Received: 15 December 2023 | Revised: 13 March 2024 | Accepted: 16 April 2024

DOI: 10.1002/trc2.12485

Alzheimer's & Dementia
Translational Research & Clinical Interventions

PERSPECTIVE

Pooling Alzheimer's disease clinical trial data to develop personalized medicine approaches is easier said than done: A proof-of-principle study and call to action

Mark A. Dubbelman[1,2,3,4] | **Eleonora M. Vromen**[1,2] | **Betty M. Tijms**[1,2] | **Johannes Berkhof**[5] | **Lois Ottenhoff**[1,2,6] | **Everard G. B. Vijverberg**[1,2] | **Niels D. Prins**[6] | **Wiesje M. van der Flier**[1,2,5] | **Sietske A. M. Sikkes**[1,2,7]

[1]Alzheimer Center Amsterdam, Department of Neurology, Vrije Universiteit Amsterdam, Amsterdam UMC, Amsterdam, The Netherlands

[2]Amsterdam Neuroscience, Neurodegeneration, Amsterdam, The Netherlands

[3]Center for Alzheimer Research and Treatment, Department of Neurology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA

[4]Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA

[5]Department of Epidemiology & Data Science, Vrije Universiteit Amsterdam, Amsterdam UMC, Amsterdam, The Netherlands

[6]Brain Research Center, Amsterdam, The Netherlands

[7]Department of Clinical, Neuro and Developmental Psychology, Faculty of Behavioral and Movement Sciences, VU University, Amsterdam, The Netherlands

Correspondence
Sietske A. M. Sikkes, Alzheimer Center Amsterdam, Department of Neurology, Amsterdam UMC location VUmc, P.O. Box 7057, 1007 MB, Amsterdam, The Netherlands.
Email: s.sikkes@amsterdamumc.nl

Funding information
Health~Holland, Topsector Life Sciences & Health, Grant/Award Number: PPP-allowance #LSHM19051; Alzheimer Nederland, Grant/Award Number: WE.06-2023-02; ZonMW, Grant/Award Number: #73305095007; Topsector Life Sciences & Health, Grant/Award Number: #LSHM20106; Health Holland, Grant/Award Numbers: LSHM19051, LSHM20084, LSHM22026-SGF; Zon-MW, Grant/Award Numbers: #7330502051, #73305095008

Abstract

With the advent of the first generation of disease-modifying treatments for Alzheimer's disease, it is clearer now more than ever that the field needs to move toward personalized medicine. Pooling data from past trials may help identify subgroups most likely to benefit from specific treatments and thus inform future trial design. In this perspective, we report on our effort to pool data from past Alzheimer's disease trials to identify patients most likely to respond to different treatments. We delineate challenges and hurdles, from our proof-of-principle study, for which we requested access to trial datasets from various pharmaceutical companies and encountered obstacles in the process of arranging data-sharing agreements through legal departments. Six phase I–III trials from three sponsors provided access to their data (total $n = 3170$), which included demographic information, vital signs, primary and secondary endpoints, and in a small subset, cerebrospinal fluid amyloid ($n = 165$, 5.2%) and tau ($n = 212$, 6.7%). Data could be analyzed only within specific data access platforms, limiting potential harmonization with data provided through other platforms. Limited overlap in terms of outcome measures, clinical and biological information hindered analyses. Thus, while it is a commendable advancement that (some) trials now allow researchers to study their data, we conclude that gaining access to past trial datasets is

Mark A. Dubbelman and Eleonora M. Vromen contributed equally and shared the first authorship.

complicated, frustrating the field's communal effort to find the best treatments for the right individuals. We provide a plea to promote harmonization and open access to data, by urging trial sponsors and the academic research community alike to remove barriers to data access and improve collaboration through practicing open science and harmonizing outcome measures, to allow investigators to learn all there is to learn from past failures and successes.

KEYWORDS
Alzheimer's disease, amyloid, biomarkers, clinical trials, drug development, pharmaceutical companies, tau

HIGHLIGHTS

- Pooling data from past Alzheimer's disease clinical trials may help identify subgroups most likely to benefit from specific treatments and may help inform future trial design.
- Accessing past trial datasets is complicated, frustrating the field's communal effort to find the best treatments for the right individuals.
- We urge trial sponsors and the academic research community to remove data access barriers and improve collaboration through practicing open science and harmonizing outcome measures.

1 | INTRODUCTION

Alzheimer's disease is the most common cause of dementia,[1] affecting millions of people worldwide. As of 2023, over 180 clinical trials are actively investigating more than 140 unique agents of which the vast majority are disease-modifying.[2] Recent trials have shown promising effects on various pathophysiologic mechanisms.[3–6] In addition, monoclonal antibodies against amyloid-β are the first disease-modifying treatments to have shown slowing of disease progression compared to placebo in large phase III studies.[5–7] With the recent approval by the United States Food and Drug Administration for the US market of aducanumab and lecanemab,[8,9] and positive results of a recently concluded phase III trial with donanemab,[6] the Alzheimer's disease field has entered a new era.

Still, the first generation of disease-modifying treatments is by no means a panacea. It is likely these treatments, which target one aspect of Alzheimer's disease (i.e., amyloid) and are characterized by limited effect sizes and risk of side effects, will be followed by future generations of improved disease-modifying treatments targeting amyloid (e.g., larger effect sizes, fewer and less serious side effects), as well as treatments targeting other aspects of the disease, such as tau, synapse pathology, or immune processes. Thus, the vital question is: who is most likely to benefit from particular treatments? It has been well-established that Alzheimer's disease is a heterogeneous disease in terms of biology,[10,11] genetics,[12] and clinical characteristics.[13] When a variety of patients are included in the same study and treated as a single group, the irresponsiveness of certain subgroups of patients may obfuscate beneficial effects of the same treatment in other subgroups.

Past trials that were negative or had only modest effects on their own, may nonetheless still hold important information, particularly when analyzed in combination with other available datasets. For example, they can provide insight into the mechanisms underlying treatment efficacy, inform about optimal outcome measures when analyzing specific targets, and can identify subgroups of participants that might have responded well to treatment. The benefits of pooled analysis across intervention studies have been shown previously by others.[14,15] By increasing sample size and statistical power, pooling individual participant data from multiple trials may allow us to uncover characteristics conducive to responsiveness to a specific treatment, despite negative trial results.

As a proof of principle, we endeavored to gain access to trial data and then attempted to pool, harmonize, and analyze the data, to study if participant characteristics could be identified that were related to increased responsiveness to treatment. In this perspective paper, we first report on different aspects of this effort, including the timeline and methodological approach. Specifically, we describe the practical hurdles we encountered. Finally, we make recommendations for future research.

2 | GAINING ACCESS TO TRIAL DATA

We reached out to contacts of seven trial sponsors that have worked with Brain Research Center in Amsterdam, The Netherlands, to inquire

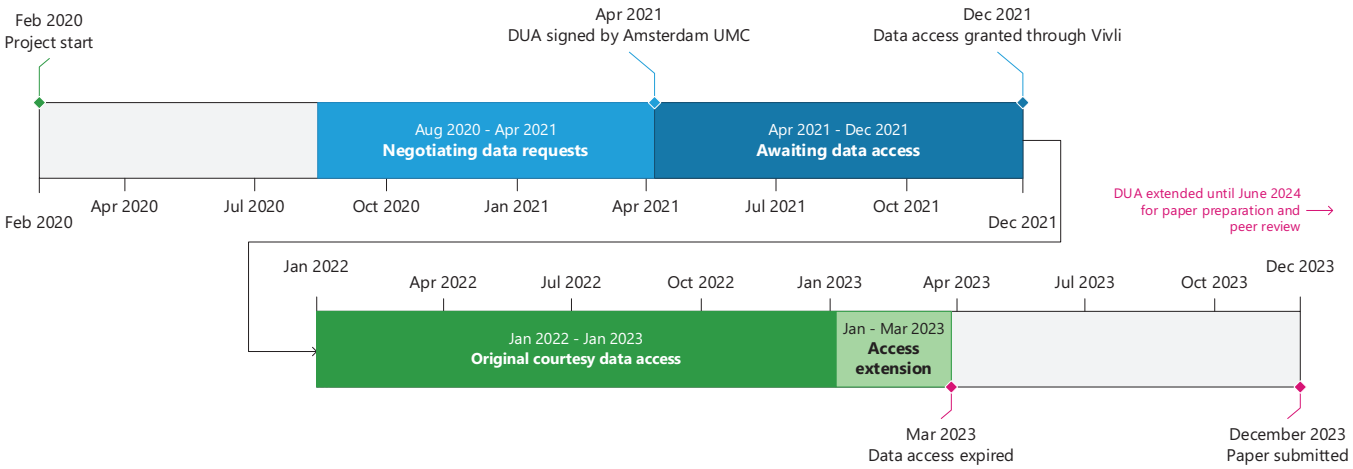

FIGURE 1 Data request timeline. Timeline to negotiate data access (shown in blue), including 8 months of waiting for the other party to sign the data use agreement. Initial courtesy data access was granted for 12 months upon approval, with a 3-month extension approved later. DUA, data use agreement; UMC, University Medical Center

about possibilities for obtaining access to data from their completed trials. Data requests were subsequently made through several platforms (Vivli, Yoda, Clinical Study Data Request) and by contacting researchers in the authors' network directly. Per current legislation, data use agreements between the party sharing the data (the platforms furnishing trial data) and the party receiving the data (Amsterdam UMC location VUmc) were a prerequisite for data access. Drafting and signing data use agreements required extensive back-and-forth between the legal departments of both parties over the course of many months. Complicating factors in signing the data use agreements included the applicability of United States versus Dutch law, liability concerns, and a shortage of legal personnel and lack of priority to process the agreements. In some cases, negotiations ended in stalemate, where the data use agreement was non-negotiable on the data providers' end, while the Dutch receiving academic organization could not accept all terms in the default American documents. Figure 1 displays a timeline of this project; from the moment it started in February of 2020 until the moment of submission of this paper in late 2023. All in all, it took over 2 years to obtain access to some trial data, while for others, the process was never concluded (that is, no access gained). Access was eventually granted to, and data provided from, six trial data sets from three sponsors investigating three different compounds. Access was granted to a trial dataset from one sponsor, but the data were never provided on the Vivli platform. Further, access was not granted to trials from two sponsors due to legal reasons, and we did not obtain access to another four trials due to other reasons.

3 | HARMONIZATION AND POOLING OF DATA

With access to several datasets, the next step was to investigate whether it is feasible to run analyses on pooled datasets in which the data from multiple trials were merged. Several platforms for trial data sharing exist, including Vivli (https://vivli.org/), the Yale University Open Data Access (YODA) project (https://yoda.yale.edu/), and ClinicalStudyDataRequest.com (https://clinicalstudydatarequest.com/). Briefly, all platforms host data on secure servers and researchers can submit proposals to obtain data access. Access is granted through a sealed desktop environment, that is, data are not downloaded directly to the researcher's device, but are available through the cloud. Data may not be exported, although it is possible to export tables and figures upon approval of the platform. For our study, data access was granted through the Vivli platform; thus, only data that were provisioned through the Vivli platform could be pooled for analysis. In our case, four datasets were made available through the Vivli platform directly, and two were transferred from the Yoda to the Vivli platform. The harmonization process included finding overlapping variables across datasets, restructuring datasets, and renaming variables so they share the same names in preparation for merging datasets, and aligning categories for categorical variables (such as race and ethnicity). To aid in this endeavor, anonymized study protocol documents and data dictionaries, which served as references for harmonizing the data, were also made available through Vivli. On a technical note, it was difficult to use the search function on these documents that were often hundreds of pages long, which meant all relevant variables and coding needed to be located manually. Within each study, rather than one dataset containing everything, separate datasets for each subset of data (e.g., demographic information, vital signs, cognitive endpoints, biomarkers, etc.) were provided.

The following trials granted access to their Individual patient data: Boehringer Ingelheim (BI) trials 1289.5 and 1289.7, Eli Lilly and Company (EL) trials LFAN and LFBC, and GlaxoSmithKline (GSK) trials BA1106006 and BA113043. The BI trials investigated compound BI409306, which is an inhibitor of phosphodiesterase 9A.[16] The GSK trials investigated compound GSK933776, which is a humanized mouse IgG1 monoclonal antibody for the N-terminus of the amyloid β peptide.[17] The EL trials investigated compound LY450139, or semagacestat, which is a γ-secretase inhibitor.[18] Trials showed differences

in study phase, type of compound, duration, and included population. None of the included trials were successful in changing cognitive functioning.

Table 1 displays descriptive information regarding study design of all trials that allowed access to their data. Measures that were shared across trials included participant age, sex, weight, height, and education. Outcome measures overlapped in part, with four out of six studies (the BI and EL trials) including repeated measures of the sum of boxes (SB) of the Clinical Dementia Rating (CDR),[19] and also repeatedly assessed the Alzheimer's Disease Assessment Scale—Cognitive (ADAS-Cog) subscale[20] and the Alzheimer's Disease Cooperative Study—Activities of Daily Living (ADCS-ADL) instrument.[21] All trials included the Mini-Mental State Examination (MMSE).[22] All trials except the BI trials also reported apolipoprotein ε4 carriership.

Biomarker data from cerebrospinal fluid (CSF) were scarce, only available for a subset of participants from the EL and GSK trials (n = 212, 6.7%). These included baseline CSF concentrations of amyloid β1-42, amyloid β1-40, phosphorylated tau_{181} (p-tau), and total tau (t-tau). It was difficult to harmonize these data because it was not clearly stated in the study documentation made available what assay was used to measure levels of amyloid and tau. As a result, we were unable to determine whether different studies used the same assays. Furthermore, no cutoffs to determine biomarker abnormality were provided, which could have been an alternative way to harmonize fluid biomarker data. This hindered pooled analysis of cerebrospinal fluid data of relevant Alzheimer's disease-related proteins. Similarly, data on hippocampal volume from brain magnetic resonance (MR) imaging could not be jointly analyzed, because only a small subset of participants in the EL trials underwent MR imaging.

3.1 | Proof-of-principle analysis

As a proof-of-principle, we conducted causal forest analysis on a total of 3170 participants included from six trials, of whom 1740 (54.9%) were female. Causal Forest analysis is a machine learning technique that can be used to find a minimal subset of patient characteristics that is related to heterogeneity in outcomes.[23,24] It can be used to model treatment effects and is a flexible method that allows for modelling higher order interactions, although it cannot model time, which is a previously identified limitation of this analysis technique.[25] As outcome for each trial, we used the change in the score at baseline subtracted from the score at the end-of-trial visit for: (i) MMSE, (ii) CDR-SB, and (iii) ADAS-Cog. For MMSE, a negative change represents a decline in cognitive functioning, whereas a positive change on ADAS-Cog and CDR-SB represents a decline in cognitive functioning. Treatment group (placebo vs. compound) was included as an effect modifier, while all shared patient characteristics, as well as treatment duration and dosage, were entered in the analysis as predictors (an overview of measures included is shown in Table 1). Causal Forests were built with 1000 trees to estimate confidence intervals for the predictions, using the total number of available baseline variables to randomly sample at each split. The minimum node size was set to 10. The random seed was set to the same specific number in all analyses to facilitate reproducibility. Causal forest analyses were run both on pooled trial data, and individual compounds. These analyses were performed within the Vivli platform, using R version 4.1.1.[26]

Causal forest analyses on the pooled data across all trials, or on the data within each single study, did not lead to the identification of participant characteristics that were related to heterogeneity in treatment effects on MMSE (estimate for differential forest prediction on pooled data = −2.58, SE = 1.46, p = .961), CDR (estimate on pooled data = −4.83, SE = 1.69, p = .998), or ADAS-Cog (estimate on pooled data = −1.86, SE = 1.28, p = .928). We did not find any evidence for characteristics that might have promoted treatment effectiveness among the data analyzed in this proof-of-principle study, which could be related to the fact that all studies were negative trials. However, we cannot conclude that such characteristics do not exist. Below, we will discuss various current hurdles in the use of existing trial data that may, at least in part, explain why the above analyses did not reveal evidence for any subgroup of participants that may have responded well to treatment.

4 | DISCUSSION

Completed clinical trial data sets, whether from positive or negative trials, hold a wealth of information, from which the scientific community can learn a lot. Given the complexity of Alzheimer's disease and the heterogeneity among patients, we need to improve strategies for patient stratification to work toward tailored treatment strategies and past trials can help us improve future trial design, characterize subgroups that might benefit from specific treatments, and identify theragnostic biomarkers or determinants of treatment response. For a long time, such trial data were inaccessible for researchers, and it is commendable that multiple companies have decided to share this valuable information. In our effort to gain access to, combine, and jointly analyze data from previous clinical trials investigating disease-modifying treatments for Alzheimer's disease, we identify and describe several hurdles and make recommendations for future research.

4.1 | Hurdles using existing trial data

In our attempt to gain access to and reanalyze existing data, we encountered several hurdles, mostly involving factors before the actual analysis could start. First, obtaining access to the data was a slow, complicated process, mostly due to extensive reviewing of data use agreements by legal departments on both the sending and receiving ends and one of the obstacles was the applicability of Dutch versus United States law. For us, it took 9 months to have the data use agreements signed on the receiving end, and then waited 8 more months to receive a signature from the sending party. This project was funded for 3 years, which means that funding received for almost half of that time could not go toward the actual analysis of the data. Unless the data-sharing process changes, funding agencies should adjust timelines for these types of projects.

TABLE 1 Information about trials that provided data

ClinicalTrials.gov study ID	NCT00762411	NCT00594568	NCT02337907	NCT02240693	NCT01424436	NCT00459550
Sponsor	EL	EL	BI	BI	GSK	GSK
Compound	LY450139	LY450139	BI409306	BI409306	GSK933776	GSK933776
Phase	III	III	II	II	I	I
Study dates[a]	9/2008–4/2011	3/2008–5/2011	1/2015–10/2017	1/2015–10/2017	5/2010-12/2011	3/2007–5/2011
Population						
N	1111	1537	329	128	19	50
AGE RANGE	≥55 years	≥55 years	≥55 years	≥55 years	55–85	55–80
MMSE	16–26	16–26	18–26	≥24	20–26	18–26
Duration	76 weeks	88 weeks	12 weeks	12 weeks	8 weeks[b]	34 weeks[c]
Primary outcomes	ADAS-Cog, ADCS-ADL	ADAS-Cog, ADCS-ADL	NTB	NTB	CSF amyloid-β	Adverse events
Secondary outcomes	CDR-SB, NPI, RUD-Lite, EQ-5D, QoL-AD, MMSE	CDR-SB, NPI, RUD-Lite, EQ-5D, QoL-AD, MMSE	ADAS-Cog, ADCS-ADL, CDR-SB	ADAS-Cog, ADCS-ADL, CDR-SB	CSF amyloid-β, CSF p-tau	Plasma and CSF biomarkers
Other measures	Age, sex, education, race/ethnicity, pulse, blood pressure, weight, length, smoking and alcohol use, APOE carriership, MR (in subset), CSF (in subset)	Age, sex, education, race/ethnicity, pulse, blood pressure, weight, length, smoking and alcohol use, APOE carriership, MR (in subset), CSF (in subset)	Age, sex, education, race, pulse, blood pressure, weight, length, smoking and alcohol use	Age, sex, education, race, pulse, blood pressure, weight, length, smoking and alcohol use	Age, sex, education, race/ethnicity, pulse, blood pressure, weight, length, APOE carriership, CSF	Age, sex, education, race/ethnicity, pulse, blood pressure, weight, length, APOE carriership, CSF

Abbreviations: ADAS-Cog, Alzheimer's Disease Assessment Scale—Cognitive; ADCS-ADL, Alzheimer's Disease Cooperative Study—Activities of Daily Living; APOE, apolipoprotein; BI, Boehringer Ingelheim; CDR-SB, Clinical Dementia Rating—Sum of Boxes; CSF, cerebrospinal fluid; EL, Eli Lilly and Company; EQ-5D, EuroQol 5-Dimensionional Health-Related Quality of Life Scale; GSK, GlaxoSmithKline; MMSE, Mini-Mental State Examination; MR, magnetic resonance; NTB, Neuropsychological Test Battery; QoL-AD, Quality of Life in Alzheimer's Disease; RUD, resource utilization in dementia.

[a]Dates shown as month/year.

[b]The primary outcome was measured after 22 hours, and again after 8 weeks (as a secondary outcome).

[c]This study was divided into two parts, with the first part concluding after 12 weeks.

Second, joint analysis of data was difficult to do because data were distributed across different data-sharing platforms that cannot interact, of which Vivli, Yoda, and ClinicalStudyDataRequest.com are a few examples. Distributing data through independent channels reduces the number of individual data use agreements that are required: rather than setting up agreements with each trial separately, a single agreement can be made with the data-sharing platform, covering multiple trials at once. However, current practice limits the possibility to pool data to just the datasets available within the same platform. This was also the case for our proof-of-principle analyses, which were limited to those data sets that were furnished through the Vivli platform. Unless platforms have agreements that allow for cross-platform sharing, it is not allowed to import or export individual patient data to or from the platform, which makes it difficult to perform direct statistical comparisons. As a result, this practice might be counterproductive: instead of making the process of data sharing more effective and efficient, there is a risk that the data-sharing platforms become too rigid in the way they share data, while flexibility is key when innovation is sought.

Third, an important factor to consider when attempting to pool and reanalyze existing trial data is heterogeneity in trial design, with respect to the research question, study population, target, and drug mechanism, among other aspects. While potential advantages of pooling data from multiple completed trials include the possibility to increase statistical power to detect subgroup differences, investigate new hypotheses, and use the heterogeneity in study populations to create a more representative sample, there are also potential disadvantages. For example, if differences between trials are too large, statistical power may be reduced, and pooled analyses may lead to inconclusive results.[14,25] Ultimately, the researcher requesting trial data should carefully consider which data to pool, depending on their research question and aims, to avoid ending up with a patchwork of datasets that do not form a cohesive collection that serves the intended research. In our proof-of-principle study, the trials showed large heterogeneity in study set-up, including differences in follow-up durations, intervals between study visits, and assessments performed at each visit. Moreover, the overlap of variables between the datasets was quite limited. This is partly because we included trials from all phases in the drug development pipeline, yet it reduces the number of variables that may be included in any analyses. Demographic variables like age, sex, race and ethnicity, education, height, weight, and vital signs like heart rate and blood pressure were the only variables shared across all datasets. Cognitive and functional measures used as endpoints included some of the same instruments, like the commonly used CDR, but did not always overlap. Furthermore, biomarker data like amyloid and tau levels in cerebrospinal fluid could not be pooled, because based on the information provided in the platform it was not evident what assays were used in the studies. Also, these were only measured in a limited number of participants, and the sample was too small to use data-driven approaches to determine cut-points. When additional biomarker measurements, such as proteomics, would be available within the trial datasets, they are not immediately made available for use. While there is a wealth of data available from past trials, differences in study design and data collected make it difficult to pool data. These factors may have contributed to the negative results of our proof-of-principle analyses. It is also possible that there either was no heterogeneity in treatment effects to be found in these negative trials, or that the sample size was still too small to detect such effects. We should also note that the trials included in our proof-of-principle analyses ran between 2008 and 2017, when the knowledge of Alzheimer's disease mechanisms was not as advanced as it is today. Considering the large number of drugs currently under investigation,[2] many of which share the same drug mechanism (i.e., anti-amyloid treatments) and have included similar study populations, we expect that some of these aspects of heterogeneity that hampered our data analysis will be less of a problem in future research pooling data from more recently concluded trials.

Last, and perhaps most importantly, individuals who participate in clinical trials do so first and foremost to benefit future generations and to advance the field.[27] The data they contribute to the trials, whether in the active or in the placebo group, and whether the outcome of the trial is positive or negative, should be used to move forward our search for a cure for Alzheimer's disease. Thus, trial sponsors have a moral obligation to maximize the use of the data they collect in their trials. Restricting access to these data not only poses the risk of the data going to waste, especially in the face of a negative trial, but it also perpetuates a system where only a select few may contribute to the questions being investigated.

4.2 | Recommendations

We have several recommendations that may help overcome the hurdles described above. First, the administrative hassle to gain access to data should be minimized. Academic researchers are increasingly often encouraged to share their data with the community to promote transparent science by funding agencies like the National Institutes of Health,[28] yet trial sponsors lag behind when it comes to fostering open science. One solution would be that pharmaceutical companies make their individual patient data available as soon as trials have closed and results have been published, under the most liberal conditions feasible. This could help avoid unnecessary administrative delays. Furnishing open access to individual patient data will require companies to build new data-sharing infrastructures if they do not already have such platforms set up, and this might incur additional expenses that might not yield a direct return on investment. However, by collaborating with existing data platforms and initiatives, such as the Alzheimer's Disease Data Initiative workbench (https://www.alzheimersdata.org/) in which various trial sponsors already participate, expenses may be reduced. Moreover, data sharing between stakeholders has been identified as a conduit for advancing the development of medical products, including therapeutic drugs.[29] If pooled trial data analyses reveal patient subgroups where a certain treatment is (more) effective, it is possible that drugs from a previously negative trial might still be marketed for the treatment of Alzheimer's disease in specific subgroups of patients.

Second, rather than only allowing researchers to access data within a sealed-off platform, trial sponsors should follow the example set

by the Anti-Amyloid Treatment in Asymptomatic Alzheimer's (A4) study.[30] The A4 study is an example of a public-private partnership clinical trial that allows open access to their baseline data to any qualified researcher who, once approved, may download the data from a server and analyze them on their computer systems. Dozens of publications have already resulted from the baseline data and the full A4 trial data are expected to be released to the public in the near future. Following the A4 study's example, clinical trials could allow researchers to download data in accessible formats (e.g., comma-separated value files) onto their devices to analyze using the software of their choosing, rather than providing data access within a sealed-off platform. Naturally, the proper consent for data sharing should be obtained from participants prior to their enrollment, and the data should be properly deidentified to ensure that the privacy of the participant is maintained. Other examples of open data-sharing initiatives are the aforementioned Alzheimer's Disease Data Initiative workbench and the Critical Path of Alzheimer's Disease (https://c-path.org/program/critical-path-for-alzheimers-disease/). This latter initiative currently provides individual patient data from different trials, including demographic characteristics and cognitive test scores (e.g., MMSE). However, data regarding the treatment arm and AD biomarkers are still limited. Furthermore, this platform does not (yet) provide details regarding individual trials such as drug names and study protocols, because of which it is not yet fit for the purpose of finding heterogeneity in treatment effects. Still, it serves as an admirable example of how individual patient data from different trail datasets can be combined.[31] As an alternative to allowing researchers to download data from a server to their research environments, sponsors, and pharmaceutical companies should collaborate to make their individual patient data available on either a single, or on all data sharing platforms.

Third, some of the limitations for jointly analyzing data across trials are created when a study is designed. Regarding study design, two points need to be addressed here: (i) routine assessment of Alzheimer's disease biomarkers in cerebrospinal fluid, plasma, or on positron emission tomography imaging should be the standard for any trial in any of the phases; (ii) inclusion of a core outcome set. First, collecting biomarker data in a standardized manner and including assays of markers that might not be directly relevant to the treatment target can help ensure that heterogeneity in participants' biological characteristics can be adequately analyzed across trials. Second, while we acknowledge the importance of trials retaining their unique study designs to align with a compound, study phase, and trial population-specific needs,[32] employing a uniform core set of well-validated outcome measures will be the key to assessing treatment efficacy across trials. This core outcome set[33] would allow analyses across multiple trials. Importantly, trial-specific measures should supplement a core outcome set and should be tailored to the sample (e.g., adequate for participants' age and disease stages), trial duration, and target treatment. As an example, the Alzheimer's Disease Neuroimaging Initiative (ADNI) collects both biomarker and clinical data in a harmonized manner, allowing for the joint analysis of data from many patients across numerous sites in the United States.[34] ADNI has helped answer many questions in the Alzheimer's disease field and proves that the power is in the numbers. Further, to detect clinically meaningful changes, a critical revisit of the clinical outcome measures is warranted[32,35]: it may be time to retire legacy instruments that are insensitive to cognitive changes over relatively short amounts of time, particularly in very early disease stages. Together, this will allow future trial data to be pooled more easily and may pave the way for more meaningful cross-trial analyses.

Last, while pharmaceutical companies can aid by standardizing some aspects of their study design and by making data more easily accessible, it will be the scientific community's responsibility to use the full potential of trial data by carefully considering (a) what scientific questions they intend to answer using trial data, and (b) what statistical methods to use when attempting to analyze and/or pool these data. For example, while pooling individual patient data across trials may be preferable to a meta-analysis when studying heterogeneity, the latter might serve well for other research questions, is less expensive and less time-consuming, and does not require pharmaceutical companies to provide individual patient data.[15] On the other hand, with individual patient data, statistical methods to investigate heterogeneity in treatment effects can be considered.

4.3 | Future perspective

We envision a future in which open science is the norm, for both academia and trialists, open access data are provided on request, and all trials include a core outcome set and routinely measure Alzheimer's disease biomarkers, while also incorporating unique trial features. This would facilitate comparisons between trials and could perhaps aid in identifying factors in the patient population that contribute to heterogeneity. We may only reach the finish line when data is openly accessible to all, regardless of their affiliation, funding, or background.

5 | CONCLUSION

In our proof-of-principle study, we did not succeed to meaningfully pool data from Alzheimer's disease clinical trials. Although it seemed, in theory, that quite a few datasets should be readily accessible, it turned out that we were too optimistic about the current state of the field. We thus conclude that, under the current circumstances, the outcome does not outweigh the effort. Nonetheless, joining efforts and pooling data to increase the pace at which we, as a scientific community, learn about this devastating disease is the future. We therefore plea for open science, not just for observational studies, but also for pharmacological trials. We advocate for removing the barriers that hinder data sharing and for making trial data more easily and openly accessible, so the research field may learn from past failures and successes. That way, we can ultimately work toward personalized treatment for Alzheimer's disease.

ACKNOWLEDGMENTS

This publication is based on research using data from data contributors Eli Lilly and Company, Boehringer Ingelheim, and GlaxoSmithKline

that has been made available through Vivli, Inc. Vivli has not contributed to or approved, and Vivli, Eli Lilly and Company, Boehringer Ingelheim, and GlaxoSmithKline are not in any way responsible for, the contents of this publication. This study, carried out under YODA Project #2021-4567, used data obtained from the Yale University Open Data Access Project, which has an agreement with Janssen Research & Development, L.L.C. The interpretation and reporting of research using these data are solely the responsibility of the authors and do not necessarily represent the official views of the Yale University Open Data Access Project or Janssen Research & Development, L.L.C. This work was supported through OTAPA, a collaboration project that is co-funded by the PPP Allowance made available by Health-Holland, Top Sector Life Sciences & Health to stimulate public-private partnerships, and Brain Research Center (grant number LSHM19051). Research of Alzheimer Center Amsterdam is part of the Neurodegeneration research program of Amsterdam Neuroscience. Alzheimer Center Amsterdam is supported by Stichting Alzheimer Nederland and Stichting Steun Alzheimercentrum Amsterdam. Mark A. Dubbelman received grant support from Alzheimer Nederland (WE.06-2023-02), paid to his institution. Eleonora M. Vromen, Betty M. Tijms„ and Lois Ottenhoff report no disclosures relevant to the manuscript. Everard G. B. Vijverberg has received consultancy fees (paid to his university) for New Amsterdam Pharma, Treeway, ReMynd, Vivoryon, Biogen, Vigil Neuroscience, ImmunoBrain Checkpoint, and Roche. Within his university affiliation, EGBV is PI of studies of DIAN, AC immune, Alnylam, CogRX therapeutics, New Amsterdam Pharma, Janssen, UCB, Roche, Vivoryon, ImmunoBrain, GSK, Biogen, and Alector. Sub-Investigator from Eli Lilly, Fuij Film Toyama, GemVax. Niels D. Prins performed consultancy work for Aribio, Amylyx, Eli-Lilly, and Janssen, and received a speaker fee from Biogen. Niels D. Prins is CEO and co-owner of Brain Research Center, the Netherlands. Johannes Berkhof, Everard G. B. Vijverberg, Niels D. Prins, and Wiesje M. van der Flier are recipients of ABOARD (A Personalized Medicine Approach for Alzheimer's Disease); a public-private partnership receiving funding from ZonMW (#73305095007) and Health~Holland, Topsector Life Sciences & Health (PPP-allowance #LSHM19051). More than 30 partners participate in ABOARD. Research programs of WMF have been funded by ZonMW, NWO, EU-JPND, EU-IHI, Alzheimer Nederland, Hersenstichting CardioVascular Onderzoek Nederland, Health~Holland, Topsector Life Sciences & Health, stichting Dioraphte, Gieskes-Strijbis fonds, stichting Equilibrio, Edwin Bouw fonds, Pasman stichting, stichting Alzheimer & Neuropsychiatrie Foundation, Philips, Biogen MA Inc, Novartis-NL, Life-MI, AVID, Roche BV, Fujifilm, Eisai, Combinostics. Wiesje M. van der Flier holds the Pasman chair. Wiesje M. van der Flier has been an invited speaker at Biogen MA Inc, Danone, Eisai, WebMD Neurology (Medscape), NovoNordisk, Springer Healthcare, and European Brain Council. Wiesje M. van der Flier is consultant to Oxford Health Policy Forum CIC, Roche, Biogen MA Inc, and Eisai. Wiesje M. van der Flier is a member of the steering commission of NovoNordisk evoke/evoke+. Wiesje M. van der Flier participated in advisory boards of Biogen MA Inc, Roche, and Eli Lilly. All funding is paid to her institution. Wiesje M. van der Flier is member of the steering committee of PAVE and Think Brain Health. Wiesje M. van der Flier was associate editor of Alzheimer, Research & Therapy in 2020/2021. Wiesje M. van der Flier is an associate editor at Brain. SAMS received grant support from Health Holland (LSHM19051, LSHM20084, LSHM22026-SGF), and Zon-MW (#7330502051 and #73305095008). Sietske A. M. Sikkes provided consultancy services for Biogen, Boehringer, and Toyama. Sietske A. M. Sikkes is the developer of the Amsterdam IADL Questionnaire, and received license fees from Green Valley, VtV Therapeutics, Alzheon, Vivoryon, and Roche. All funding was paid to her institution.

CONFLICT OF INTEREST STATEMENT

Author disclosures are available in the supporting information.

ORCID

Mark A. Dubbelman https://orcid.org/0000-0002-5708-4925

REFERENCES

1. Scheltens P, De Strooper B, Kivipelto M, et al. Alzheimer's disease. *Lancet*. 2021;397(10284):1577-1590. doi:10.1016/s0140-6736(20)32205-4
2. Cummings J, Zhou Y, Lee G, Zhong K, Fonseca J, Cheng F. Alzheimer's disease drug development pipeline: 2023. *Alzheimers Dement (N Y)*. 2023;9(2):e12385. doi:10.1002/trc2.12385
3. Cummings J, Lee G, Zhong K, Fonseca J, Taghva K. Alzheimer's disease drug development pipeline: 2021. *Alzheimers Dement (N Y)*. 2021;7(1):e12179. doi:10.1002/trc2.12179
4. Cummings J, Lee G, Nahed P, et al. Alzheimer's disease drug development pipeline: 2022. *Alzheimers Dement (N Y)*. 2022;8(1):e12295. doi:10.1002/trc2.12295
5. McDade E, Cummings JL, Dhadda S, et al. Lecanemab in patients with early Alzheimer's disease: detailed results on biomarker, cognitive, and clinical effects from the randomized and open-label extension of the phase 2 proof-of-concept study. *Alzheimers Res Ther*. 2022;14(1):191. doi:10.1186/s13195-022-01124-2
6. Sims JR, Zimmer JA, Evans CD, et al. Donanemab in early symptomatic Alzheimer disease: the TRAILBLAZER-ALZ 2 randomized clinical trial. *JAMA*. 2023;330(6):512-527. doi:10.1001/jama.2023.13239
7. Budd Haeberlein S, Aisen PS, Barkhof F, et al. Two randomized phase 3 studies of aducanumab in early Alzheimer's disease. *J Prev Alzheimers Dis*. 2022;9(2):197-210. doi:10.14283/jpad.2022.30
8. FDA converts novel Alzheimer's disease treatment to traditional approval. July 6, 2023. Accessed August 8, 2023. https://www.fda.gov/news-events/press-announcements/fda-converts-novel-alzheimers-disease-treatment-traditional-approval
9. FDA grants accelerated approval for Alzheimer's drug. June 7, 2021. Accessed August 8, 2023. https://www.fda.gov/news-events/press-announcements/fda-grants-accelerated-approval-alzheimers-drug
10. Chen P, Yao H, Tijms BM, et al. Four distinct subtypes of Alzheimer's disease based on resting-state connectivity biomarkers. *Biol Psychiatry*. 2023;93(9):759-769. doi:10.1016/j.biopsych.2022.06.019
11. Tijms BM, Gobom J, Reus L, et al. Pathophysiological subtypes of Alzheimer's disease based on cerebrospinal fluid proteomics. *Brain*. 2020;143(12):3776-3792. doi:10.1093/brain/awaa325
12. Ringman JM, Goate A, Masters CL, et al. Genetic heterogeneity in Alzheimer disease and implications for treatment strategies. *Curr Neurol Neurosci Rep*. 2014;14(11):499. doi:10.1007/s11910-014-0499-8
13. Badhwar A, McFall GP, Sapkota S, et al. A multiomics approach to heterogeneity in Alzheimer's disease: focused review and roadmap. *Brain*. 2020;143(5):1315-1331. doi:10.1093/brain/awz384

14. Bangdiwala SI, Bhargava A, O'Connor DP, et al. Statistical methodologies to pool across multiple intervention studies. *Transl Behav Med*. 2016;6(2):228-235. doi:10.1007/s13142-016-0386-8
15. Blettner M, Sauerbrei W, Schlehofer B, Scheuchenpflug T, Friedenreich C. Traditional reviews, meta-analyses and pooled analyses in epidemiology. *Int J Epidemiol*. 1999;28(1):1-9. doi:10.1093/ije/28.1.1
16. Wunderlich G, Thamer C, Roehrle M, Garcia M, Frölich L, Dubois B. P3-009: study design and characteristics of two phase II proof-of-concept clinical trials of the PDE9 inhibitor BI 409306 in early Alzheimer's disease. *Alzheimers Dement*. 2016;12(7S_Part_16):pp820-821. doi:10.1016/j.jalz.2016.06.1666
17. Andreasen N, Simeoni M, Ostlund H, et al. First administration of the Fc-attenuated anti-beta amyloid antibody GSK933776 to patients with mild Alzheimer's disease: a randomized, placebo-controlled study. *PLoS One*. 2015;10(3):e0098153. doi:10.1371/journal.pone.0098153
18. Doody RS, Raman R, Farlow M, et al. A phase 3 trial of semagacestat for treatment of Alzheimer's disease. *N Engl J Med*. 2013;369(4):341-350. doi:10.1056/NEJMoa1210951
19. Morris JC. Clinical dementia rating: a reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. *Int Psychogeriatr*. 1997;9(Suppl 1):173-176.
20. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. *Am J Psychiatry*. 1984;141(11):1356-1364. doi:10.1176/ajp.141.11.1356
21. Galasko D, Bennett D, Sano M, et al. An inventory to assess activities of daily living for clinical trials in Alzheimer's disease. The Alzheimer's Disease Cooperative Study. Article. *Alzheimer Dis Assoc Dis*. 1997;11(Suppl 2):S33-39.
22. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12(3):189-198. doi:10.1016/0022-3956(75)90026-6
23. Liu Y, Li Z, Ge Q, Lin N, Xiong M. Deep feature selection and causal analysis of Alzheimer's disease. *Front Neurosci*. 2019;13:1198. doi:10.3389/fnins.2019.01198
24. Wager S, Athey S. Estimation and inference of heterogeneous treatment effects using random forests. *J Am Stat Assoc*. 2018;113(523):1228-1242. doi:10.1080/01621459.2017.1319839
25. Brantner CL, Nguyen TQ, Tang T, Zhao C, Hong H, Stuart EA. Comparison of methods that combine multiple randomized trials to estimate heterogeneous treatment effects. *Stat Med*. 2024;43(7):1291-1314. doi:10.1002/sim.9955
26. R: A language and environment for statistical computing. Version 4.3.1. 2023. https://www.R-project.org/
27. Ottenhoff L, Vijverberg EGB, Visser LNC, et al. Experiences of and recommendations on clinical trial design in Alzheimer's disease from the participant's point of view: a mixed-methods study in two clinical trial centers in the Netherlands. *Alzheimers Res Ther*. 2023;15(1):72. doi:10.1186/s13195-023-01190-0
28. Final NIH policy for data management and sharing. October 29, 2020. Accessed February 21, 2024. https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html
29. Karpen SR, White JK, Mullin AP, et al. Effective data sharing as a conduit for advancing medical product development. *Ther Innov Regul Sci*. 2021;55(3):591-600. doi:10.1007/s43441-020-00255-8
30. Sperling RA, Rentz DM, Johnson KA, et al. The A4 study: stopping AD before symptoms begin? *Sci Transl Med*. 2014;6(228):228fs13. doi:10.1126/scitranslmed.3007941
31. Sivakumaran S, Hanan NJ, Cui Z, et al. The Critical Path for Alzheimer's Disease (CPAD) Consortium: a platform for pre-competitive data sharing, standardization, and analysis to support quantitative tools for AD drug development. *Alzheimers Dement*. 2021;17(S9):e051903. doi:10.1002/alz.051903
32. Jutten RJ, Papp KV, Hendrix S, et al. Why a clinical trial is as good as its outcome measure: a framework for the selection and use of cognitive outcome measures for clinical trials of Alzheimer's disease. *Alzheimers Dement*. 2023;19(2):708-720. doi:10.1002/alz.12773
33. Prinsen CA, Vohra S, Rose MR, et al. How to select outcome measurement instruments for outcomes included in a "Core Outcome Set"—a practical guideline. *Trials*. 2016;17(1):449. doi:10.1186/s13063-016-1555-2
34. Petersen RC, Aisen PS, Beckett LA, et al. Alzheimer's Disease Neuroimaging Initiative (ADNI): clinical characterization. *Neurology*. 2010;74(3):201-209. doi:10.1212/WNL.0b013e3181cb3e25
35. Dubbelman MA, Hendriksen HMA, Harrison JE, et al. Cognitive and functional change over time in cognitively healthy individuals according to Alzheimer disease biomarker-defined subgroups. *Neurology*. 2024;102(2):e207978. doi:10.1212/WNL.0000000000207978

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Dubbelman MA, Vromen EM, Tijms BM, et al. Pooling Alzheimer's disease clinical trial data to develop personalized medicine approaches is easier said than done: A proof-of-principle study and call to action. *Alzheimer's Dement*. 2024;10:e12485. https://doi.org/10.1002/trc2.12485