is not informative. The field has invoked the centiloid method as a means to compare the impact of various tracers, and this should be considered by the authors.

Most trials used florbetapir. We separated our results to show findings are very similar when evaluating only studies that used the same tracer. We now present in the Appendix results stratified by tracer and a scaled overall estimate. We added the following text to the methods:

*We also conducted sensitivity analyses stratifying by the tracer used in amyloid PET and applying a scaling transformation to address the possibility that estimates from different tracers were not comparable.*
The authors have also chosen to use the MMSE as the single outcome metric, and while they admit this is insensitive and of marginal utility, it was the one most commonly used measure across trials. As such, they are obligated to use this type of a measure, but the MMSE is hardly a measure of efficacy of clinical trials designed to lower amyloid.

As has been demonstrated recently, more sensitive measures, that combine multiple instruments, including the MMSE, such as PACC measures, have been thought to be more relevant. Nevertheless, this does not negate the conclusion of the authors that lowering amyloid thus far has not been demonstrated to be efficacious. It is not a problem with the metric as such, but the MMSE is hardly appropriate.

Many prior researchers have noted that the MMSE is not sensitive to cognitive changes among high functioning individuals, because of the ceiling effect. This limitation is less relevant for the trials because most enrolled individuals with subtle to moderate cognitive impairment. Thus, with the populations enrolled in these trials, cognitive declines or improvements would likely have been detectable with the MMSE. To further confirm this, we repeated our analysis using various versions of the ADAS-Cog, the next most commonly reported outcome. See Appendix 1. We have added the following text to the methods:

> We performed sensitivity analyses restricting to antibody drugs and with and without the unpublished trials of BAN2401 and Aducanumab. Additional sensitivity analyses are given in Appendix 1. We repeat the analysis using alternative cognitive outcomes (CDR-SB and various versions of the ADAS-Cog), where available.

We also note that for BAN2401, one of two potentially effective drugs, the cognitive outcome we use is the ADCOMS, rescaled to approximate MMSE points.

The most recent trials, including BAN2401 and aducanumab were the two trials that suggested possible clinical efficacy. The authors included BAN2401 results in their analyses and took into account the Apolipoprotein E4 effect. However, the authors were limited to press release data for the aducanumab trials. The aducanumab trials were the most recent studies that suggested possible efficacy with an impact on the primary and all three secondary measures in one of the two trials. The other trial, however, was negative. As such, the most suggestive data on the role of a relationship between reduction of amyloid and cognitive benefit came from the trial on which they have the least data.

We looked for updated data on aducanumab and found an updated report from April 2020. (See https://investors.biogen.com/static-files/f91e95d9-2fce-46ce-9115-0628cfe96e83.) This report contained the same data that we used in our initial analysis. We also point out that we obtained aggregated estimates for changes in SUVr and cognition by treatment arm. While this data is unpublished, it is no different in quantity than data obtained for other trials.

As stated above, we repeated our analysis using various versions of the ADAS-Cog and CDR-SB (see Appendix 1) and did see a significant effect of amyloid on change for the EMERGE trial.
and for Aducanumab overall. We added text to our manuscript to address this issue. We also note that our web application allows the user to update data, should additional data on Aducanumab become available during the FDA approval process. We have added the following to the results and discussion to address these issues:

*Biogen submitted an application for FDA approval of Aducanumab based on new analysis of updated data from EMERGE and ENGAGE studies. Should new data become publicly available during this process, estimates can be updated using our web application. Evaluating only data from Aducanumab for the ADAS-Cog 13 and CDR-SB cognitive outcomes, amyloid reduction appears to be effective at slowing cognitive decline (see Appendix 1). We cannot rule out the possibility that targeting amyloid-beta by a specific mechanism will slow cognitive change slows cognitive change while amyloid reductions achieved by other mechanisms have no benefit. However, our results are also consistent with the possibility that the single successful Aducanumab trial was a chance result or that aducanumab has off-target effects, i.e., aducanumab benefits cognition via mechanisms other than amyloid reduction. The potential for chance findings increases when data from individual trials are analyzed at multiple time points or with multiple cognitive measures. Our meta-analyzed findings thus call into question the importance of amyloid reduction, despite promising results in a single study. Hypothetically, to demonstrate that cognition improves cognition in the pooled analysis would require a new trial of similar scale to EMERGE and with large, dose-dependent effects on cognition.*

Finally, while all of these trials have been negative with regard to clinical impact, that does not necessarily negate the rationale for addressing amyloid. That is, it would simply take one trial with a beneficial clinical outcome suggesting that amyloid lowering does, in fact, have a subsequent impact on clinical efficacy to warrant its continued investigation.

Since an average of 1 in 20 trials of a non-effective target will appear effective due to chance, this effect of the biomarker on the disease outcome should be aggregated across multiple trials to evaluate the plausibility of an overall effect of targeting that biomarker. It is possible that other trials did not appear effective because many, but not all, did not lower amyloid as effectively as Aducanumab. However, trials which did not lower amyloid very much have large uncertainties associated with them and would contribute overall less to overall estimates. The following text addresses these concerns:

*Among all trials, unpublished reports on Aducanumab and BAN2401 indicated statistically significant benefit in at least one cognitive measure in intent-to-treat analyses. These associations should be interpreted with caution because of the potential for a chance result, which becomes more likely as the number of trials targeting the same mechanism increases. With 14 trials, if we assume no effect of amyloid on cognition, there is a 51% chance that at least one result would meet the typical threshold of p<0.05 for statistical significance. If we consider all 34 trials that met our inclusion
criteria, plus two additional trials of Aducanumab, there is an 84% chance. It is therefore critical to consider pooled evidence from all available trials to contextualize results for any single trial.

We did consider the effect on our pooled estimate of a large trial that successfully moved amyloid and for which there was a clear, and significant dose response in cognition. No trials examined met all of these criteria. Indeed, such a trial might lead to a significant pooled estimate. However, such a pooled estimate would not lead to clinically significant changes in cognition with realistic reductions in amyloid. We added the following to the discussion:

Since some trials changed amyloid very little (see Appendix 1), our overall null findings could be sensitive to the addition of a highly successful trial. We evaluated how our pooled estimate would hypothetically change with the addition of a new, hypothetical trial (of similar size and with similar effects on amyloid as the EMERGE trial). We defined a hypothetical trial with a low-dose treatment group that decreased SUVr by -0.15 and increased MMSE by 0.7 relative to the placebo group and a high-dose treatment group that decreased SUVr by -0.30 and increased MMSE by 1.4 relative to the placebo group. The pooled estimate of the effect of 0.1 reduction in SUVr on MMSE incorporating this trial plus all prior results would be 0.094, 95% CI: (0.0038, 0.18).

The rationale in this study also assumes that amyloid is the dominant factor in having an impact on clinical outcome when, in fact, there are likely numerous mediating events that intervene between amyloid processing and cognition. In this sense, the amyloid cognition relationship is an oversimplification of a likely very complicated process.

Our analysis is premised on the assumption that these drugs primarily act through reducing amyloid. If a drug both effectively reduces amyloid and improves cognition via other mechanisms, our estimates would attribute this to efficacy of amyloid reduction on reduction in cognitive decline. We have added the following text to the methods to address this:

Furthermore, our analysis is premised on the assumption that these drugs primarily act through reducing amyloid. If a drug both effectively reduces amyloid and improves cognition via other mechanisms, our estimates would attribute this to efficacy of amyloid reduction in reducing cognitive decline.

In summary, this is a scholarly exercise in evaluating clinical trials data on amyloid lowering drugs. However, it does not add any significant insight into clinical trials design and interpretation of amyloid lowering drugs than the field already appreciates from the trials that have been completed.

We disagree. Trial results have not been synthesized to offer direct evidence either for or against the amyloid cascade hypothesis. There are multiple randomized trials with evidence of engagement with amyloid, which can and should be used to draw conclusions about the mechanistic importance of this biological target on cognitive outcomes. Information on target
engagement of new medications is routinely used to encourage support and pursuit of those medications. We should similarly take advantage of that evidence to evaluate whether changes in the proposed targets induce improvements in clinical outcomes. The following text addresses this concern:

This is the first report combining all available results into a single estimate of the effect of amyloid reduction on cognitive change.

This combined result is much more precise than results from any individual trial and indicates that substantial cognitive benefits of amyloid reduction are unlikely within the time frame of the conducted trials.

The potential for chance findings increases when data from individual trials are analyzed at multiple time points or with multiple cognitive measures. Our meta-analyzed findings thus call into question the importance of amyloid reduction, despite promising results in a single study.

These findings suggest amyloid reduction per se is unlikely to have notable cognitive benefits for patients within the time frame of typical trials.
Reviewer: 4

Comments:
This study aims to determine if amyloid reduction intervention improves cognitive outcome during Alzheimer’s disease by comparing evidence across all randomized clinical trials. The progressive formation of amyloid during Alzheimer’s disease prompted the amyloid cascade hypothesis that suggests build-up of amyloid may cause neuronal pathology and cognitive decline. Therefore therapeutics targeted at blocking this build-up should result in preventing this decline. The authors show through cross-comparison of as many trials as possible that fit into their criteria (including accurate measure of amyloid build-up and cognitive assessment) that amyloid reduction did not substantially improve or cause greater decline in cognition.

This article represents an important contribution to this field as it indicates that reducing amyloid is not beneficial and casts doubt on the amyloid cascade hypothesis itself. In general the reviewer agrees with the findings in this study and found it well-performed, well-written and informative. Furthermore the provision of an app to continue to include and assess future trials as the data become available allows future re-assessment of these findings. I would recommend accepting this article for publication in its current form.

Thank you.
The aim of this study was to examine whether amyloid-beta reduction improves cognitive outcomes using pooled summary information from randomized trials of amyloid targeting therapies. The researchers reviewed ClinicalTrials.gov and identified 34 trials which met their inclusion criteria and were able to analyze information of 14 trials of 8 different drugs using maximum likelihood to estimate the effect of amyloid-PET reduction on cognitive change on the MMSE. Their aggregated results show that amyloid reduction strategies did not substantially improve cognition. The study adds to the field by overcoming the lack of statistical power to estimate the effect of amyloid reduction on cognitive outcomes by combining results from multiple amyloid-targeting trials. The online interface created by the authors that allows for recalculation of results with new or hypothetical results is another valuable addition to the research field. The study has several limitations that should be addressed in light of the findings.

Major comments:
None.

Minor comments:
- In the abstract, the authors should rephrase to explain better what is meant by ‘more precise’ (page 3, line 30).

We have added the following sentence:

For example, findings from any one study may have wide confidence intervals that would not allow us to rule out either substantial benefit or substantial harm. Combining evidence from multiple trials provides more precise effect estimates (i.e., narrower confidence intervals).

- The authors could mention the lack of data on individuals in even earlier stages of disease (preclinical AD) and potential implications of this for their results. Perhaps the results from failed AD anti-amyloid clinical trials in patients in early symptomatic phases of AD suggest that clearing of amyloid fails to improve cognitive functioning once clinical expression is present.

We have added the following to the discussion to address this limitation:

We also note trials were typically restricted to individuals who were amyloid positive and were already demonstrating cognitive impairment. If amyloid reduction strategies are only effective in individuals who have not accumulated significant amyloid or in the preclinical stages of AD, we would fail to find that amyloid reduction slows cognitive decline.

- It seems contradictory to state in the introduction (page 5, line 9-10) that the plausibility of differential effects of amyloid reduction on cognitive change by drug type will be evaluated, and
at the same time assume that the effect of reducing amyloid on cognition does not vary with mechanism by which amyloid-beta is targeted (page 7, line 8-10). The authors should write this down more clearly.

We have revised this.

- ‘Cognition’ is spelled wrong multiple times throughout the article (‘cognition’).

We have corrected this misspelling.

- The authors should include more information on whether they accounted for drug dosage in their analyses, as well whether they accounted for potentially differing drug-dosages for APOE-e4 carriers and non-carriers (it was only stated that analyses were adjusted for proportion of APOE-e4 carriers in each group).

Our analysis implicitly takes into account different drug dosages across randomization arms to the extent the different doses elicited different amyloid effects. Since we were examining the effect of reductions in amyloid on cognitive decline, differing drug dosages provided an instrument for changes in amyloid. As such, under some linearity assumptions, drug dosages and trial length were irrelevant to evaluating the effect of amyloid reduction on cognitive change. Appendix 1 now contains additional methodological information, as well as graphs of the raw data by drug dosage and fits.

- Did the authors obtain any information on patient age that could be incorporated in the analyses?

Only aggregated data were available and trials typically examined similar age ranges. As such, we did not adjust for mean age.

- The authors should elaborate on why they opted for a 0.1 change in SUVR.

The drugs most effective at reducing amyloid did so by 0.2 to 0.3 SUVr units, while some drugs reduced amyloid by far less. We therefore felt that 0.1 reductions in SUVr were a reasonable benchmark. Changes in cognition for smaller changes in SUVr can be determined by rescaling our estimates. We added the following text to the results to address this:

The drugs most effective at reducing amyloid did so by 0.2 to 0.3 SUVr units, while some drugs reduced amyloid by far less. We therefore present estimated effects on cognition associated with a 0.1 unit reduction in SUVr.

- Did the authors obtain any additional information on discontinued trials?

Some terminated trials are included in our analysis if trial data was posted on clinicaltrials.gov. Of trials for which we could not obtain data a majority were terminated or not yet completed.
- It is unclear whether the authors accounted for a smaller range of MMSE scores on the crosswalks due to the inclusion criteria of some trials that use a minimum and/or maximum score on the MMSE for inclusion.

Our crosswalks converted changes in another cognitive measure to changes in MMSE. Since these trials did not provide pre- and post-MMSE scores (otherwise we would have used the MMSE data), we could not account for MMSE inclusion criteria in our crosswalks.

- Was there a difference in findings for trials on MCI patients and trials on AD dementia patients?

We were also interested in this question. However, trials typically included patients MMSE scores consistent with both MCI and AD. Unfortunately, no meaningful stratification by baseline cognitive status was possible.

- On page 12, lines 46-48, it is mentioned that the MMSE has low sensitivity to cognitive deterioration in cognitively normal adults. Since the present study did not include any cognitively normal adults, I would suggest omitting this sentence.

Several reviewers asked us to address why we used the MMSE. Instead of removing this sentence, we have revised the sentence in question as follows:

    Though MMSE has low sensitivity to cognitive deterioration in cognitively normal adults, nearly all of these studies enrolled cognitively impaired patients who manifested MMSE deterioration, indicating that the MMSE is sensitive to changes in cognition in the trial populations.

- Page 13, lines 34-40: the sentence starts with ‘this covariance is negligible’ and ends with ‘then this covariance term is relatively negligible’.

We have deleted the word relatively.

- The discussion could be improved by adding more perspective and in-depth discussion of results, and by relating these findings to a more holistic view that incorporates the great number of genes and processes involved such as tau, neurovascular factors, inflammatory factors, and oxidative stress in AD pathogenesis.

We have added the following conclusions section to address this:

    Medications that effectively reduce the level of amyloid in the brain comprise a large fraction of therapeutics in the pipeline for Alzheimer’s disease prevention. Evidence of amyloid reduction is routinely touted as evidence of target engagement for new medications and support for the promise of ultimate efficacy in improving cognitive
outcomes. Over the last two decades, numerous amyloid-targeting medications have been evaluated in Phase 1, 2, or 3 trials. Although two medications appear promising in post-hoc analyses, these data remain unpublished on ClinicalTrials.gov or in a peer-reviewed journal. The repeated failures of anti-amyloid drugs call into question whether amyloid is a viable target for the prevention and treatment of Alzheimer’s disease. In prior analyses, however, each of these trials has been evaluated individually, with no prior research combining the evidence from all trials to evaluate whether amyloid reduction is likely to improve cognitive outcomes.

These findings suggest amyloid reduction per se is unlikely to have notable cognitive benefits for patients within the time frame of typical trials. This does not conclusively invalidate the amyloid cascade hypothesis because amyloid reduction may have delayed effects on cognition that do not manifest until years later. If this is the case, however, amyloid reduction trials would need to substantially extend the typical follow-up period to detect any benefit. Amyloid reduction strategies might be effective in different patient populations, for example people with much earlier stages of disease. The hypothesis that individuals with lower baseline amyloid burden reap greater cognitive benefit from amyloid reduction could be evaluated with additional data from extant trials.