Does amyloid-beta reduction improve cognitive outcomes? Integrated evidence from randomized trials of amyloid targeting therapies

<table>
<thead>
<tr>
<th>Journal:</th>
<th>BMJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>BMJ-2020-057847</td>
</tr>
<tr>
<td>Article Type:</td>
<td>Research</td>
</tr>
<tr>
<td>BMJ Journal:</td>
<td>BMJ</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>01-May-2020</td>
</tr>
</tbody>
</table>
| Complete List of Authors: | Ackley, Sarah; University of California San Francisco, Epidemiology and Biostatistics  
Zimmerman, Scott; UCSF  
Brenowitz, Willa; University of California San Francisco Department of Epidemiology and Biostatistics  
Gold, Audra; University of California San Francisco  
Tchetgen Tchetgen, Eric; University of Pennsylvania  
Manly, Jennifer; Columbia University  
Mayeda, Elizabeth Rose; University of California Los Angeles Fielding School of Public Health, Los Angeles, CA  
Filshiein, Teresa; University of California San Francisco Department of Epidemiology and Biostatistics  
Power, Melinda; George Washington University School of Public Health and Health Services  
Elahi, Fanny; University of California San Francisco  
Brickman, Adam; Columbia University  
Glymour, M. Maria; University of California San Francisco, Department of Epidemiology and Biostatistics; University of California, San Francisco |
| Keywords:           | Alzheimer's Disease, Amyloid, Meta-analysis, Antibody Drug, Randomized Controlled Trial, Amyloid-PET |
Does amyloid-beta reduction improve cognitive outcomes? Integrated evidence from randomized trials of amyloid targeting therapies

April 30, 2020

Sarah F. Ackley, PhD, Department of Epidemiology and Biostatistics, University of California, San Francisco

Scott C. Zimmerman, MPH, Department of Epidemiology and Biostatistics, University of California, San Francisco

Willa D. Brenowitz, PhD, Department of Psychiatry, University of California, San Francisco

Audra L. Gold, MS, Department of Epidemiology and Biostatistics, University of California, San Francisco

Eric J. Tchetgen Tchetgen, PhD, Department of Statistics, University of Pennsylvania

Jennifer J. Manly, PhD, Taub Institute for Alzheimer’s Disease and the Aging Brain, G.H. Sergievsky Center, Department of Neurology, Columbia University

Elizabeth Rose Mayeda, PhD, Department of Epidemiology and Biostatistics, University of California, Los Angeles

Teresa J. Filshtein, PhD, 23&Me

Melinda C. Power, PhD, Department of Epidemiology, George Washington University Milken Institute School of Public Health

Fanny Elahi, MD, UCSF Weill Institute for Neurosciences, University of California, San Francisco

Adam M. Brickman, PhD, Taub Institute for Alzheimer’s Disease and the Aging Brain, G.H. Sergievsky Center, Department of Neurology, Columbia University

M. Maria Glymour,* ScD, Department of Epidemiology and Biostatistics, University of California, San Francisco

*Corresponding Author: M. Maria Glymour, 550 16th Street, San Francisco, maria.glymour@ucsf.edu, 415-514-8014

https://mc.manuscriptcentral.com/bmj
Abstract

Reduction of amyloid beta has been a primary focus of new therapies for prevention or treatment of Alzheimer’s Disease (AD). No amyloid-targeting therapies have progressed sufficiently to receive FDA approval, bringing into question amyloid’s hypothesized role in AD development. Trials of these drugs have been analyzed individually to validate specific therapies but have not been fully leveraged to evaluate whether and on what time scale reductions in amyloid are likely to improve cognition. In this analysis, we pool summary information from 14 randomized trials of amyloid-targeting therapies to estimate the effect of amyloid reductions on cognitive change.

We reviewed ClinicalTrials.gov to identify randomized controlled trials of therapies for prevention or treatment of AD targeting an amyloid mechanism. Analyses included trials for which we could obtain information on both change in brain levels of amyloid measured with amyloid-PET and change in at least one cognitive test score reported for each randomization arm. Using randomization as an instrument, we used maximum likelihood to estimate the effect of amyloid reduction on cognitive change using a fixed-effects model. We estimate the effect of reducing amyloid by 0.1 standardized uptake value ratio units on change in the Mini-Mental State Examination (MMSE).

Aggregated results from all trials were more precise than estimates from any single trial. The pooled estimate for the effect of reducing amyloid by 0.1 standardized uptake value ratio units was an improvement in the MMSE of 0.02, 95% CI: (-0.05, 0.09) points. We provide an R Shiny app allowing for the re-estimation of our results when new data become available, and to illustrate the magnitude of the new evidence that would be necessary to achieve a pooled estimate supporting benefit of amyloid reduction.

Pooling evidence from all available trials reporting both amyloid reduction and change in cognition, we find that amyloid reduction strategies, in aggregate, did not substantially improve cognition.
Introduction

Amyloid plaques and oligomers are hypothesized to cause a cascade of pathological events resulting in cognitive decline in Alzheimer’s disease (AD) (1, 2, 3). Motivated by the amyloid cascade hypothesis, reducing amyloid-\(\beta\) in the brain has been a primary aim of new therapies for prevention or treatment of AD (4). While the presence of amyloid plaques and oligomers in the brain is highly correlated with the progression of AD (5, 6), the mechanisms by which amyloid might mediate neuronal pathology are currently not well understood (7). To date, no anti-amyloid therapies have progressed sufficiently to receive FDA approval (8). Drugs have targeted amyloid plaques, amyloid oligomers, and soluble oligomers and have been performed in populations with mild to moderate AD, as well as earlier stages of disease (prodromal AD) (9). The majority of trials of amyloid targets failed to produce positive results in either early or late stages of disease. The negative findings from these trials have prompted skepticism about amyloid’s role in producing neuronal pathology, and many have instead argued that amyloid may be a marker for other pathological processes and therefore is not a viable drug target (10, 11).

However, no single trial provides conclusive evidence about the potential impact of amyloid reduction on cognitive decline. Most trials are powered to evaluate specific therapeutics, and are not designed to accurately estimate the effect of amyloid reduction per se on cognitive outcomes. Individual trials are generally small, and often data on amyloid burden is assessed only in a subsample of participants (e.g. trials reviewed in 12). For any individual trial of an amyloid-targeting drug that fails to deliver cognitive benefits, there are several alternative explanations: First, amyloid reduction does not improve cognition. Second, the mechanism by which amyloid is targeted may affect whether and the extent to which reductions in brain amyloid slow cognitive decline (e.g. 13). Alternatively, 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 (9). This lack of statistical power is an important problem because even a small benefit for cognition could be of substantial clinical interest and would lend evidence to the amyloid cascade hypothesis. This lack of statistical power to detect an effect could be overcome by combining results from multiple trials of different medications that all target amyloid. To date, evidence from multiple trials of amyloid targets has yet to be systematically combined, nor have trial data been leveraged to evaluate whether changes in amyloid are likely to improve cognition. Using a modification of intent-to-treat meta-analysis based on instrumental variable analyses (IV) (14), it is possible to aggregate results of multiple studies into one combined estimate of the effect of decreasing amyloid on cognitive decline.

Here, we analyze individually and in aggregate changes in brain amyloid and changes in cognition from...
randomized trials of amyloid reducing drug therapies to estimate the effect of amyloid reductions on cognitive change. Using randomization as an instrumental variable for amyloid reduction, we evaluate the plausibility of the hypothesis that reductions in amyloid will slow cognitive decline. We also aggregate results across multiple studies of multiple drugs and evaluate the plausibility of differential effects of amyloid reduction on cognitive change by drug type (antibody vs. small-molecule drugs).
Methods

Data Sources

On May 9, 2019, The Alzheimer Research Forum (alzforum.org) was searched for “Amyloid-Related” therapeutics for “Alzheimer’s Disease” and “Mild Cognitive Impairment” (MCI), which yielded a comprehensive list of drugs. We then searched for trials of these drugs on ClinicalTrials.gov and restricted to trials for “Alzheimer’s Disease” or “Mild Cognitive Impairment” that were completed, terminated, or “Active, not recruiting.” Based on the information available on ClinicalTrials.gov, we excluded trials that did not have a placebo control and trials that did not have measures of brain amyloid quantified using the standardized uptake value ratio (SUVR) obtained from amyloid-PET and change in a cognitive score within randomization arms.

Study Selection

We excluded any trials for which the evidence indicated no effective reduction in SUVR as a result of treatment (one trial of ACC-001: NCT01227564). We obtained data, when available, from ClinicalTrials.gov, peer-reviewed publications, or other publicly available materials, such as press releases. When data were unavailable online, we contacted pharmaceutical companies directly using telephone and email contact information posted on ClinicalTrials.gov. For the BAN2401 trial (NCT01767311, 15), we obtained the data from the study contact at Eisai pharmaceuticals. For the two trials of Aducanumab, EMERGE and ENGAGE (NCT02484547 and NCT02477800, respectively), a press release was available in a portable document format with raw data (standard errors were estimated from graphs). This analysis is based on all data available to us on December 20, 2019.

Data Extraction and Synthesis

Standard meta-analysis methods cannot be applied here as we are estimating the effect of change in amyloid on change in cognition across heterogeneous drug treatments. We derived a maximum-likelihood estimator to estimate the effect of amyloid reduction on cognitive change. Based on the principles of instrumental variable analyses, we made the following standard assumptions (14): randomization to drug treatment plausibly affected the change in amyloid, randomization is independent of plausible confounders such as APOE-ε4, and randomization to drug treatment does not affect cognition through mechanisms other than amyloid reduction.
Additionally, we assumed that change in cognition due to drug treatment to be proportional to the change in amyloid resulting from drug treatment, following a linear dose-response association with change in amyloid. This approach allowed us to combine results for trials with different durations of follow-up, regardless of the duration of follow-up. In pooling results across trials, we assume that the effect of reducing amyloid on cognition does not vary with mechanism by which amyloid-β is targeted, i.e. by drug. We did not account for the covariance between measured mean change in cognition and measured mean change in SUVr as this information was not reported.

Since randomization depended on APOE-ε4 carrier status in the trial of BAN2401 (adaptive randomization altered mid-trial to exclude carriers from the highest treatment group), we required a second estimator for the effect of change in amyloid adjusting for the proportion of APOE-ε4 carriers in each group, because APOE-ε4 carrier status is known to affect change in cognition. For this trial, since data were collected at two time points, we additionally assumed that mean change in amyloid was linear with respect to time. See the directed acyclic graphs (16) in the Appendix for further details.

All statistical analysis was performed in R version 3.6.1. The likelihood for each trial was the product of probabilities of the observed change in cognition for each arm conditional on an intercept (the change in cognition associated with no change in SUVr) and a slope (the effect of change in SUVr on cognition). To obtain pooled estimates, these likelihoods were then multiplied assuming one common slope and a trial-specific intercept. Minimization of the negative log-likelihood was performed to obtain maximum-likelihood estimates for slopes and intercepts. Standard errors were obtained using observed Fisher information (17). Derivations of the maximum likelihood estimators and R code are given in the Appendix.

We obtained estimates pooled by drug and overall of the effect of change in amyloid on change in cognition. Because the populations enrolled and time in followup in each trial may differ, we allowed each trial to have its own intercept, which gives the expected change in cognition with no change in SUVr for participants in that trial over the followup period. We performed sensitivity analyses restricting to antibody drugs and with and without the unpublished trials of BAN2401 and Aducanumab. For the sake of comparison, we use a similar maximum-likelihood estimation procedure to estimate the effect of APOE-ε4 carriage on annual change in cognition.

Because evidence from new trials or updates of existing trials are frequently released, we provide a web-based interface at https://amyloidintegratingevidence.shinyapps.io/application/, which allows for a recalculation of results either with updated data or new data that may become available.
Figure 1: Flow chart of trial exclusions and inclusions. Our search identified 57 drugs studied in 2822 trials registered on ClinicalTrials.gov; 196 of these studies were in populations with AD and/or MCI and were “Active, not recruiting,” “Terminated,” or “Completed,” and 34 trials appeared likely to meet the inclusion criteria based outcomes listed on ClinicalTrials.gov. We were able to obtain data for 14 (15 clinical trial numbers) of these 34 trials.
Results

As shown in Figure 1, our search of AlzForum.org identified 57 drugs studied in 2822 trials registered on ClinicalTrials.gov; 196 of these studies were in populations with AD and/or MCI and were “Active, not recruiting,” “Terminated,” or “Completed.” Two authors reviewed the 196 trials and 34 trials appeared likely to meet the inclusion criteria based outcomes listed on ClinicalTrials.gov. We were able to obtain data for 14 (15 clinical trial numbers) of these 34 trials, and called or emailed pharmaceutical companies for data from the other 20 trials, but were unable to obtain additional data for the following reasons: the trial was terminated (and thus no additional data was available) (2), the trial was not yet completed (2), change in SUVr was not measured (1), the data could not be converted to electronic format to be shared (1), the PI retired and there was no other contact (2), an extensive proposal was required without guarantee the data we needed was available (1), or we did not get a response (11). The 14 trials for which we obtained publicly available data tested the following drugs: Bexarotene, Solanezumab (3 trials, 18), LY450139 (2 trials), Gantenerumab, Bapineuzumab (2 trials, 3 clinical trial numbers: 12, 19, 20, 21), Verubecestat (2 trials), BAN2401 (13, 22, 15), and Aducanumab (23). All but three of these trials reported change in MMSE as a cognitive outcome; one trial of LY450139 (NCT00762411) reported ADAS-Cog11, one trial of Verubecestat (MK-8931) (NCT01953601), and the trial of BAN2401 trial reported ADCOMS. To provide a common scale, changes in these cognitive measures were converted to change in MMSE points (see Appendix 5 for further details) using a crosswalk (24, 25). When pooling trials of Solanezumab, only two independent trials were used, since the third trial was an extension of an earlier trial. Table S1 in the appendix gives summaries of the trials included in the aggregated analysis and where data were obtained.

Figure 2 shows the estimated effect of a 0.1 reduction in SUVr on change in MMSE for 8 drugs and 14 trials, with 95% confidence intervals (CIs). Estimates from all trials are consistent with no effect of changing amyloid on performance on the MMSE. These estimates combine information based on both the effect of randomization on change in SUVr and the effect of randomization on change in MMSE. Pooling all 14 trials, the estimated effect of a 0.1 unit reduction in amyloid SUVr on MMSE is 0.02, 95% CI: (0.05, 0.09) points. Pooling by drug type (antibody and small-molecule) also yield results consistent with the null.

To contextualize the effect estimates, we compare them to the effect of APOE-ε4 carriage on annual cognitive decline. We used trials 2a and 2b of bapineuzumab and the trial of BAN2401 to estimate the effect of APOE-ε4 carriage on annual cognitive change. The estimated effect on annual change in MMSE score of APOE-ε4 carriage from trials 2a and 2b of bapineuzumab is -0.7, 95% CI: (-1.09, -0.31) and from the trial of BAN2401 -1.33, 95% CI: (-1.99, -0.67) (table S3).
had approximately a 0.3 reduction in SUVr over 18 months, one of the largest changes in SUVr seen across these trials. Such a reduction in SUVr would correspond to a difference in change in MMSE of 0.054, 95% CI: (-0.16, 0.27), expected to be significantly smaller in magnitude than the annual effect of APOE-ε4 carriage.
Figure 2: Forest plot of the estimated effects (95% confidence intervals) of an 0.1 decrease in SUVr on MMSE. A) for each trial and drug and B) pooled across all drugs and by drug type. The trials of BAN2401 and Aducanumab are unpublished and were excluded from the “All Published Antibody” category. Estimates pooled across drugs are shown as diamonds, where the center reflects the pooled estimate and the width gives the 95% CI. A numbered key is given for multiple trials of the same drug (see table S1 for clinical trial numbers).
Discussion

Aggregated results from 40 randomization arms in 14 trials indicate that reducing amyloid with drug treatment has, at most, a small effect on cognition as measured with the MMSE. Findings from each of these trials individually were consistent with no effect of change in amyloid on cognitive decline, but point estimates were inconsistent and usually had wide confidence intervals. Even the largest potential benefit consistent with the upper bound of the 95% confidence interval of our pool estimates, were small compared to the annual effect of APOE-ε4 on decline in MMSE.

Chance findings of a statistically significant beneficial effect for any pharmaceutical become likely as more and more trials are conducted. 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, there is an 82% chance. It is therefore critical to consider pooled evidence from all available trials to contextualize results for any single trial.

This is the first report combining all results into a single estimate of the effect of amyloid reduction on cognitive change. In addition to obtaining data from published trials or posted on ClinicalTrials.gov, we obtained data from two recent unpublished trials reporting significant benefit of amyloid reduction on cognitive decline: BAN2401 has been reported to be effective (AAIC 2018), as well as Aducanumab at higher doses in a reanalysis of terminated studies (26).

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. There are some caveats to this interpretation. First, we assumed that changes in cognition due to randomization to drug treatment were fully mediated by reductions in amyloid-β. However, it is plausible that reductions in amyloid-β improve cognitive outcomes, but that the drugs evaluated harmed cognition via other mechanisms. We cannot rule out this possibility, but given the precision of our null effect estimate, the adverse direct effects would need to precisely counterbalance the positive effects of amyloid reduction, which is unlikely.

Second, we focused on the MMSE because it was reported by nearly all trials, but MMSE is known to have low sensitivity to cognitive deterioration in cognitively normal adults. However, nearly all of these studies showed an average deterioration in MMSE, indicating that it may be reasonably sensitive to changes in cognition in the trial populations.

The pooled result across all trials is only relevant if we assumed that the effect of reducing amyloid on cognition does not vary with mechanism by which amyloid-β is targeted, i.e. by drug. We therefore also stratified our results by antibody and non-antibody drugs, with and without the recent unpublished trials of...
BAN2401 and Aducanumab. While we cannot rule out the possibility that targeting amyloid-β by a specific mechanism will slow cognitive change, our results are consistent with no cognitive benefit of any amyloid reducing drug treatment.

We were also unable to obtain data for 20 trials that met our eligibility criteria after contacting pharmaceutical companies directly. Lack of access to full results was an important limitation of our analysis, as only one company responded to requests for additional data (BAN2401, Eisai pharmaceuticals). However, only 7 of the 20 trials not included in this analysis were “completed”; 8 were “active, not recruiting” and 5 were “terminated.” Terminated trials may have been terminated due to futility and thus would be unlikely to successfully slow cognitive decline. Data for the “active, not recruiting” trials may become available at a future date.

Without full reports, another limitation of this study is possible error in the input data. Other crosswalks between ADCOMS and MMSE and ADAS-Cog11 and MMSE 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 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.

We did not account for the covariance between measured cognition and measured SUVr since this information was not reported. This covariance is negligible if error in measured SUVr is large compared to the variance in true 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, then this covariance term is relatively negligible. Both of these are plausible assumptions because amyloid-PET produces noisy measurements in SUVr (27) and variance in amyloid most plausibly accounts for only a minority of the variance in cognition across individuals (e.g. 28, 29). More details are given in Appendix 3. If estimated covariances between measured SUVr and measured cognition become available, such information could be easily incorporated into the proposed estimation procedure.

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. We could not evaluate this without long-term cognitive follow-up results from each trial. If this is the case, however, amyloid reduction trials would need to substantially extend the typical follow-up period to detect any benefit. In conclusion, these results provide evidence that amyloid reduction alone is unlikely to substantially slow
cognitive decline within the follow-up period of most current trials.
Patient and Public Involvement

This research was done without patient involvement. Patients were not invited to comment on the study design and were not consulted to develop patient relevant outcomes or interpret the results. Patients were not invited to contribute to the writing or editing of this document for readability or accuracy.

Dissemination of results to study participants is not applicable.

Contributions Statement

SFA, MCP, and MMG conceived of the project. MMG supervised the work. SFA, MMG, ETT, SCZ discussed and planned the analysis and contributed to the supplemental information. SFA collected the data from online sources and SCZ checked this work. SFA and WDB independently reviewed trials for meeting inclusion criteria. ALG contacted pharmaceutical companies by phone and email to attempt to obtain additional data. SFA performed the analysis. SFA, MCP, and MMG contributed to data visualization. SZ reviewed all code and built the web application. AMB, JJM, ERM, TJF, and FE contributed to the interpretation of results and writing of the discussion section. All authors discussed the results and contributed to the writing and editing of manuscript.

Funding

NIH NIA, grant number R01AG057869.

Ethics Committee Approval

No individual-level data is included in this manuscript. All data is aggregated data from clinical trials that is publicly available, with the exception of summary data for the trial of BAN2401 obtained from Esai pharmaceuticals.
Appendices
Appendix 1: Data Sources

1. Bexarotene NCT01782742: 30
2. Solanezumab NCT00904683 & NCT01127633: 31
5. LY450139 NCT00762411: https://clinicaltrials.gov/ct2/show/results/NCT00762411
6. LY450139 NCT00594568: https://clinicaltrials.gov/ct2/show/results/NCT00594568
7. Gantenerumab NCT01224106: 32
8. Bapineuzumab NCT00575055: 21
9-10. Bapineuzumab NCT00676143, NCT00667810: 19
13. BAN2401 (NCT01767311): presented at the Alzheimer’s Association International Conference, 2018; summary data from Eisai pharmaceuticals
15. Aducanumab (ENGAGE) NCT02477800: Press Release
## Data sources for trials

<table>
<thead>
<tr>
<th>Clinical Trial Number</th>
<th>Drug(-key)</th>
<th>Drug Classification</th>
<th>Number of Treatment Arms</th>
<th>Number with Cognitive Assessment</th>
<th>Number with Amyloid PET</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCT01782742</td>
<td>Bexarotene Small</td>
<td>molecule</td>
<td>2</td>
<td>20</td>
<td>20</td>
<td>30</td>
</tr>
<tr>
<td>NCT00904683</td>
<td>Solanezumab-1 &amp; 2</td>
<td>Antibody</td>
<td>2</td>
<td>251</td>
<td>1322</td>
<td>31</td>
</tr>
<tr>
<td>NCT01127633</td>
<td>Solanezumab-1 &amp; 2</td>
<td>Antibody</td>
<td>2</td>
<td>860</td>
<td>90</td>
<td>clinicaltrials.gov</td>
</tr>
<tr>
<td>NCT01900665</td>
<td>Solanezumab-3</td>
<td>Antibody</td>
<td>2</td>
<td>1769</td>
<td>1596</td>
<td>clinicaltrials.gov</td>
</tr>
<tr>
<td>NCT00762411</td>
<td>LY450139-1 Small</td>
<td>molecule</td>
<td>2</td>
<td>1108</td>
<td>1108</td>
<td>clinicaltrials.gov</td>
</tr>
<tr>
<td>NCT00594568</td>
<td>LY450139-2 Small</td>
<td>molecule</td>
<td>3</td>
<td>939</td>
<td>125</td>
<td>clinicaltrials.gov</td>
</tr>
<tr>
<td>NCT01224106</td>
<td>Gantenerumab Antibody</td>
<td>3</td>
<td>797</td>
<td>55</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>NCT00575055</td>
<td>Bapinezumab-1 Antibody</td>
<td>1</td>
<td>26</td>
<td>26</td>
<td>26</td>
<td>21</td>
</tr>
<tr>
<td>NCT00676143</td>
<td>Bapinezumab-2 Antibody</td>
<td>2a</td>
<td>1090</td>
<td>115</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>NCT00667810</td>
<td>Bapinezumab-3 Antibody</td>
<td>2b</td>
<td>3</td>
<td>1114</td>
<td>39</td>
<td>19</td>
</tr>
<tr>
<td>NCT01739348</td>
<td>Verubecestat-1 Small</td>
<td>molecule</td>
<td>3</td>
<td>1838</td>
<td>44</td>
<td>clinicaltrials.gov</td>
</tr>
</tbody>
</table>

18
<table>
<thead>
<tr>
<th>Clinical Trial Number</th>
<th>Drug(-key)</th>
<th>Drug Classification</th>
<th>Number of Treatment Arms</th>
<th>Number with Cognitive Assessment</th>
<th>Number with Amyloid PET</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCT01953601</td>
<td>Verubecestat-2</td>
<td>Small molecule</td>
<td>3</td>
<td>1392</td>
<td>187</td>
<td>clinicaltrials.gov</td>
</tr>
<tr>
<td>NCT01767311</td>
<td>BAN2401 Antibody</td>
<td>5</td>
<td>854</td>
<td>306</td>
<td>Alzheimer’s Association International Conference, 2018</td>
<td></td>
</tr>
<tr>
<td>NCT02484547</td>
<td>Aducanumab-1 Antibody</td>
<td>3</td>
<td>879</td>
<td>317</td>
<td>Press Release</td>
<td></td>
</tr>
<tr>
<td>NCT01953601</td>
<td>Aducanumab-2 Antibody</td>
<td>3</td>
<td>923</td>
<td>317</td>
<td>Press Release</td>
<td></td>
</tr>
</tbody>
</table>

*Table S1:* Description of trials included in the aggregated analysis. A numbered key is given for multiple trials of the same drug.
Appendix 2: Effect of APOE-ε4

<table>
<thead>
<tr>
<th>Drug</th>
<th>Effect of APOE-ε4 on MMSE Change per Year, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAN2401</td>
<td>-1.33, (-1.99, -0.67)</td>
</tr>
<tr>
<td>Bapineuzumab</td>
<td>-0.70, (-1.09, -0.31)</td>
</tr>
</tbody>
</table>

Table S2: The estimated effect of APOE-ε4 on annual change in MMSE. All values are reported to two decimal places.
Appendix 3: Derivation of the Maximum Likelihood Estimator

A

\[ Z \rightarrow \Delta \rightarrow Y \]

\[ U \]

B

\[ Z \rightarrow \Delta \rightarrow Y \]

\[ U \]

\[ \epsilon \]

Figure S1: DAGs representing experimental design for all trials except the trial of BAN2401 (A) and the BAN2401 trial (NCT01767311) (B). A: Nodes represent the following variables: \( Z \): randomization arm; \( \Delta \): amyloid change; \( Y \): cognitive change; \( U \): disease severity or other shared causes of amyloid and cognition change. B: Nodes represent the following variables: \( Z \): randomization arm (one of 6 possible groups); \( \Delta \): amyloid change; \( Y \): cognitive change; \( U \): disease severity or other shared causes of amyloid and cognition change; \( \epsilon \): APOE-\( \varepsilon 4 \) status, positive or negative.
Figure S2: DAGs representing experimental design for non-BAN2401 trials with other sources of variance for amyloid reduction ($\gamma_\delta$) and cognition ($\gamma_y$), as well as measurement error for amyloid ($\varepsilon_\delta$) and cognition ($\varepsilon_y$).
We assume the causal structure in figure S2, with the following random variables: $Z$ is the instrumental variable, $\Delta$ is the endogenous variable, $U$ is an unmeasured confounder of amyloid accumulation and cognition, and $Y$ is the outcome. Lowercase $z$, $\delta$, $u$, and $y$ are used to represent specific values these random variables.

We assume the $Y$, $\Delta$, and $U$ are continuous random variables. Similar results can be obtained assuming discrete random variables by replacing integrals with sums over the support of these variables. While we assume only one unmeasured confounder $U$, these results can easily be extended for multiple unmeasured confounders. We consider the possibility that measures of the endogenous and outcome variables are only available for a subsample of individuals within each study arm $N_{z,\delta}$ and $N_{z,y}$ (as is often the case when the endogenous variable is an expensive biomarker).

We begin with a standard IV linearity assumption: we assume $E[Y|\Delta = \delta, U = u] = \alpha \delta + \beta u$, where $\alpha$ is the parameter of interest, giving the effect of $\Delta$ on $Y$. Since treatment $z$ is randomly assigned and only affects $Y$ through $\Delta$, we can write:

$$E[Y|\Delta = \delta, U = u, Z = z] = \alpha \delta + \beta u.$$  \hspace{1cm} (1)

We obtain the unconditional expectation for a given treatment arm $z$ by integrating the conditional expectation over the support of $u$ and $\delta$, multiplied by the probability densities of $u$ and $\delta$:

$$E[Y|Z = z] = \int_{u,\delta} (\alpha \delta + \beta u) \Pr(U = u) \Pr(\Delta = \delta|Z = z) d\delta du$$  \hspace{1cm} (2)

The integrand is absolutely integrable, so we can expand and change the order of integration, which gives:

$$E[Y|Z = z] = \int_u \Pr(U = u) du \int_\delta \alpha \delta \Pr(\Delta = \delta|Z = z) d\delta + \int_u \beta u \Pr(U = u) du \int_\delta \Pr(\Delta = \delta|Z = z) d\delta$$  \hspace{1cm} (3)

Using the fact that $\int_u \Pr(U = u) du = \int_\delta \Pr(\Delta = \delta|Z = z) d\delta = 1$, we obtain:

$$E[Y|Z = z] = \int_\delta \alpha \delta \Pr(\Delta = \delta|Z = z) d\delta + \int_u \beta u \Pr(U = u) du$$  \hspace{1cm} (4)

Finally, using the definition of the expected value:

$$E[Y|Z = z] = \alpha E[\Delta = \delta|Z = z] + \beta E[U].$$  \hspace{1cm} (5)

$\beta E[U]$ is a constant so, we let $\beta E[U] = b_0$. We note that $b_0$ may be study dependent since the distribution of confounders may vary with study population. We then have
\[ E[Y|Z=z] = \alpha E[\Delta = \delta|Z=z] + b_0 \] (6)

Therefore, the expected value of \( Y \) changes linearly with the expected value of \( \Delta \) within randomization arms \( z \) of a study.

Let

\[ \mu_z = E[Y|Z=z] \]

and

\[ \theta_z = E[\Delta|Z=z]. \]

We measure sample means \( \hat{\mu}_z \) and \( \hat{\theta}_z \), with sample standard errors \( \hat{\sigma}_z \) and \( \hat{\rho}_z \), respectively. As sample size increases, the distributions of \( \hat{\mu}_z \) and \( \hat{\theta}_z \) approach normal distributions with standard deviations given by the unmeasured population standard errors \( \sigma_z \) and \( \rho_z \). If \( Y \) and \( \Delta \) are normally distributed, we do not have to consider this a large-sample approximation. Therefore, \( \alpha \hat{\theta}_z + b_0 \), the value of \( \hat{\mu}_z \) we would predict based on \( \hat{\theta}_z \), also approaches a normal distribution with mean \( \alpha \hat{\theta}_z + b_0 \) and standard deviation \( \alpha \rho_z \). The discrepancy between \( \hat{\mu}_z \) and \( \alpha \hat{\theta}_z + b_0 \) is a measure of how far values of \( \hat{\mu}_z \) based on \( \hat{\theta}_z \) are from \( \hat{\mu}_z \), and this discrepancy will be expected to be smaller the better estimates of the parameters \( \alpha \) and \( b_0 \). Therefore, based on a Welch’s two-sample \( t \)-test, we can write the likelihood function of the parameters \( \alpha \) and \( b_0 \) given the observed data \( D \) (\( \hat{\theta}_z \), \( \hat{\mu}_z \), \( \hat{\sigma}_z \), and \( \hat{\rho}_z \) for each treatment level \( z \)) as:

\[
L(\alpha; b_0|D) = \prod_z f \left( \frac{\alpha \hat{\theta}_z + b_0 - \hat{\mu}_z}{\sqrt{\sigma_z^2 + \alpha^2 \rho_z^2}} \right) \] (7)

where \( f \) is a \( t \)-distribution with mean zero and standard deviation 1, with degrees of freedom given by:

\[
\frac{(\sigma_z^2 + \alpha^2 \rho_z^2)^2}{\sigma_z^4 + \frac{\alpha^4 \rho_z^4}{N_z,\sigma^4} + \frac{\alpha^2 \rho_z^2}{N_z,\delta^2} - 1}. \] (8)

We note that in equation 7 we are overestimating the variance of

\[ \alpha \hat{\theta}_z + b_0 - \hat{\mu}_z. \]

Specifically, the correct variance is obtained from the variance of the difference of two non-independent random variables multiplied by constant coefficients and is given by:

\[
\frac{(\sigma_z^2 + \alpha^2 \rho_z^2)^2}{\sigma_z^4 + \frac{\alpha^4 \rho_z^4}{N_z,\sigma^4} + \frac{\alpha^2 \rho_z^2}{N_z,\delta^2} - 1}.
\]
\[
\text{Var}[\alpha \hat{\theta}_z + b_0 - \hat{\mu}_z] = \sigma_z^2 + \alpha^2 \rho_z^2 - 2\alpha \text{Cov}(\hat{\theta}_z, \hat{\mu}_z)
\] (9)

With only aggregated data, we do not know, nor can we estimate, the value of \( \text{Cov}(\sigma_z, \rho_z) \). However, it is possible to argue that the covariance is small compared with the other terms in this expression.

Specifically,
\[
\hat{\Delta} = \beta_1 U + \zeta + \gamma_\delta + \epsilon_\delta
\] (10)

where \( \beta_1 \) is the effect of \( U \) on \( \Delta \) and \( \zeta \) is the effect of randomization on \( \Delta \).

Similarly,
\[
\hat{Y} = \alpha \Delta + \beta_2 U + \gamma_y + \epsilon_y
\] (11)

where \( \beta_2 \) is the effect of \( U \) on \( \Delta \). We can rewrite this as:
\[
\hat{Y} = \alpha \hat{\Delta} - \alpha \epsilon_\delta + \beta_2 U + \gamma_y + \epsilon_y
\] (12)

Now, we can write an expression for the covariance between \( \hat{Y} \) and \( \hat{\Delta} \):
\[
\text{Cov}(\hat{Y}, \hat{\Delta}) = \text{Cov}(\alpha \hat{\Delta}, \hat{\Delta}) - \text{Cov}(\alpha \epsilon_\delta, \hat{\Delta}) + \text{Cov}(\beta_2 U, \hat{\Delta}) + \text{Cov}(\gamma_y, \hat{\Delta}) + \text{Cov}(\epsilon_y, \hat{\Delta})
\] (13)

Simplifying,
\[
\text{Cov}(\hat{Y}, \hat{\Delta}) = \alpha \text{Cov}(\hat{\Delta}, \hat{\Delta}) - \alpha \text{Cov}(\epsilon_\delta, \hat{\Delta}) + \beta_1 \beta_2 \text{Cov}(U, U)
\] (14)

\[
\text{Cov}(\hat{Y}, \hat{\Delta}) = \alpha \text{Cov}(\Delta + \epsilon_\delta, \Delta + \epsilon_\delta) - \alpha \text{Cov}(\epsilon_\delta, \epsilon_\delta) + \beta_1 \beta_2 \text{Cov}(U, U)
\] (15)

\[
\text{Cov}(\hat{Y}, \hat{\Delta}) = \alpha \text{Cov}(\Delta, \Delta) + \alpha \text{Cov}(\epsilon_\delta, \epsilon_\delta) - \alpha \text{Cov}(\epsilon_\delta, \epsilon_\delta) + \beta_1 \beta_2 \text{Cov}(U, U)
\] (16)

\[
\text{Cov}(\hat{Y}, \hat{\Delta}) = \alpha \text{Cov}(\Delta, \Delta) + \beta_1 \beta_2 \text{Cov}(U, U)
\] (17)
Where,

\[
\text{Cov}(\sigma_z, \rho_z) = \text{Cov}(\frac{\hat{Y}}{\sqrt{N_y}}, \frac{\hat{\Delta}}{\sqrt{N_{\delta,y}}}) = \frac{1}{\sqrt{N_{\delta,y}N_{\delta,\delta}}} \text{Cov}(\hat{Y}, \hat{\Delta}) \tag{18}
\]

Therefore, if error in measured SUVr is large compared to the variance in true SUVr and the non-\(U\) predictors of cognition (e.g. education, vascular risk factors, and other non-amyloid pathologies) account for more variance in cognition than \(U\), then we can argue that this covariance term is small compared with the standard error in mean cognition and the standard error in measured SUVr.
Appendix 4: Analyses of the BAN2401 trial

BAN2401 targets protofibrils of amyloid-β and, in an ongoing trial, has produced statistically significant reductions in both amyloid-β in the brain as well as reductions in cognitive decline, according to a recent press release (AAIC 2018). However, there is a highly variable proportion of APOE-ε4 carriers across treatment arms, ranging from 30% to 91%. After the start of the trial it was determined that APOE-ε4 carriers had an unacceptably high risk of amyloid-related imaging abnormalities, specifically edema (ARIA-E). As a result, no additional APOE-ε4 carriers were assigned to this group. This prompted concerns that differences in cognitive outcomes between groups is due depletion of APOE-ε4 carriers in the highest treatment group and not drug efficacy. In a recent press release, Eisai pharmaceuticals stated their results were robust to adjustment for differences in APOE-ε4 carrier status across treatment arms. They additionally stated that this trial was the first of its kind to provide evidence for the amyloid cascade hypothesis since BAN2401 reduced-β amyloid and was associated with slower cognitive decline.

It was assumed that changes in ADCOMS in the BAN2401 trial were linear with respect to changes in MMSE and a conversion was derived using the following data (derived from table 4 in 24). Differential weights are used to account for differences in sample size.

<table>
<thead>
<tr>
<th>Change in ADCOMS</th>
<th>Change in MMSE</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.027</td>
<td>0.51</td>
<td>276.40</td>
</tr>
<tr>
<td>0.057</td>
<td>1.58</td>
<td>102.60</td>
</tr>
<tr>
<td>0.082</td>
<td>1.24</td>
<td>90.43</td>
</tr>
<tr>
<td>0.002</td>
<td>0.12</td>
<td>294.20</td>
</tr>
</tbody>
</table>
Appendix 5: Conversion of ADAS-Cog11 & CDR-SB to MMSE

ClinicalTrials.gov Identifier: NCT00762411

ADAS-Cog11 to MMSE

We use the crosswalk between the ADAS-Cog11 and MMSE in table 1 of 25.

We fit a logistic curve with an upper bound at 30 to obtain a function that predicts a mean MMSE score as function of mean ADAS-Cog score. We then calculate the derivative of this function. We use the following midpoint approximation to calculate how the change in ADAS-Cog with respect to changing amyloid is related to the change in MMSE with respect to changing amyloid:

\[
\text{Mean change in MMSE} = \left( \frac{d\text{MMSE}}{d\text{ADAS-Cog}} \right)_{(\text{Starting ADAS-Cog} + \text{Change in ADAS-Cog})/2} \text{Mean change in ADAS-Cog}
\]

We plot the crosswalk between between ADAS-Cog11 and MMSE (table 1 of 25) and the fitted model (top) and the derivative of the fitted model (bottom).
ClinicalTrials.gov Identifier: NCT01953601

CDR-SB to MMSE

We use the crosswalk between the CDR-SB and MMSE in table 1 of 25 and an identical procedure as above.

We plot the crosswalk between CDR-SB and MMSE (table 1 of 25) and the fitted model (top) and the derivative of the fitted model (bottom).
R Code

- data.R
- analysis.R
References


## PRISMA 2009 Checklist

<table>
<thead>
<tr>
<th>Section/topic</th>
<th>#</th>
<th>Checklist item</th>
<th>Reported on page #</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TITLE</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>1</td>
<td>Identify the report as a systematic review, meta-analysis, or both.</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>See cover letter: This is not a traditional meta-analysis</em></td>
<td></td>
</tr>
<tr>
<td><strong>ABSTRACT</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structured summary</td>
<td>2</td>
<td>Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.</td>
<td>2-3</td>
</tr>
<tr>
<td><strong>INTRODUCTION</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rationale</td>
<td>3</td>
<td>Describe the rationale for the review in the context of what is already known.</td>
<td>4</td>
</tr>
<tr>
<td>Objectives</td>
<td>4</td>
<td>Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).</td>
<td>4-5</td>
</tr>
<tr>
<td><strong>METHODS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protocol and registration</td>
<td>5</td>
<td>Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.</td>
<td>NA</td>
</tr>
<tr>
<td>Eligibility criteria</td>
<td>6</td>
<td>Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.</td>
<td>7-8</td>
</tr>
<tr>
<td>Information sources</td>
<td>7</td>
<td>Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.</td>
<td>7-8</td>
</tr>
<tr>
<td>Search</td>
<td>8</td>
<td>Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.</td>
<td>7-8</td>
</tr>
<tr>
<td>Study selection</td>
<td>9</td>
<td>State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).</td>
<td>7-8</td>
</tr>
<tr>
<td>Data collection process</td>
<td>10</td>
<td>Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.</td>
<td>7-8</td>
</tr>
<tr>
<td>Data items</td>
<td>11</td>
<td>List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.</td>
<td>7-8</td>
</tr>
<tr>
<td>Risk of bias in individual studies</td>
<td>12</td>
<td>Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.</td>
<td>NA</td>
</tr>
<tr>
<td>Summary measures</td>
<td>13</td>
<td>State the principal summary measures (e.g., risk ratio, difference in means).</td>
<td>7-8</td>
</tr>
</tbody>
</table>
### PRISMA 2009 Checklist

**Synthesis of results**

Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., $I^2$) for each meta-analysis.

Page 7-8

<table>
<thead>
<tr>
<th>Section/topic</th>
<th>#</th>
<th>Checklist item</th>
<th>Reported on page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk of bias across studies</td>
<td>15</td>
<td>Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).</td>
<td>NA</td>
</tr>
<tr>
<td>Additional analyses</td>
<td>16</td>
<td>Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.</td>
<td>NA</td>
</tr>
</tbody>
</table>

### RESULTS

**Study selection**

Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.

Page 9

<table>
<thead>
<tr>
<th>Section/topic</th>
<th>#</th>
<th>Checklist item</th>
<th>Reported on page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study selection</td>
<td>17</td>
<td>Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.</td>
<td>9</td>
</tr>
<tr>
<td>Study characteristics</td>
<td>18</td>
<td>For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.</td>
<td>Appendix</td>
</tr>
<tr>
<td>Risk of bias within studies</td>
<td>19</td>
<td>Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).</td>
<td>NA</td>
</tr>
<tr>
<td>Results of individual studies</td>
<td>20</td>
<td>For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.</td>
<td>Figure 2</td>
</tr>
<tr>
<td>Synthesis of results</td>
<td>21</td>
<td>Present results of each meta-analysis done, including confidence intervals and measures of consistency.</td>
<td>Figure 2</td>
</tr>
<tr>
<td>Risk of bias across studies</td>
<td>22</td>
<td>Present results of any assessment of risk of bias across studies (see Item 15).</td>
<td>NA</td>
</tr>
<tr>
<td>Additional analysis</td>
<td>23</td>
<td>Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).</td>
<td>NA</td>
</tr>
</tbody>
</table>

### DISCUSSION

**Summary of evidence**

Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).

Page 12-13

<table>
<thead>
<tr>
<th>Section/topic</th>
<th>#</th>
<th>Checklist item</th>
<th>Reported on page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary of evidence</td>
<td>24</td>
<td>Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).</td>
<td>12-13</td>
</tr>
<tr>
<td>Limitations</td>
<td>25</td>
<td>Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).</td>
<td>12-13</td>
</tr>
<tr>
<td>Conclusions</td>
<td>26</td>
<td>Provide a general interpretation of the results in the context of other evidence, and implications for future research.</td>
<td>12-13</td>
</tr>
</tbody>
</table>

### FUNDING

**Funding**

Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.

Page NA

<table>
<thead>
<tr>
<th>Section/topic</th>
<th>#</th>
<th>Checklist item</th>
<th>Reported on page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Funding</td>
<td>27</td>
<td>Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.</td>
<td>NA</td>
</tr>
</tbody>
</table>
PRISMA 2009 Checklist

For more information, visit: www.prisma-statement.org.