



Transferability of Alzheimer's disease progression subtypes to an independent population cohort

Hanyi Chen^a, Alexandra Young^b, Neil P. Oxtoby^a, Frederik Barkhof^{a,c,d}, Daniel C. Alexander^a, Andre Altmann^{a,*}, for the ADNI investigators¹

^a Centre for Medical Image Computing Department of Medical Physics and Biomedical Engineering and Department of Computer Science, University College London, UK

^b Department of Neuroimaging Institute of Psychiatry Psychology and Neuroscience, King's College London, UK

^c Queen Square Institute of Neurology, University College London, UK

^d Department of Radiology and Nuclear Medicine, Amsterdam University Medical Center, The Netherlands

ARTICLE INFO

Keywords:

Alzheimer's disease
Subtypes
Modelling
Early risk factors

ABSTRACT

In the past, methods to subtype or biotype patients using brain imaging data have been developed. However, it is unclear whether and how these trained machine learning models can be successfully applied to population cohorts to study the genetic and lifestyle factors underpinning these subtypes. This work, using the Subtype and Stage Inference (SuStaIn) algorithm, examines the generalisability of data-driven Alzheimer's disease (AD) progression models.

We first compared SuStaIn models trained separately on Alzheimer's disease neuroimaging initiative (ADNI) data and an AD-at-risk population constructed from the UK Biobank dataset. We further applied data harmonization techniques to remove cohort effects. Next, we built SuStaIn models on the harmonized datasets, which were then used to subtype and stage subjects in the other harmonized dataset.

The first key finding is that three consistent atrophy subtypes were found in both datasets, which match the previously identified subtype progression patterns in AD: 'typical', 'cortical' and 'subcortical'. Next, the subtype agreement was further supported by high consistency in individuals' subtypes and stage assignment based on the different models: more than 92% of the subjects, with reliable subtype assignment in both ADNI and UK Biobank dataset, were assigned to an identical subtype under the model built on the different datasets. The successful transferability of AD atrophy progression subtypes across cohorts capturing different phases of disease development enabled further investigations of associations between AD atrophy subtypes and risk factors. Our study showed that (1) the average age is highest in the typical subtype and lowest in the subcortical subtype; (2) the typical subtype is associated with statistically more-AD-like cerebrospinal fluid biomarkers values in comparison to the other two subtypes; and (3) in comparison to the subcortical subtype, the cortical subtype subjects are more likely to associate with prescription of cholesterol and high blood pressure medications.

In summary, we presented cross-cohort consistent recovery of AD atrophy subtypes, showing how the same subtypes arise even in cohorts capturing substantially different disease phases. Our study opened opportunities for future detailed investigations of atrophy subtypes with a broad range of early risk factors, which will potentially lead to a better understanding of the disease aetiology and the role of lifestyle and behaviour on AD.

1. Introduction

Alzheimer's disease (AD) is a global health and economic burden affecting around 47 million individuals worldwide. Currently, there exists only one FDA-approved drug with some disease-modifying potential in a subset of patients. In general, a key confound

that prevents successful outcomes in most treatment trials to date has been AD's high variability in onset, mechanism, and clinical expression.

In the literature, the amyloid cascade hypothesis (Hardy and Higgins, 1992; Selkoe and Hardy, 2016) posits that a build-up in the amyloid- β protein is the pathological origin, followed by formation of

* Corresponding author.

E-mail address: a.altmann@ucl.ac.uk (A. Altmann).

¹ Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in the analysis or the writing of this report. A complete list of ADNI investigators can be found at: http://adni.loni.usc.edu/wpcontent/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

<https://doi.org/10.1016/j.neuroimage.2023.120005>.

Received 12 December 2022; Received in revised form 22 February 2023; Accepted 6 March 2023

Available online 11 March 2023.

1053-8119/© 2023 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

intracellular neurofibrillary tangles (NFTs) consisting of the tau protein, decline in brain glucose metabolism, atrophy of grey matter (in particular in the hippocampus) and finally cognitive decline (Jack Jr et al., 2010). This theoretical model of biomarker progression has been supported by longitudinal studies conducted in asymptomatic mutations carriers for autosomal dominant AD (Bateman et al., 2012) and disease progression models applied to cross-sectional studies of sporadic, late-onset AD (Donohue et al., 2014; Young et al., 2014; Lorenzi et al., 2019; Koval et al., 2021; Venkatraghavan et al., 2019). However, the evolution of biomarkers in AD and many other brain disorders remains uncertain and there is heterogeneity in the order in which biomarkers show abnormality. For instance, different pattern of hypometabolism, brain atrophy and tau distribution have been linked to AD subtypes with different cognitive profiles (Laforce et al., 2014; Ossenkoppele et al., 2015, 2016).

Studies tackling the heterogeneity can be broadly categorised into two streams: staging methods that emphasize the temporal progressions of disease process (Bilgel et al., 2016; Donohue et al., 2014; Young et al., 2014), and subtyping methods that focus on identifying distinct groups of patients based on their phenotypic heterogeneity (Nettiksimmons et al., 2010; Scheltens et al., 2017; Tijms et al., 2020; Whitwell et al., 2009). The recently developed Subtype and Stage Inference (SuStaIn) algorithm (Young et al., 2018) is an unsupervised learning technique that identifies disease progression subtypes in cross-sectional data. It uniquely defines subtypes by a trajectory of change, thereby avoiding confounds of temporal change and phenotypic difference. Using SuStaIn, Young et al. (2018) identified three data driven atrophy subtypes with distinct temporal progression patterns based on cross-sectional brain MRI scans from the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset. Furthermore, within AD, SuStaIn has been successfully used to identify different subtypes of temporal progression of NFTs and amyloid plaques (Aksman et al., 2020) as well as spatio-temporal spread of cortical tau (Vogel et al., 2021). Beyond AD, SuStaIn has identified atrophy subtypes in genetic frontotemporal dementia (Young et al., 2018, 2021) reflecting the causative mutation, in multiple sclerosis (Eshaghi et al., 2021) and in Chronic Obstructive Pulmonary Disease (Young et al., 2020). The reason for the emergence of distinct disease subtypes in AD remains elusive and is likely linked to genetic risk factors, co-morbidities, and environmental factors; many of which may act upstream of amyloid- β .

While recent studies have uncovered many factors that increase the risk for developing AD such as diabetes (Andrews et al., 2020), blood pressure complexity and variability (Ma et al., 2019), various lifestyle factors including smoking status, alcohol consumption, physical, social and leisure activities (Lourida et al., 2019; Sommerlad et al., 2019; Yates et al., 2016) and genetic risk factors (Kunkle et al., 2019), no such study has yet been performed for AD progression subtypes. One of the challenges is that available datasets are either disease-focused (with deep phenotyping using disease-related biomarkers, but lacking decades of pre-clinical data) or are (longitudinal) population cohorts (with deep phenotyping on risk factors, but lacking disease-related biomarkers). Thus, neither dataset can be used to train subtyping tools and to explore early risk factors influencing subtype development. To tackle this challenge, we aim to connect datasets that offer windows on different phases of disease development. We proposed to train atrophy-based subtyping models on disease-focused datasets such as ADNI and then 'transfer' them to population cohorts such as UK Biobank (UKBB). This is not a form of transfer learning in the traditional sense, where a model trained on one dataset is adjusted to solve a different, related task on a different dataset. To quantify the success of a model transfer, we define consistency of subtype and stage assignments at subject-level between the transferred model and a reference model, which was trained on the target dataset. Successful model transfers will enable the investigation of subtype-related risk factors that go beyond what can be discovered in each dataset individually.

2. Methods

Working across the ADNI and UKBB datasets raises many interesting research questions: (1) do consistent atrophy subtypes emerge from these two datasets? (2) how to eliminate cohort effects on biomarkers for the two datasets while preserving biological information? (3) can a model trained on ADNI effectively subtype and stage participants in UKBB? Fig. 1 gives a summary of the analysis strategy as well as the datasets used for modelling and the models built on the datasets. We briefly outline the overall organisation of the analysis with detailed information being provided later in the Methods section. We addressed these issues by first comparing the SuStaIn models for measures of cortical and subcortical volumes trained separately on the original ADNI dataset and an AD-at-risk population constructed from UKBB dataset. Then we applied the harmonization technique ComBat (Johnson et al., 2007; Fortin et al., 2017) to these measures of cortical and subcortical volumes to remove the cohort effect while preserving the biological covariates effects. We next trained the SuStaIn model on the two harmonized datasets. We hypothesized that the ADNI dataset, due to its focus on AD pathology, would constitute the better training set and that harmonization would be critical to a successful model transfer. In total, we trained four SuStaIn models: an ADNI SuStaIn, a UKBB SuStaIn, a harmonized ADNI SuStaIn and a harmonized UKBB SuStaIn. For each subject in the two datasets, we compared the subtype and stage assignments under three models, namely, the model trained on the original dataset it belongs to and the models trained on the two harmonized datasets. In addition, to highlight the central role of harmonization, we also applied the ADNI SuStaIn model to the unharmonized UKBB data and studied the UKBB subjects' subtype assignments under the ADNI SuStaIn model and under the UKBB SuStaIn model.

Finally, to illustrate the usage of transferring SuStaIn models, we studied associations of AD subtypes with factors including: age, the CSF biomarkers ($A\beta$, tau and p-tau), and medication history variables.

2.1. ADNI dataset

Data used in the preparation of this article was obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (<http://adni.loni.usc.edu>). ADNI was initially launched in 2004 led by Principal Investigator Michael W. Weiner, MD, funded as a public-private partnership, that has made major contributions to AD research, enabling the sharing of data between researchers around the world. The primary goal of ADNI has been to support studies of AD early detection and tracking disease's progression using biomarkers including magnetic resonance imaging (MRI), positron emission tomography (PET) and clinical and neuropsychological assessments. For up-to-date information, see www.adni-info.org.

Demographic information (i.e., age, sex, intracranial volume [ICV] and years of education), diagnostic labels (CN: cognitively normal, SMC: substantial memory complaint, EMCI: early mild cognitive impairment, LMCI: late mild cognitive impairment, AD: Alzheimer's disease), and CSF biomarker values ($A\beta$, tau and p-tau) were extracted from the table 'ADNIMERGE'. MRI-based features were extracted from T1-weighted MRIs obtained on 3T scanners at baseline for ADNI1/GO/2 participants. The scans were processed with FreeSurfer 5.1 to obtain 80 cortical and subcortical regional measures based on the Desikan-Killiany atlas. Only scans that passed overall quality control carried out by ADNI were retained (see here for details: <http://adni.loni.usc.edu/methods/mri-tool/mri-analysis/>).

2.2. UKBB dataset

This research has been conducted using the UK Biobank under application 70047.

UK Biobank is a major UK collaborative research project that was initially established as a charitable company in 2003. The large-scale

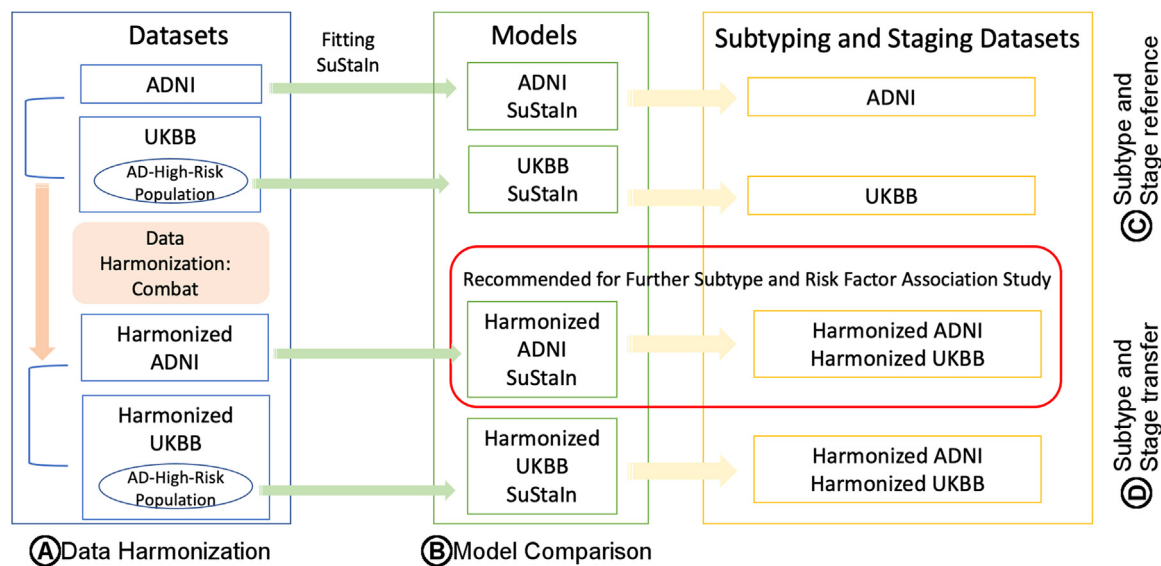


Fig. 1. Analysis overview. To explore the transferability of SuStaln models we train models on different datasets with or without using feature harmonization (A). The training data for SuStaln is either the original ADNI dataset or a high-risk AD population constructed from the UKBB. The resulting models are referred to as ADNI SuStaln and UKBB SuStaln models, respectively. This allows us to investigate how similar the models from the two datasets are (B). Next, these models were applied to assign subtype and stage to individuals within the same dataset to create a subtype and stage reference (C). To explore the transferability of SuStaln models we first harmonized the features in the training datasets using ComBat (A). Then two additional SuStaln models were trained on the harmonized versions of the two datasets. The resulting Harmonized ADNI SuStaln and Harmonized UKBB SuStaln models were used to assign subtype and stage to individuals within and across the two harmonized datasets (D). This setup enabled us to quantify the consistency of using un-harmonized and harmonized training data as well as assessing the consistency of subtype and stage assignment on the subject level. Our study suggested to use the Harmonized ADNI SuStaln model for individual subtype and stage for both datasets for further subtype and risk factor association study.

biomedical database and research resource, containing in-depth genetic and follow-up health information from half a million UK participants aged between 40 - 69 years, has enabled scientific researchers to investigate meaningful associations between subgroups of the population and certain diseases, which further allows improvement of healthcare according to the varying causes, prognosis and response to treatment of common diseases including determining the risks of disease in different groups, providing direct evidence of the scope for prevention and leading to the development of improved diagnostic methods and treatments. For up-to-date information, see <https://www.ukbiobank.ac.uk>.

We downloaded the regional volume data based on T1-weighted MRI scans (80 cortical and subcortical volume metrics based on the Desikan-Killiany atlas) from UKBB. All scans were acquired on the same type of 3T scanner (Siemens Skyra) and processed with FreeSurfer 6.0 and FreeSurfer 7.0 for cortical and for subcortical regions, respectively (see http://www.fmrib.ox.ac.uk/ukbiobank/protocol/V4_23092014.pdf and Alfaro-Almagro et al., 2018). Quality control was carried out prior to the upload to UKBB and measures for the image data include signal-to-noise-ratio for individual modalities and ‘discrepancy’ between a given pair of images after they have been co-registered (for details see: https://biobank.utsu.ac.uk/crystal/crystal/docs/brain_mri.pdf). To define control and AD-at-risk populations for the UKBB dataset, we also obtained AD-related variables including number of APOE- ϵ 2 and APOE- ϵ 4 alleles (from genotyping data), family history of dementia, self-reported neurological or psychiatric disorders and cognitive test results. From the eleven cognitive tests available in UKBB, we focused on three tests that assess reaction time (UKBB RT), executive function (UKBB Trail Making Test A and B) and delayed memory (UKBB Prospective Memory Test [PMT]). The UKBB PMT is a simple test, where early during the touchscreen cognitive section, the participant is shown a message with an instruction for a later point in the session. The first and final answer, the history of attempts and the time it took to answer the later question are recorded. To illustrate the practical benefit of transferring trained SuStaln models and being able to leverage the deep phenotyping of other non-AD cohorts, we further obtained

the data on cholesterol and blood pressure medication use from UKBB.

2.3. Defining control, case (training) study populations in ADNI and UKBB

For the ADNI dataset, the control set consisted of SMC and CN subjects with negative amyloid, where amyloid negativity was defined based on Florbetapir PET using the ADNI criteria, i.e., cortical SUVR normalized by whole cerebellum < 1.11 . We regarded all the subjects passing overall quality control and who were not in the control dataset as cases (training set) for SuStaln modelling.

To assess the quality of the model transfer we require a SuStaln model trained on the UKBB data to establish a subtyping reference. However, the UKBB dataset lacks AD-specific phenotyping such as diagnostic labels and CSF or PET measurements. Consequently, we can neither define a control set free of AD pathology nor a training set with a clear AD pathology. However, SuStaln is an unsupervised method and therefore it is sufficient to enrich the training for participants with beginning AD pathology to form AD-related subtypes. Likewise, the control set is used to standardize biomarkers and therefore needs only to be ‘mostly healthy’ as a small number of subjects exhibiting pathological aging is unlikely to substantially affect the z-scoring. We therefore use the following definitions for a ‘healthy control’ set and ‘AD-at-risk’ training set.

The control set of UKBB consisted of subjects not carrying the APOE- ϵ 4 allele and without family history of dementia, or self-reported neurological or psychiatric disorders (see Table S1 in the supplementary materials for the list of neurological or psychiatric disorders examined), and who performed well in all the AD-related cognitive tests according to the following set of standards: (1) no error made in the two trail making tests; (2) time to complete the trail making path is less than the UKBB sample median; (3) the mean time to correctly identify items in the reaction time test is less than UKBB sample median; (4) had the correct recall in the UKB PMT on first attempt and the time to answer is less than the first decile of the UKBB imaging subset.

Table 1
ADNI and UKBB training and control sets demographics.

	ADNI Training	ADNI Control	UKBB Training	UKBB Control	UKBB All
Total	632	163	616	262	36,494
CN/SMC/Early MCI/Late MCI/AD	75/31/244/164/118	108/55/0/0/0	NA	NA	NA
Female/Male	323/309	78/85	319/297	117/145	19,351/17,143
APOE-4 (0/1/2/NA)	309/236/76/0	127/33/3/0	188/77/294/10	262/0/0/0	21,377/7662/662/1730
APOE-2 (0/1/2/NA)	NA	NA	530/39/3/10	217/44/1/0	26,490/4642/169/1730
AGE (min/med/max)	55/72/90	56/71/89	48/68/81	49/65/77	45/65/82
Amyloid (neg/pos/NA)	174/366/92	163/0/0	NA	NA	NA

The ‘AD at-risk’ training population was defined as the union of two sets: (1) subjects with high genetic risk (i.e., at least one family member with dementia and APOE- ϵ 4 homozygotes); (2) subjects with high cognitive risk according to the following criteria: (i) number of errors made in trail making is larger than four (which is the 90th percentile) and (ii) not being able to correctly recall on the first or the second attempt in the UKBB PMT, plus self-reported AD subjects and excluding those with non-AD self-reported neurological or psychiatric disorders.

2.4. Data harmonization and data pre-processing

2.4.1. Data harmonization

Applying machine learning models, which were trained on one training dataset, to a new application dataset that was generated under different conditions (scanner hardware, scanning protocols, and scan site etc.) is hampered by a change in distribution in the derived features. In our case, both ADNI and UKBB brain T1-weighted MRI data were acquired on 3T scanners. However, the ADNI data was processed with FreeSurfer 5.1, while the UKBB data was processed with FreeSurfer 6.0 and FreeSurfer 7.0. Both datasets provide the same 80 cortical and subcortical volume metrics. In addition, there are biological differences, e.g., the ADNI cohort is older on average (see Table 1 for more details on the demographics). Our first goal was to harmonize the 80 cortical and subcortical volume metrics with the aim to remove the differences in cohorts due to variations in acquisition and processing and to preserve the biological information. Of note, the ADNI cohort features different sites and scanners, and previous work demonstrated that harmonization across sites improves machine learning models (Chen et al., 2020).

We selected the ComBat harmonization technique to eliminate the cohort effect after following a similar analysis suggested by Fortin et al. (2018) (see Supplementary materials).

2.4.2. Data pre-processing

The computational time complexity of the SuStaIn algorithm scales exponentially with the number of input features. Thus, to make computations feasible, the 80 cortical and subcortical regional volumes from the Desikan-Killiany atlas were grouped to derive the volume of 13 regions (accumbens area, amygdala, caudate, hippocampus, insula, pallidum, putamen, thalamus, cingulate, frontal lobe, parietal lobe, temporal lobe, occipital lobe) according to the lobe groupings provided by FreeSurferWiki (<https://surfer.nmr.mgh.harvard.edu/fswiki/CorticalParcellation>). For each of the 13 regional volume variables, we removed the effect of age, sex and ICV using a linear regression model. Then we expressed the values of these variables as z-scores relative to the cohort’s control population (see above). The z-scores’ signs were adjusted such that increases correspond to increasing pathology.

2.5. Subtype and stage inference (SuStaIn)

The SuStaIn algorithm, introduced by Young et al. (2018), uses unsupervised learning to uncover data-driven disease subtypes with distinct temporal progression patterns. SuStaIn is a continuous generalisation of the event-based model (EBM; Fonteijn et al., 2012; Young et al., 2014). While the EBM describes disease progression as a single event sequence

(E_1, E_2, \dots, E_N) where each event is defined as the value of a biomarker turning from normal to abnormal, SuStaIn allows for multiple event sequences. Each event is the value of a biomarker reaching/exceeding a z-score cut-off, each sequence defines a disease progression subtype, and the position in an event sequence defines disease progression stage as accumulated events. During training, SuStaIn simultaneously finds the optimal solutions of subtype membership, subtype trajectory and the posterior distributions of both. In brief, one can think of SuStaIn as a clustering algorithm that takes the progressive nature of the disease into account.

The uncertainty in the subtype progression patterns and the proportion of individuals belonging to each subtypes is estimated using a Markov Chain Monte Carlo (MCMC)-based procedure (Young et al., 2018). The sequence, or event order, is visualized and examined in a positional variance diagram (PVD) chart. For further mathematical details see the Supplementary Materials and Young et al. (2018) for a detailed description.

Like other clustering machine learning models, the number of subtypes in SuStaIn is a hyperparameter and the optimal number of subtypes is generally unknown beforehand. To find the optimal number of subtypes for a SuStaIn model, we used tenfold cross-validation to compute the Cross-Validation Information Criterion (CVIC) (Gelman et al., 2014), which captures out-of-sample generalizability. We selected the number of subtypes by comparing the values of CVIC as well as inspecting the PVDs.

To assign subjects to subtypes and stages, SuStaIn first chooses the subtype with the highest likelihood (after evaluating the likelihood that a subject belongs to each subtype by summing over all disease stages). Next, it selects the stage with the highest likelihood within that subtype. We expect small subtype maximum likelihoods (i.e., low subtype certainty) for subjects assigned to stage zero or to late stages. The zero-stage subjects are those with no obvious abnormalities in the features, therefore there is not enough signal to confidently support any subtype. Likewise, subjects at late stages of the disease progression have abnormalities distributed across nearly all brain regions in which case the differences in the MRI patterns between subtypes become less distinguishable.

2.6. SuStaIn models

We applied SuStaIn to the ADNI, harmonized ADNI, UKBB and the harmonized UKBB training populations, respectively. The z-score events and the distribution of biomarkers in the four training datasets are listed in Supplementary Table S2. For each of these training populations, the SuStaIn model was fitted up to a maximum of five subtypes and we obtained four SuStaIn models: ADNI SuStaIn, Harmonized ADNI SuStaIn, UKBB SuStaIn and Harmonized UKBB SuStaIn.

2.7. SuStaIn subtyping and staging

For each subject in the ADNI and UKBB datasets, we assigned subjects to subtypes and stages predicted by three SuStaIn models (Fig. 1). For example, for a subject in the ADNI dataset, subtype and stage assignment was computed using ADNI SuStaIn, Harmonized ADNI SuStaIn and

Harmonized UKBB SuStaIn, respectively. To quantify transferability, we measured the consistency of subtype and stage assignments by different models.

2.8. Association studies

In our association study, we focused on individuals with atrophy patterns that can be reliably assigned to a SuStaIn, i.e., non-zero stages and high subtype certainty ($Prob > 0.8$) under a model, and we compared the individuals' subtype and stage assignments under different SuStaIn models. Additionally, we repeated the analysis at a lower subtype certainty ($Prob > 0.5$).

We investigated the hypothesis that AD atrophy subtypes suggested by SuStaIn are associated with risk factors and biomarkers including age, cerebrospinal fluid (CSF) biomarkers and medication history. We applied the Kruskal–Wallis test to study the association between SuStaIn subtypes and age and CSF biomarkers. In addition, we used linear regression to adjust for age and sex when testing for associations between subtype and CSF biomarkers. Chi-Squared Tests, contingency tables and logistic regressions were used to assess the associations of SuStaIn subtypes and cholesterol and high blood pressure medications, before and after controlling for age and sex.

3. Results

Our ADNI sample size consists of 795 subjects in total (183 CN, 86 SMC, 244 early MCI, 164 late MCI and 118 CE at baseline). Of these, 163 amyloid-negative SMC and CN subjects were defined as a control set and the remaining 632 subjects were used for SuStaIn model training. In the UKBB dataset, we have 36,494 subjects with MRI scans. According to our training set and control set criteria, 616 subjects were selected as the training population for SuStaIn model fitting (age ranges from 48 to 81 with a median of 68; 319 females; 288 had at least one family member with dementia and were APOE- ϵ 4 homozygotes; 333 performed poorly in the AD-related cognitive tests); 262 subjects (age ranges from 48 to 77 with a median of 65; 117 females; with no family member reporting dementia and not carrying an APOE- ϵ 4 allele) formed the control set (Table 1).

Our data harmonization analysis showed that all three harmonization techniques performed well in reducing site effects. However, ComBat presented its unique advantage in harmonizing the variances in the features over other harmonization techniques (Figs. S1, S2, Tables S3–S10).

3.1. Three consistent subtypes with distinct atrophy patterns were identified from ADNI and UKBB before and after data harmonization

We first investigated whether harmonization affects the estimated models on the same dataset and whether training using a disease group (ADNI) and the AD at-risk group (UKBB) result in similar SuStaIn models. We summarised the findings in Fig. 2, which shows the PVDs from the SuStaIn models with 13 biomarkers.

The CIVC on the ADNI model suggested three subtypes, in line with the original work (Young et al., 2018). The CIVC on the UKBB model suggested four subtypes (Fig. S1). However, two of the subtypes in the UKBB model were nearly identical (Fig. S2). This was supported by a Bhattacharyya coefficient (BC) of 0.99 between these subtypes (a BC of 1.0 indicates identical distributions). Thus, we also selected three subtypes for the UKBB models. Moreover, having the same number of subtypes simplifies the comparison between ADNI and UKBB subtypes, which is a key objective of this work.

These three subtypes agree with the original analysis by Young et al. (2018): the 'typical' subtype has atrophy starting in the hippocampus and amygdala, atrophy in the 'cortical' subtype originates in the lobes, cingulate and insula, while in the 'subcortical' subtype atrophy is first observed in the pallidum, putamen and caudate.

Further, harmonization did not change the subtype patterns in a dataset as can be seen from the almost identical PVDs for each dataset before and after harmonization in Fig. 2. We also noted that for all subtypes emerging from each of the SuStaIn models, the positional variance, which corresponds to the level of uncertainty, increases as disease progresses to later stages, which will increase staging uncertainty of the model towards the end of the disease trajectory. Fig. 3 depicts the disease progression pattern rendered as brains for each subtype in the Harmonized ADNI SuStaIn and Harmonized UKBB SuStaIn models. Further, Fig. 4 shows the correlation of the most likely biomarker ordering for each subtype between the models trained on ADNI and UKBB data. The correlation was stronger for the typical ($r = 0.85$) and cortical ($r = 0.91$) subtype compared to the subcortical subtype ($r = 0.61$).

3.2. Individual subtype and stage assignments under different sustain models were highly consistent

After observing that the estimated models are similar between datasets and are not affected by harmonization, we investigated whether the subtype and stage assignments are consistent on the subject level. A high degree of agreement in both dimensions indicates a successful model transfer.

3.2.1. High subtype consistency

Over 91% of subject assignments to subtypes are consistent across three models built on different datasets.

Table 2 shows the subtype consistency for ADNI participants between the Harmonized ADNI and Harmonized UKBB models. As a silver standard, we focused on the non-zero-stage population with high certainly subtype assignments (>0.8) under a model. For the 320 subjects that meet the silver standard (i.e., non-zero stage with high certainly subtype assignments) under at least one of the two models, 292 have identical subtype assignments between the two models, leading to a 91% consistency. Requiring the silver standard only for subtyping by the Harmonized ADNI (UKBB) model, the consistency increases to 96% (92%). Further, requiring the silver standard by both models, only one out of 153 participants received a discordant subtype assignment.

Similarly, for UKBB subjects Table 2 suggests a 95% subtype consistency under the two harmonized models. For the 9865 subjects that meet the silver standard under at least one model, 9356 have identical subtype assignments between the two models (95% consistency). Requiring the silver standard only for subtyping by the Harmonized ADNI (UKBB) model, the consistency increases to 98% (95%). Further, requiring the silver standard by both models, only 12 out of 4845 participants received a discordant subtype assignments (99.8% consistency).

Using the lower subtype certainty cutoff at 0.5, doubled and tripled the number of subtyped participants in ADNI and UKBB, respectively. While the consistency decreased by about 10%, it remained above 80% (Table S12).

To further highlight the importance of data harmonization when applying a SuStaIn model trained on an alternative dataset to another, we examined results of UKBB individual subtype and stage assignments under the ADNI SuStaIn model *without* any data harmonization (Table S3). Only 4% of the total population (1449 subjects) were staged as non-zero, and 1.7% (631 subjects) received a non-zero stage with high subtype certainty – compared to 19% (6956 subjects) when using harmonization. Only 424 of 631 (67%) participants received concordant subtype assignments with the UKB SuStaIn model – compared to 6838 of 6956 (98%) when using harmonization. Thus, harmonization results in a larger fraction of successfully subtyped participants by the transferred model and it increases the concordance between subtype assignments by the original and by the transferred models.

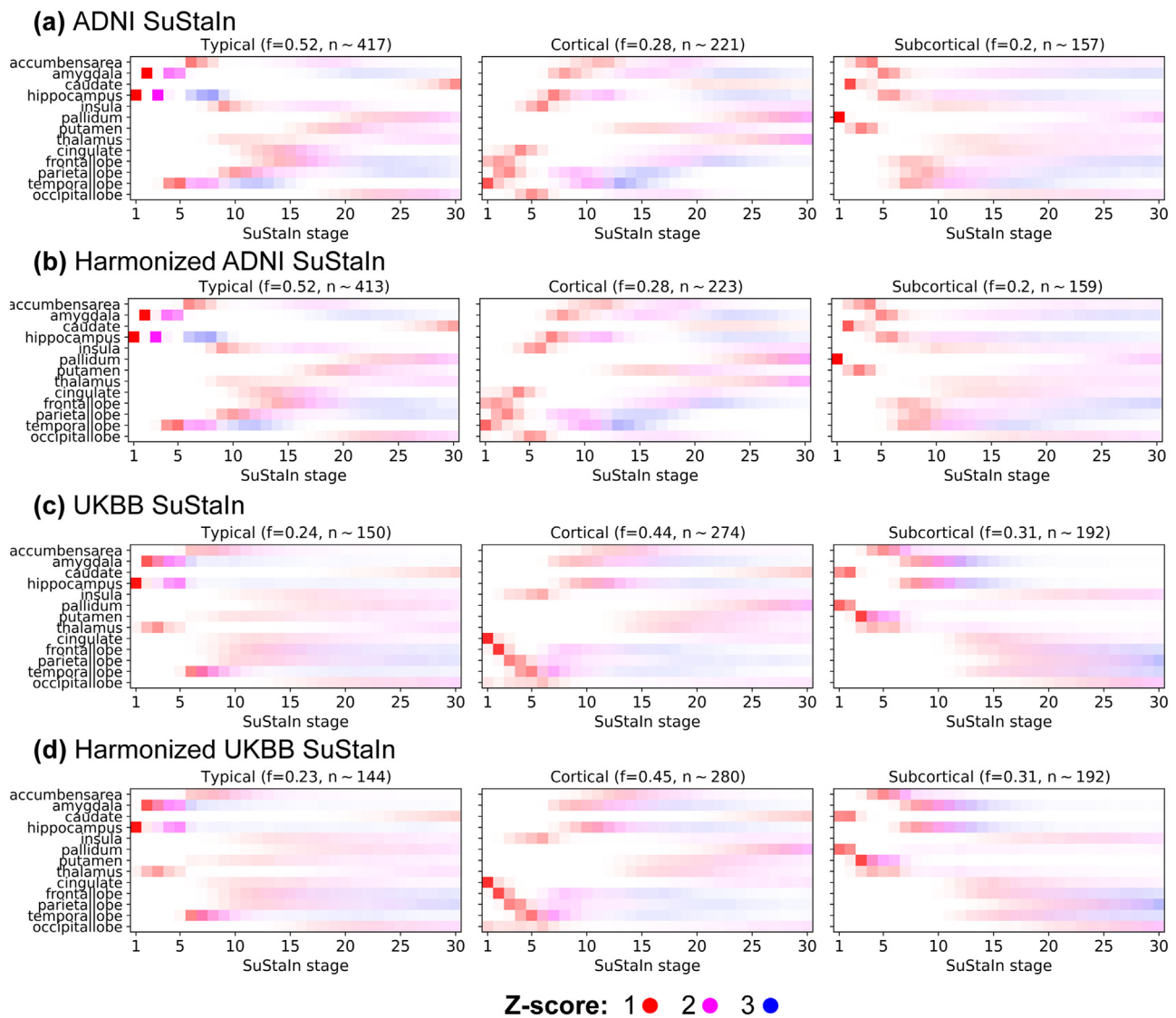


Fig. 2. Three consistent subtypes were identified by all SuStain models. In the Positional Variance Diagrams (PVDs) of the models each row corresponds to a biomarker (regional brain volume) and each coloured entry in a PVD marks the probability that the biomarker has surpassed an event score (here: z-scores) threshold. The figures display z-score thresholds of 1.0 (red), 2.0 (magenta) and 3.0 (blue). While a higher z-score reflects a more severe abnormality of the biomarker, a low positional variance, which is marked in a clearer colour, and is typically observed at the early stage of a diagram, associates with a higher degree of certainty. For instance, in the Harmonized ADNI PVD (b), hippocampal volume reaches high abnormality (z-score > 2, magenta) very early (stage 3) in the typical subtype with high certainty (low positional variance, a clear magenta). The PVDs correspond to different training datasets (ADNI and UKBB) and harmonization options: ADNI (a), harmonized ADNI (b), UKBB (c) and harmonized UKBB (d). We learned from the following four PVDs that: (1) harmonization did not change the subtype patterns in a dataset (see (a) and (b)); (c) and (d)); (2) Indistinguishable patterns between subtypes in the late stages of disease trajectory evidenced by the increasing positional variance as disease progresses to later stages; (3) Focusing on the early stages, three consistent subtypes were identified by all SuStain models: the ‘typical’ subtype has atrophy starting in the hippocampus and amygdala, atrophy in the ‘cortical’ subtype originates in the lobes, cingulate and insula, while in the ‘subcortical’ subtype atrophy is first observed in the pallidum, putamen and caudate .

3.2.2. High stage consistency

Fig. 5 shows the staging consistency across the models for all 795 ADNI and all 36,493 UKBB subjects regardless of their subtype certainty. The harmonization did not influence the staging with the ADNI model ($r^2=0.999$; Fig. 5a left panel) or the UKBB model ($r^2=0.996$; Fig. 5b left panel). Applying harmonized models across datasets led to reduced stage correlations on ADNI ($r^2=0.83$; Fig. 5a right panel) and UKBB ($r^2=0.68$; Fig. 5b right panel). By comparison, using the unharmonized ADNI model to stage UKBB participants the correlation with UKBB SuStain staging drops to $r^2=0.28$. Further, on average ADNI subjects received higher stages by Harmonized UKBB model compared to the Harmonized ADNI model (regression slope > 1), while this observation was reversed for UKBB subjects (regression slope < 1).

In contrast to using all available subjects, Fig. 6 compares the Harmonized ADNI and Harmonized UKBB models for non-zero stage subjects and with high certainty subtype assignments (>0.8) under the Harmonized ADNI Model and suggests comparable consistency between these two models.

3.3. Participants atrophy subtype was found to be associated with age, CSF biomarkers, cholesterol and high blood pressure medications

The previous results demonstrated that harmonization enables us to transfer models across datasets. Next, we investigated how risk factors and biomarkers are associated with subtypes under the harmonized SuStain models. Table 3 shows an overview of the demographics of ADNI

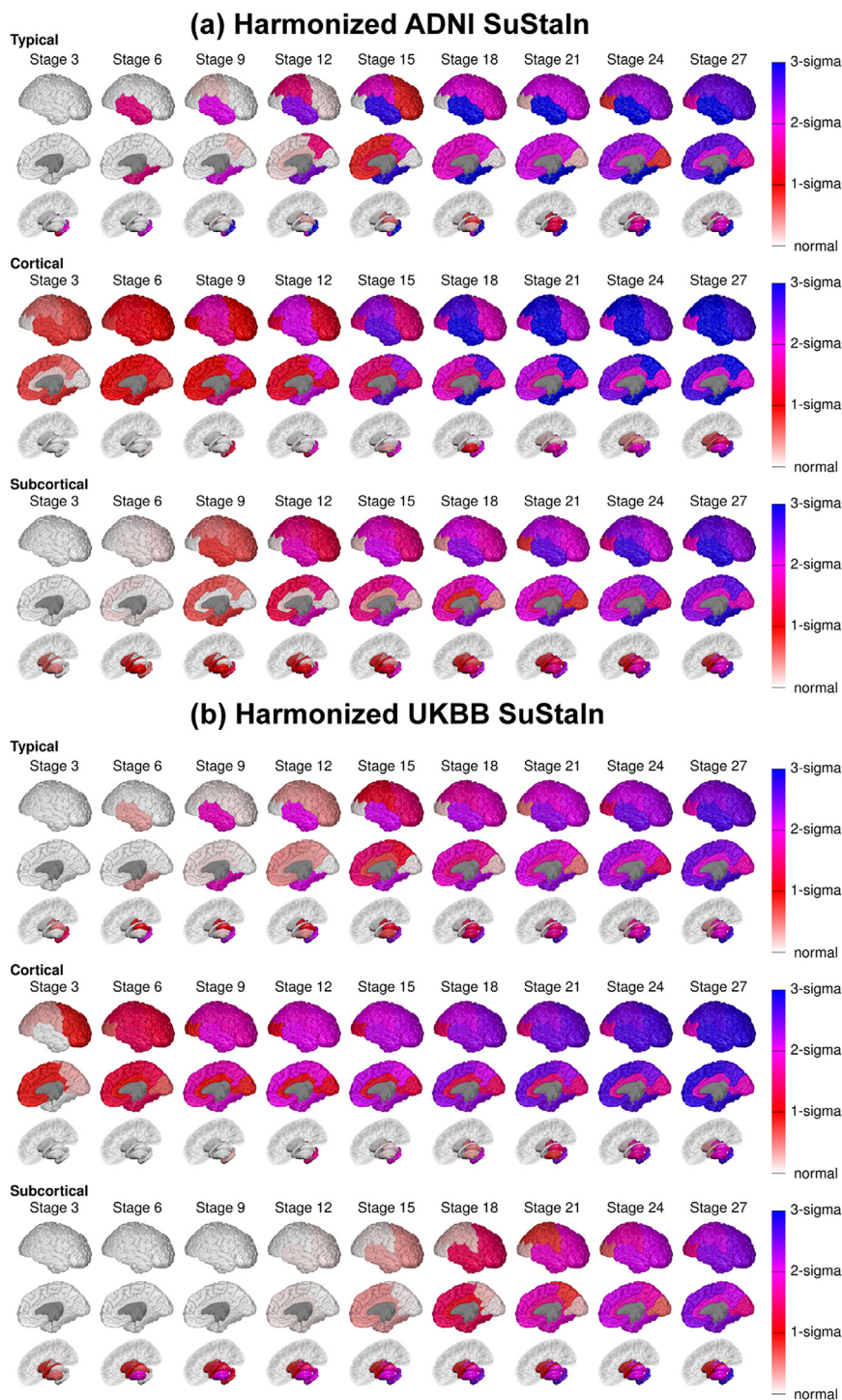


Fig. 3. Representation of the three subtypes in brain space for the Harmonized ADNI SuStain and for the Harmonized UKBB SuStain models.

and UKBB subjects assigned to each subtype that meet the silver standard under the Harmonized ADNI SuStain model (i.e., with non-zero stage and subtype certainty >0.8). In the ADNI dataset, the association study population has 219 subjects and in the UKBB dataset, there are 6956 subjects. At the more lenient subtype certainty cutoff of 0.5, there were 455 and 17,315 subtyped subjects in ADNI and UKBB, respectively.

The average age is highest in the typical subtype and lowest in the subcortical subtype. In both the ADNI and UKBB datasets, there was a statistically significant age difference between the typical and the subcortical subtypes (p-value = 0.001 in ADNI and $p < 0.0001$ in UKBB; Fig. 7 top row), where subjects assigned to the typical subtype were older com-

pared to subjects with the subcortical subtype. Similar results hold for subjects that meet the silver standard under the Harmonized UKBB SuStain model (Fig. 7; bottom row).

Subjects assigned to the typical subtype exhibit lower $A\beta$ and higher tau and p-tau values. We further compared the 185 (82 females) ADNI subjects with CSF measures which consisted of 99 subjects in the typical subtype, 65 in the cortical subtype and 21 in the subcortical subtype. The Kruskal-Wallis H-Test showed significant differences amongst the subtypes and the typical subtype subjects were associated with more AD-like $A\beta$ and tau measures (Fig. 8; top row). This result remained significant after controlling for age and sex in a regression model: $A\beta$

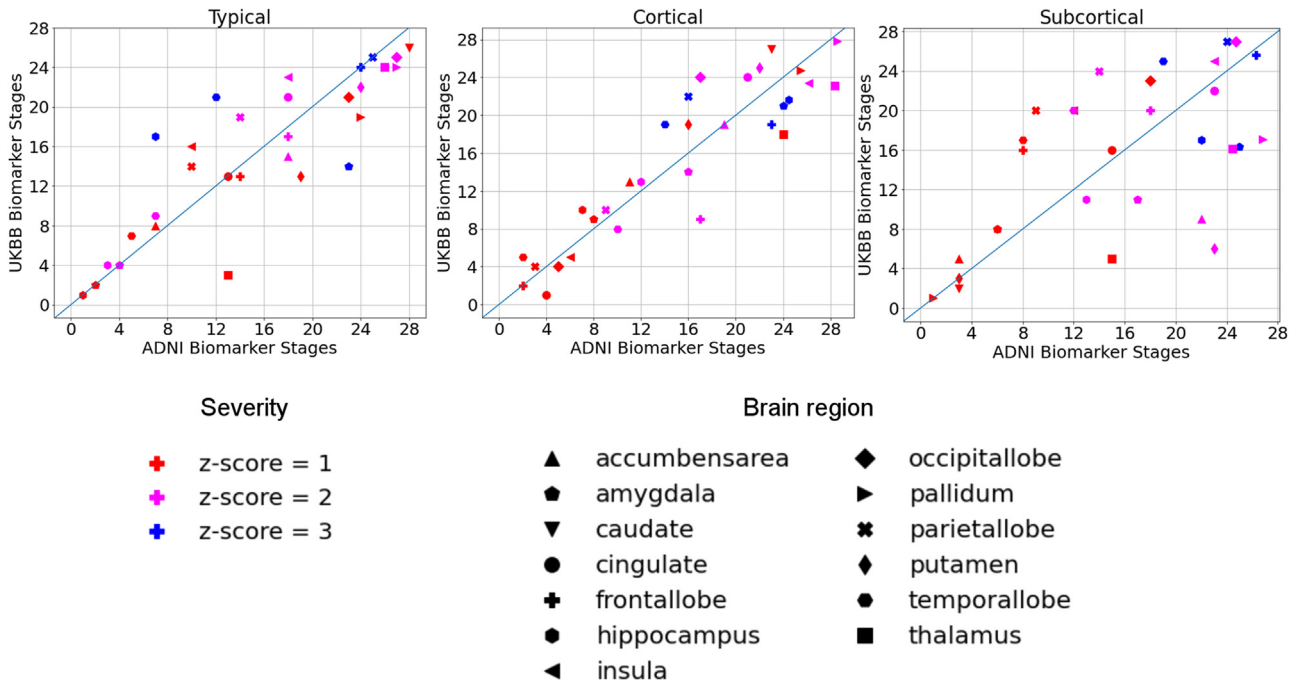


Fig. 4. Agreement between the biomarker ordering for each subtype between the SuStain models trained on ADNI and UKBB data. In the scatter plots, each mark represented the expected values of the stage of a biomarker at a particular z-score (1.0: red, 2.0: magenta and 3.0: blue) in the Harmonized ADNI (x-axis) and Harmonized UKBB (y-axis) models for a particular subtype. Different shapes correspond to different brain regions. The expected value of the stage was calculated as the expectation of the stage under the subtype model: $\bar{k} = \sum_1^k k p(k)$. As an example, the first chart shows that for the typical subtype, amygdala reaching a z-score of 1 at stage 2 in both the Harmonized ADNI and the Harmonized UKBB models, and reaching a z-score of 2 at stage 4 in both models and a z-score of 3 at stage 23 in the Harmonized ADNI model and at stage 14 in the Harmonized UKBB model. The high correlation between stage orderings (i.e., close to the diagonal line) indicates strong agreement between the models for the typical and cortical subtypes, and weaker for the subcortical subtype.

Table 2
Consistency of subject-level subtype assignments under different sustain models.

		Typical		Cortical		Subcortical		
		silver	ML	silver	ML	silver	ML	
UKBB	Typical	ADNI						
		silver	75	0	0	0	0	
	ML	41	–	0	–	0	–	
	Cortical	silver	1	9	64	54	0	7
		ML	3	–	9	–	1	–
	Subcortical	silver	0	3	0	0	13	18
ML		4	–	0	–	11	–	
ADNI	Typical	UKBB						
		silver	429	359	0	22	0	3
	ML	106	–	152	–	22	–	
	Cortical	silver	0	1	2732	908	0	0
		ML	1	–	1638	–	0	–
	Subcortical	silver	0	9	12	66	1672	743
ML		0	–	221	–	769	–	

Columns indicate the subtypes assigned by the reference model, i.e., the model trained on the dataset where it is applied to. Rows indicate the subtypes assigned by the transferred model (e.g., the model trained on UKBB but applied to ADNI data). Subtype assignments are separated into whether they satisfy the ‘silver’ standard (subtype *Prob* > 0.8; silver) or whether they are simply the maximum likelihood assignment (ML). Bold font along the diagonals indicate agreement by the two models.

Table 3
Demographics of ADNI and UKBB subjects assigned to each subtype under the Harmonized ADNI model with non-zero stage and high subtype certainty.

Dataset	ADNI (219)			UKBB (6956)		
	Typical	Cortical	Subcortical	Typical	Cortical	Subcortical
Total	119	75	25	813	3641	2502
CN/SMC/Early MCI/Late MCI/AD	6/2/24/44/43	18/8/15/19/15	5/3/8/6/3	NA	NA	NA
Female/Male	51/68	39/36	9/16	285/528	2399/1242	1334/1168
APOE-4 (0/1/2/NA)	43/48/26/0	45/18/12/0	18/6/1/0	451/183/36/143	2117/764/63/697	1444/565/49/444
APOE-2 (0/1/2/NA)	NA	NA	NA	604/93/6/110	2601/472/16/552	1881/274/15/332
AGE(min/med/max)	56/74/89	55/72/85	55/69/79	46/66/80	46/66/81	45/61/80
Amyloid (neg/pos/NA)	19/75/25	31/32/12	17/07/1	NA	NA	NA

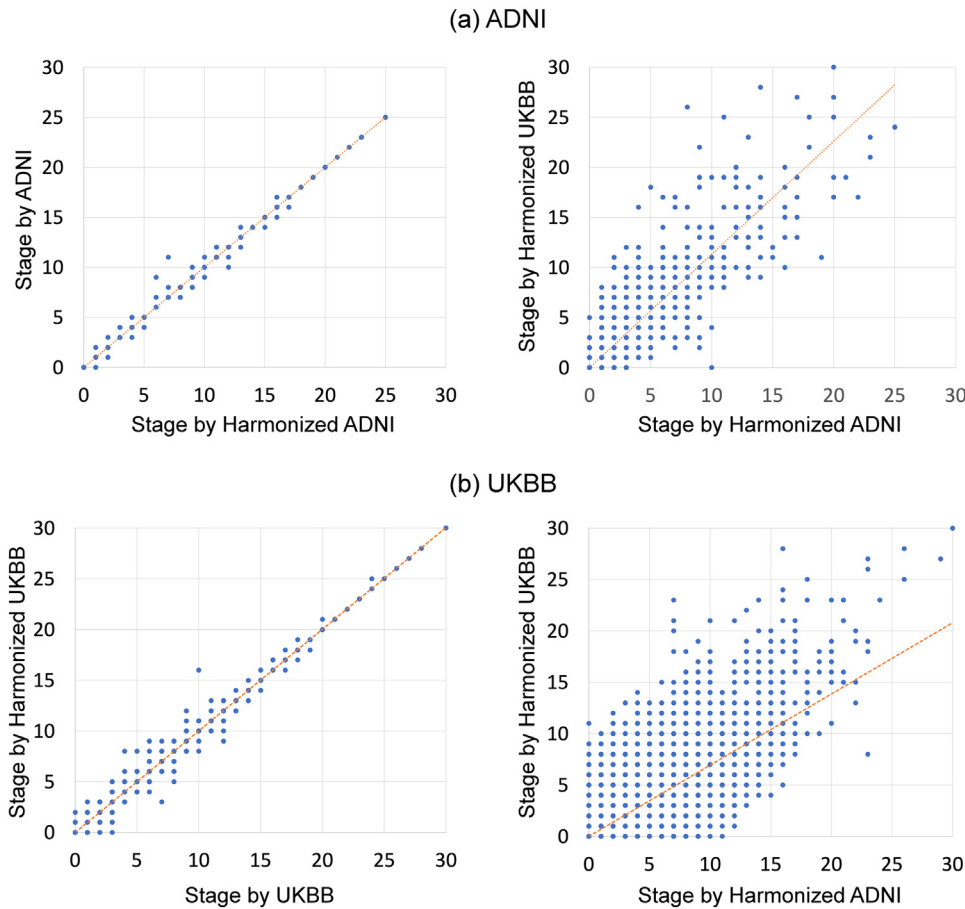


Fig. 5. Subject-level stage consistency across different *SuStaln* Models. The left panel in (a) shows that the subject-level stage assignments for all 795 ADNI subjects under the ADNI and Harmonized ADNI model were almost identical: the dots in the scatter plot distribute closely around the 45-degree line that goes through the origin, with an OLS regression coefficient being 0.9999 and a r-square of 99.9%. The right panel in (a) shows the correlation between stages provided the Harmonized ADNI and Harmonized UKBB models, with an OLS regression coefficient being 1.1297 and a r-square of 83.0%. Although the individuals, especially those in the late stages, were assigned to a slightly higher stage in the Harmonized UKBB model compared to in the Harmonized ADNI model (resulting in a coefficient being larger than one in the OLS regression line). Similarly, the left panel in (b) shows that the subject-level stage assignments for all 36,493 UKBB subjects under the UKBB and Harmonized UKBB model were almost identical: the dots in the scatter plot distribute closely around the 45-degree line that goes through the origin, with an OLS regression coefficient being 1.0016 and an r-square of 99.6%. The right panel in (b) shows the alignment of stages between the Harmonized ADNI and Harmonized UKBB models, with an r-square of 68.1% and an OLS coefficient of 0.6932.

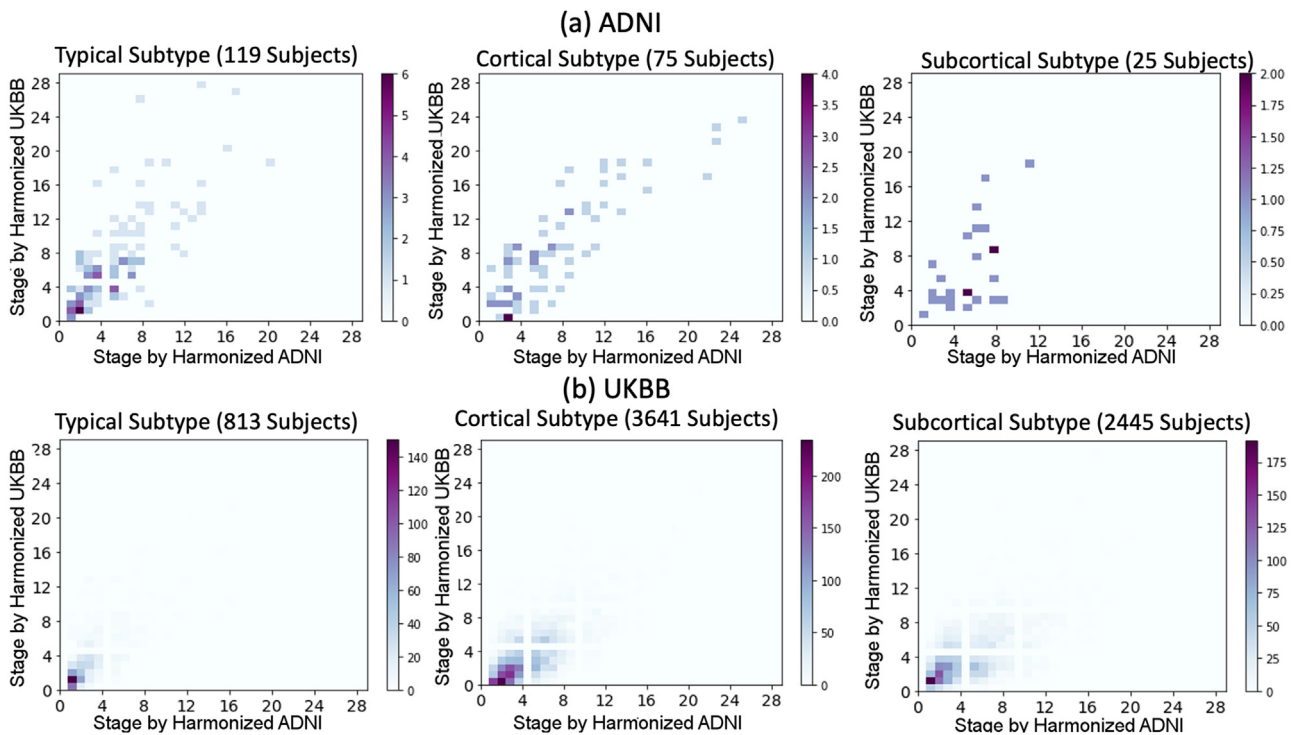


Fig. 6. Subject-level stage consistency between the Harmonized ADNI *SuStaln* and the Harmonized UKBB *SuStaln* Models stratified by subtype. Density plots grouped by subtypes for ADNI (219 subjects) and UKBB (6956) non zero-stage subjects with high certainty (>0.8) subtype under the Harmonized ADNI Model: The majority of the subjects have consistent stage assignments under the Harmonized ADNI *SuStaln* and Harmonized UKBB *SuStaln* models, especially those at early stages.

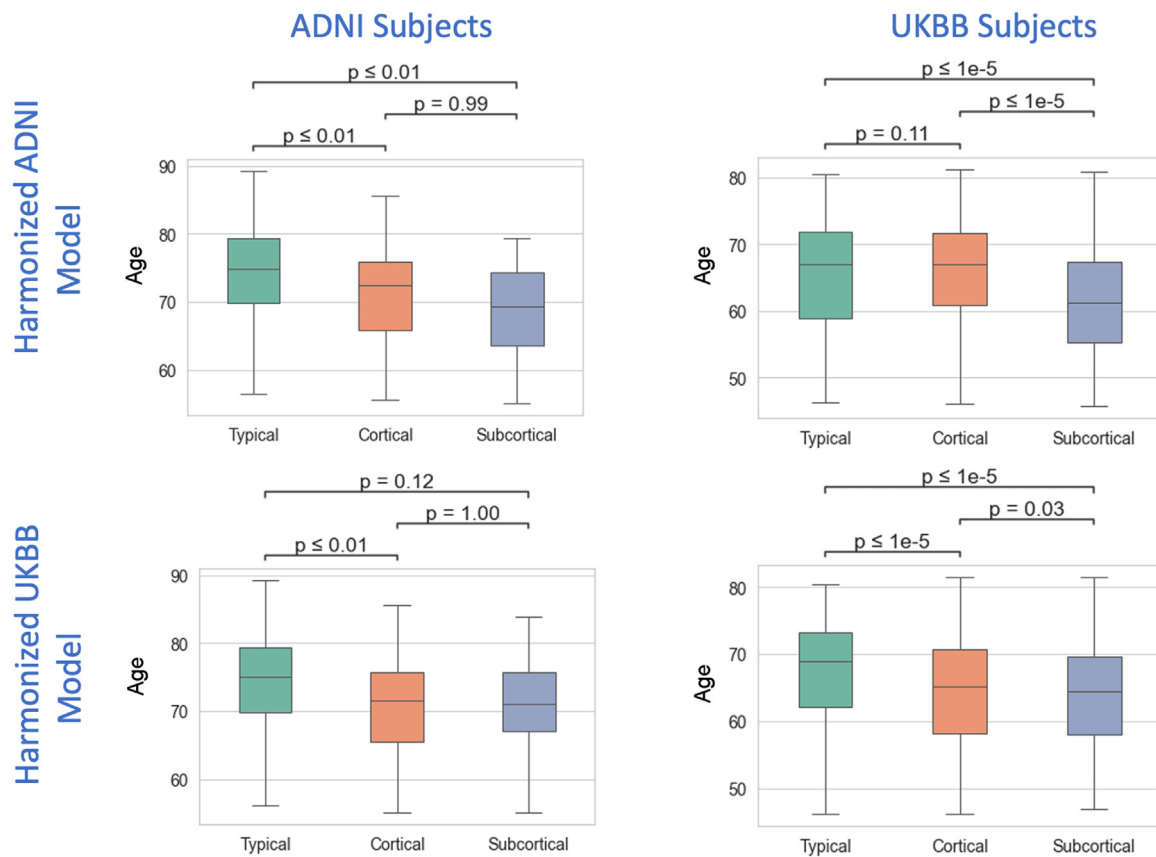


Fig. 7. The average age is highest in the typical subtype and lowest in the subcortical subtype. Subtype and age association for ADNI subjects (left column) and UKBB subjects (right column) meeting the silver standard under the Harmonized ADNI model (top row; ADNI: 219 in total: 119 typical, 75 cortical and 25 subcortical; UKBB: 6956 in total: 813 typical, 3641 cortical and 2502 subcortical) and the Harmonized UKBB model (bottom row; ADNI: 251 in total: 95 typical, 118 cortical and 38 subcortical; UKBB: 7754 in total: 536 typical, 4755 cortical and 2463 subcortical) respectively: the average age is highest in the typical subtype and lowest in the subcortical subtype.

($p = 0.014$), tau ($p = 0.026$) and p -tau ($p = 0.001$). Conclusions remained unchanged when using the Harmonized UKBB SuStaln model to subtype the ADNI subjects (Fig. 8; bottom row). Furthermore, the biomarker associations were not sensitive to the subtype certainty threshold (Table S13).

In comparison to the subcortical subtype, the cortical subtype subjects are more likely to associate with cholesterol and high blood pressure medications. SuStaln subtypes (as assigned by the Harmonized ADNI model) differed both in cholesterol and high blood pressure medications (Table 4; $p < 0.0001$). Specifically, using regression models with adjustment for age and sex, compared to the cortical subtype, the subcortical subtype demonstrated a lower proportion using both medications (cholesterol: $p = 0.005$, OR = 0.82; blood pressure $p < 0.0001$, OR = 0.75). The same holds true for the typical subtype compared to the cortical subtype (cholesterol: $p = 0.012$, OR = 0.80; blood pressure $p = 0.001$, OR = 0.74). Similar observations were drawn using the Harmonized UKBB model for subtyping (Table 4). However, after controlling for age and sex, the associations were only significant for blood pressure medication in the subcortical subtype ($p < 0.0001$, OR = 0.32). Again, the subtype certainty threshold did not affect these conclusions (Table S14).

4. Discussion

In this work, we presented results of building subtyping and staging models across ADNI and UKBB that represent different time windows during the disease to better investigate the link between AD atrophy subtypes and early risk factors. We proposed that a subtype model (i.e., a SuStaln model) is first built on the dataset with detailed disease-related

information (e.g., ADNI), and then applied to the UKBB dataset for individual subtyping and investigating of early risk factor associations.

In previous studies, it was pointed out that ADNI participants were not representative of the wider general population so that data-driven studies on ADNI data may not necessarily ensure generalizability (Veitch et al., 2021). Further, the heterogeneity in cohort datasets in the reproducibility of data-driven results has been discussed (Birkenbihl et al., 2021). Our findings in this study answered many practical questions raised in training and transferring data driven disease progression models across datasets: (1) consistent subtypes emerged from the ADNI cohort and the UKBB AD-at-risk cohort, (2) ComBat, a data harmonization technique, helped to remove cohort effects on biomarkers for these two datasets, and (3) a model trained on ADNI can be robustly applied to UKBB after appropriate data harmonization.

Across various pre-processing settings and datasets SuStaln identified three subtypes characterised by distinct temporal patterns of grey matter volume changes. These subtypes differ in their origin of atrophy and are consistent with previous analyses on the ADNI dataset (Young et al., 2018). The subtypes are referred to by their characteristic pattern of temporal progression as 'typical', 'cortical' and 'subcortical'. The three subtypes were explored in depth in their cognitive and biomarker profile in research datasets as well as in clinical datasets (Archetti et al., 2021; Young et al., 2018).

One of our interesting findings is that the SuStaln model built on the AD-at-risk population constructed from the UKBB dataset resulted in the same atrophy subtypes as those on ADNI; obviously with a different prevalence between the cohorts. This is remarkable, since there were no subjects with a clinical AD diagnosis among the UKBB training set.

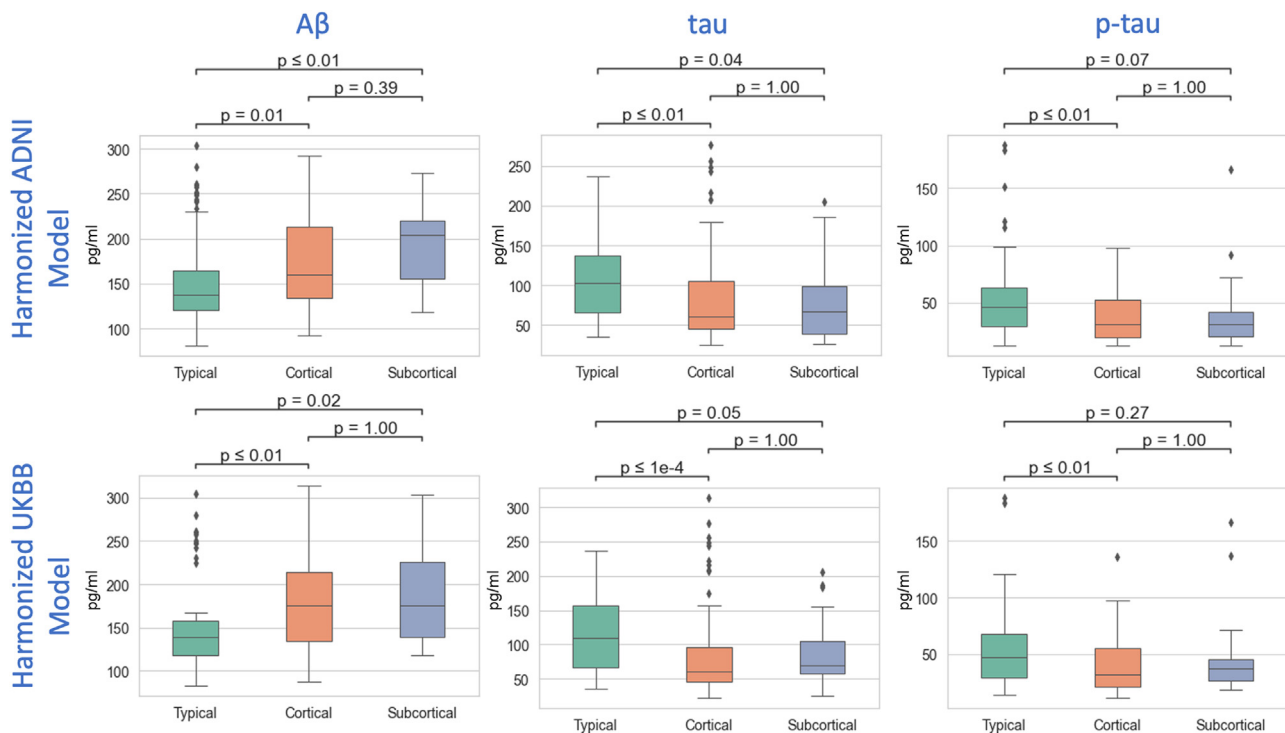


Fig. 8. The typical subtype subjects are associated with statistically worse CSF biomarkers ($A\beta$, tau and p-tau) values in comparison to the cortical subtype. Subtype and CSF biomarkers association for ADNI with CSF measures at baseline and meeting the silver standard under the Harmonized ADNI SuStain model (top row; 185 in total; 99 typical, 65 cortical and 21 subcortical) and under the Harmonized UKBB SuStain model (bottom row; 203 in total; 64 typical, 113 cortical and 26 subcortical), respectively. The typical subtype subjects are associated with lower $A\beta$ and higher tau and p-tau values in comparison to the cortical subtype.

This finding first supports the subtypes per-se as an existing pattern, and secondly, alludes to the possibility to observe the emergence of different atrophy patterns during the prodromal phase of Alzheimer's disease.

Transferring machine learning models to novel datasets in general is an ongoing challenge due to bias and covariate shifts occurring during data acquisition. It was demonstrated in a recent study that SuStain could be transferred from ADNI to a mixture of clinical and pre-clinical cohorts including OASIS, Pharma-Cog and VITA (Archetti et al., 2021). Different to OASIS, a relatively small disease cohort that covers the full spectrum of the disease progression like ADNI, UKBB is a large dataset that represents a much earlier time window even before the disease starts to emerge. Thus, our work suggests a possible route to further investigating the link between AD atrophy subtypes and early risk factors that are collected in large population cohorts such as UKBB.

To facilitate the analysis across ADNI and UK Biobank, we applied ComBat, a data harmonization technique to the two datasets and we demonstrated that harmonization is important to successfully transfer the trained models. In our harmonization analysis, we studied the impact of different cohorts (ADNI vs UKBB) where cohort differences originate from biological variation (age, disease status etc.) and technical variation (scanner manufacturer, scanning protocol, processing pipelines). Since the FreeSurfer version and cohort have a one-to-one mapping relationship in this study, the impact of differing FreeSurfer versions is addressed as part of our data harmonization study – see details in the supplementary materials (e.g., Figs S3, S4 and Table S4–S11).

Adding harmonization to the biomarker pre-processing pipeline did not affect the estimated models: models estimated on the same dataset before and after harmonization where we found almost identical (99% similarity). This was expected since SuStain converts biomarker values to z-scores with reference to a control population. Thus, any shift or scaling applied to each feature will be removed and the resulting models are identical. However, the harmonization removes cohort-specific bias and therefore mainly enables the successful transfer of the SuStain

model to external datasets, according to our definition of a successful transfer as outlined in the Introduction section. We further showed that harmonization was essential to achieve a high consistency in individuals' subtypes and stage assignment based on the different models. While subject-level subtype consistency was generally very high across experiments (>91%), the correlation of the subject-level stages was lower (e.g., $r^2=0.68$ on the UKBB data) but was a considerable improvement from using unharmonized models ($r^2=0.28$). The lower staging agreement has two main sources. Firstly, although the progression sequences estimated for each subtype from ADNI and UKBB were similar, they were not identical (Fig. 4). Secondly, discordant subtyping by models negatively impacts the agreement between staging.

A further alternative is to combine the harmonized datasets prior to model training. Such a model could potentially add value to situations where there are not enough early-stage subjects in the disease cohort. Indeed, when testing this approach, we observed that (1) the three subtypes are consistent with other models, and (2) stage assignments are highly consistent with both the Harmonized ADNI SuStain and Harmonized UKBB models (data not shown).

Linking the patterns of MRI abnormality in the subtypes to clinical phenotypes, other biomarkers, comorbidities, genetic and life-style factors could provide new insights into disease mechanisms. Thus, understanding the association of these factors with subtypes emergence may aid in stratifying patients at early stages for precision medicine and multidomain interventional studies such as those explored in the FINGER trial (Ngandu et al., 2015). We investigated the associations of subtypes with a few selected variables from ADNI and from UK Biobank. On ADNI we focused on associations with other known AD-related biomarkers such as levels of $A\beta$ and tau in the CSF. Our results suggested that the typical subtype was associated with statistically worse CSF biomarkers values (i.e., more AD-like) in comparison to the other two subtypes, before and after controlling for age and sex. This finding is in line with previous observations that MCI subjects assigned to the typical subtype

Table 4

The cortical subtype subjects are more likely to associate with cholesterol and high blood pressure medications compared to the subcortical subtype. Results remain significant after controlling for age and sex. For the non zero-stage UKBB subjects with high subtype certainty (>0.8) under the Harmonized ADNI model (6956 subjects), there were significant differences between SuStaln subtypes in both cholesterol and high blood pressure medications ($p<0.001$). Specifically, using regression models with adjustment for age and sex, compared to the cortical subtype, the subcortical subtype has the lower proportion associated with both medications (cholesterol: $p = 0.005$; blood pressure $p<0.001$). The same held true for the typical subtype compared to the cortical subtype (cholesterol: $p = 0.012$; blood pressure $p<0.001$). Some of these observations were drawn using the Harmonized UKBB and UKBB models although after controlling for age and sex, the associations were only significant for blood pressure medication in the subcortical subtype ($p<0.001$).

		Chi-Squared Tests and Contingency Table						Logistic Regression				
Model	Sample Size	P-Value	Status	Typical	Cortical	Subcortical		Typical	Subcortical	AGE	SEX	
Cholesterol Lowering	Harmonized ADNI	6956	<0.001	0	597	2703	2038	coef	-0.22	-0.20	0.10	1.06
				1	216	938	464	p-values	0.012	0.005	<0.001	<0.001
	SuStaln				27%	26%	19%					
Blood Pressure			<0.001	0	608	2643	2022	coef	-0.30	-0.29	0.07	0.69
				1	205	998	480	p-values	0.001	<0.001	<0.001	<0.001
					25%	27%	19%					
Cholesterol Lowering	Harmonized UKBB	9482	<0.001	0	381	4537	2372	coef	0.00	-0.11	0.10	0.96
				1	153	1475	564	p-values	0.98	0.09	<0.001	<0.001
	SuStaln				29%	25%	19%					
Blood Pressure			<0.001	0	385	4454	2375	coef	-0.12	-0.33	0.07	0.65
				1	149	1558	561	p-values	0.23	<0.001	<0.001	<0.001
					28%	26%	19%					
Cholesterol Lowering	UKBB SuStaln	7669	<0.001	0	384	3423	1968	coef	0.02	-0.10	0.09	0.97
				1	185	1174	535	p-values	0.83	0.12	<0.001	<0.001
					33%	26%	21%					
Blood Pressure			<0.001	0	392	3315	1992	coef	-0.10	-0.32	0.07	0.67
				1	177	1282	511	p-values	0.32	<0.001	<0.001	<0.001
					31%	28%	20%					

Significant results related to subtype assignments are highlighted in bold font. The column 'status' indicates the status of medication usage (0=no, 1=yes). Percentages refer to the medication users per subtype.

have a highest risk to progress from MCI to AD (Young et al., 2018). In UKBB, we focused on variables that may contribute to the subtype emergence but are not readily available in ADNI. Therefore, showcasing the benefit of transferring SuStaln models. Our results suggested that the subcortical subtype was associated with a statistically lower likelihood of cholesterol and high blood pressure medication use in comparison to the cortical subtype, before and after controlling for age and sex.

While early hippocampal involvement in the typical subtype may explain more severe CSF biomarkers pathology, the reason for participants assigned to the cortical subtype having higher likelihood of cholesterol and high blood pressure medications use is unclear. Previous studies have demonstrated a link between cholesterol and AD pathology (Di Paolo and Kim, 2011) and statins, a class of cholesterol lowering drugs, are routinely used in the treatment of AD patients. Likewise, mid-life elevated blood pressure was associated with increased atrophy in brain regions linked to AD (Lane et al., 2019; Power et al., 2016). More broadly, hypertension has been identified as one of the modifiable mid-life risk factors contributing to AD (Livingston et al., 2020). Potential explanations for the linkages between cholesterol and high blood pressure and AD subtypes could be that hypertension is associated with specific lower regional grey matter volume (Gianaros et al., 2006; Schaare et al., 2019). Further studies are needed for a better understanding of these associations.

A key limitation in this work is that, due to the lack of such information in the dataset, for models built in UKBB, we were unable to use AD diagnosis (CN, MCI or AD), AD biomarkers such as CSF and PET measurements, or cognitive tests tailored for diagnosing AD in the construction of control and training sets as well as in the subtype evaluation. We used variables that are highly correlated with the disease including the numbers of APOE- ϵ 4 allele, family history of AD cognitive tests available to identify groups that are least and most likely to develop the disease as the control and training sets, respectively. We also excluded subjects with non-AD self-reported neurological or psychiatric disorders. Also, we transferred the models only between two datasets. Future work on transferring SuStaln to additional datasets would help to generalize the findings from this study.

Another limitation is that all the harmonization methods examined and employed in this work attempt to capture mostly the linear relationship between cohorts and volume metrics. Improvements over ComBat that consider the co-variance of features have been proposed (e.g., CovBat (Chen et al., 2020)). Furthermore, while ComBat demonstrated strong efficacy in removing cohort effect for data in ADNI and UKBB, it does not support an efficient extension of the analysis to new datasets. In practice, when applying SuStaln built on ADNI to a new dataset, one would need to repeat the process of first harmonizing the new dataset with ADNI. Future study of harmonization methods could include Hierarchical Bayesian Regression (HBR) (Kia et al., 2020) which allows applying a SuStaln model on new dataset without repeating model fitting processes.

While the methodology presented in this paper was based on volume metrics, in principle following the same methodology, one can generalise to other brain morphological measures (e.g., thickness metrics) and to non-imaging metrics. In fact, the work by Fortin et al. applied ComBat to harmonize cortical thickness (Fortin et al., 2018) and diffusion tensor imaging metrics (Fortin et al., 2017). Thus, if such features were useful within a SuStaln model, then they could be successfully harmonized.

In conclusion, to study the association of AD atrophy subtypes with a broad range of risk factors, we proposed to harmonize the disease cohort and the aging cohort using ComBat, and to apply the SuStaln model on the disease cohort which is used to subtyping and staging subjects in the aging cohort. Our methodology enabled further detailed investigations on the linkage of clinicopathological late-stage to earlier risk factors, which has the potential to lead to a better understanding of the disease aetiology and to help avoid lifestyle and behaviour choices that puts individuals at risk.

Data and code availability statement

The data used in this work are available upon application to the Alzheimer's disease neuroimaging imitative (ADNI; <http://adni.loni.usc.edu/>) and the UK BioBank (UKBB; <https://www.ukbiobank.ac.uk/>). The SuStaln algorithm implemented in Python (pySuStaln) is open source and available at

<https://github.com/ucl-pond/pySuStaIn>. Further analysis scripts are available upon request.

Funding

This study was supported by the Early Detection of Alzheimer's Disease Subtypes (E-DADS) project, an EU Joint Programme - Neurodegenerative Disease Research (JPND) project (see www.jpnd.eu). The project is supported under the aegis of JPND through the following funding organizations: United Kingdom, Medical Research Council (MR/T046422/1); Netherlands, ZonMW (733051106); France, Agence Nationale de la Recherche (ANR-19-JPW2-000); Italy, Italian Ministry of Health (MoH); Australia, National Health & Medical Research Council (1191535); Hungary, National Research, Development and Innovation Office (2019-2.1.7-ERA-NET-2020-00008). This study was also supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre and EPSRC grant EP/M020533/1. A.L.Y. is supported by a Skills Development Fellowship from the Medical Research Council (MR/T027800/1). F.B. is a National Institute of Health Research (NIHR) Professor. N.P.O. is a UKRI Future Leaders Fellow (MR/S03546X/1). A.A. holds an MRC eMedLab Medical Bioinformatics Career Development Fellowship. This work was partly supported by the [Medical Research Council \(MR/L016311/1\)](http://www.mrc.ac.uk).

Declaration of Competing Interest

F.B. acts as a consultant for Biogen-Idec, IXICO, Merck-Serono, Novartis, Combinostics and Roche. He has received grants, or grants are pending, from the Amyloid Imaging to Prevent Alzheimer's Disease (AMYPAD) initiative, the Biomedical Research Centre at University College London Hospitals, the Dutch MS Society, ECTRIMS-MAGNIMS, EU-H2020, the Dutch Research Council (NWO), the UK MS Society, and the National Institute for Health Research, University College London. He has received payments for the development of educational presentations from Ixico and his institution from Biogen-Idec and Merck. He is on the editorial board of Radiology, European Neuroradiology, Multiple Sclerosis Journal and Neurology. F.B. and D.C.A. are on the board of directors of Queen Square Analytics.

Credit authorship contribution statement

Hanyi Chen: Conceptualization, Methodology, Formal analysis, Writing – original draft. **Alexandra Young:** Software, Visualization, Writing – review & editing. **Neil P. Oxtoby:** Writing – review & editing, Funding acquisition. **Frederik Barkhof:** Writing – review & editing, Funding acquisition. **Daniel C. Alexander:** Conceptualization, Writing – review & editing, Funding acquisition. **Andre Altmann:** Conceptualization, Formal analysis, Writing – original draft, Supervision, Funding acquisition.

Data availability

The authors do not have permission to share data.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.neuroimage.2023.120005](https://doi.org/10.1016/j.neuroimage.2023.120005).

References

Aksman L.M., Oxtoby N.P., Scelsi M.A., Wijeratne P.A., Young A.L., Lopes Alves I., Barkhof F., Alexander D.C., & Altmann A. (2020). Tau-first subtype of Alzheimer's disease consistently identified across *in vivo* and post mortem studies. *BioRxiv*, 2020.12.18.418004. doi:10.1101/2020.12.18.418004.

Alfaro-Almagro, F., Jenkinson, M., Bangerter, N.K., Andersson, J.L.R., Griffanti, L., Douaud, G., Sotiropoulos, S.N., Jbabdi, S., Hernandez-Fernandez, M., Vallee, E., Viddal, D., Webster, M., McCarthy, P., Rorden, C., Daducci, A., Alexander, D.C.,

Zhang, H., Dragonu, I., Matthews, P.M., Smith, S.M., 2018. Image processing and quality control for the first 10,000 brain imaging datasets from UK Biobank. *Neuroimage* 166, 400–424. doi:10.1016/j.neuroimage.2017.10.034.

Andrews, R.M., Shpitser, I., Lopez, O., Longstreth, W.T., Chaves, P.H.M., Kuller, L., Carlson, M.C., 2020. Examining the causal mediating role of brain pathology on the relationship between diabetes and cognitive impairment: the cardiovascular health study. *J. R. Stat. Soc. A Stat. Soc.* 183 (4), 1705–1726. doi:10.1111/rssa.12570.

Archetti, D., Young, A.L., Oxtoby, N.P., Ferreira, D., Mårtensson, G., Westman, E., Alexander, D.C., Frisoni, G.B., Redolfi, A., 2021. Inter-cohort validation of SuStaIn model for Alzheimer's disease. *Front. Big Data* 4, 30. doi:10.3389/fdata.2021.661110.

Bateman, R.J., Xiong, C., Benzinger, T.L.S., Fagan, A.M., Goate, A., Fox, N.C., Marcus, D.S., Cairns, N.J., Xie, X., Blazey, T.M., Holtzman, D.M., Santacruz, A., Buckles, V., Oliver, A., Moulder, K., Aisen, P.S., Ghetti, B., Klunk, W.E., McDade, E., Morris, J.C., 2012. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. *N. Engl. J. Med.* 367 (9), 795–804. doi:10.1056/NEJMoa1202753.

Bilgel, M., Prince, J.L., Wong, D.F., Resnick, S.M., Jernynak, B.M., 2016. A multivariate nonlinear mixed effects model for longitudinal image analysis: application to amyloid imaging. *Neuroimage* 134, 658–670. doi:10.1016/j.neuroimage.2016.04.001.

Birkenbihl, C., Salimi, Y., Fröhlich, H., 2021. Unraveling the heterogeneity in Alzheimer's disease progression across multiple cohorts and the implications for data-driven disease modeling. *Alzheimer's Dement.* 2021. doi:10.1002/alz.12387.

Chen A.A., Beer J.C., Tustison N.J., Cook P.A., Shinohara R.T., Shou H., & Initiative the A.D.N. (2020). Removal of scanner effects in covariance improves multivariate pattern analysis in neuroimaging data. *BioRxiv*, 858415. doi:10.1101/858415.

Di Paolo, G., Kim, T.W., 2011. Linking lipids to Alzheimer's disease: cholesterol and beyond. *Nat. Rev. Neurosci.* 12 (5), 284–296. doi:10.1038/nrn3012.

Donohue, M.C., Jacqmin-Gadda, H., Le Goff, M., Thomas, R.G., Raman, R., Gamst, A.C., Beckett, L.A., Jack Jr., C.R., Weiner, M.W., Dartigues, J.F., Aisen, P.S., Initiative, A.D.N., 2014. Estimating long-term multivariate progression from short-term data. *Alzheimer's Dement. J. Alzheimer's Assoc.* 10 (5 Suppl), S400–S410. doi:10.1016/j.jalz.2013.10.003.

Eshaghi, A., Young, A.L., Wijeratne, P.A., Prados, F., Arnold, D.L., Narayanan, S., Guttman, C.R.G., Barkhof, F., Alexander, D.C., Thompson, A.J., Chard, D., Ciccarelli, O., 2021. Identifying multiple sclerosis subtypes using unsupervised machine learning and MRI data. *Nat. Commun.* 12 (1), 2078. doi:10.1038/s41467-021-22265-2.

Fonteijn, H.M., Modat, M., Clarkon, M.J., Barnes, J., Lehmann, M., Hobbs, N.Z., Scallih, R.L., Tabrizi, S.J., Ourselin, S., Fox, N.C., Alexander, D.C., 2012. An event-based model for disease progression and its application in familial Alzheimer's disease and Huntington's disease. *Neuroimage* 60 (3), 1880–1889. doi:10.1016/j.neuroimage.2012.01.062.

Fortin, J.P., Cullen, N., Sheline, Y.I., Taylor, W.D., Aselcioglu, I., Cook, P.A., Adams, P., Cooper, C., Fava, M., McGrath, P.J., McInnis, M., Phillips, M.L., Trivedi, M.H., Weissman, M.M., Shinohara, R.T., 2018. Harmonization of cortical thickness measurements across scanners and sites. *Neuroimage* 167, 104–120. doi:10.1016/j.neuroimage.2017.11.024.

Fortin, J.-P., Parker, D., Tung, B., Watanabe, T., Elliott, M.A., Ruparel, K., Roalf, D.R., Satterthwaite, T.D., Gur, R.C., Gur, R.E., Schultz, R.T., Verma, R., Shinohara, R.T., 2017. Harmonization of multi-site diffusion tensor imaging data. *Neuroimage* 161, 149–170. doi:10.1016/j.neuroimage.2017.08.047.

Gelman, A., Hwang, J., Vehtari, A., 2014. Understanding predictive information criteria for Bayesian models. *Stat. Comput.* 24 (6), 997–1016. doi:10.1007/s11222-013-9416-2.

Gianaros, P.J., Greer, P.J., Ryan, C.M., Jennings, J.R., 2006. Higher blood pressure predicts lower regional grey matter volume: consequences on short-term information processing. *Neuroimage* 31 (2), 754–765. doi:10.1016/j.neuroimage.2006.01.003.

Hardy, J.A., Higgins, G.A., 1992. Alzheimer's disease: the amyloid cascade hypothesis. *Science* 256 (5054), 184–185. doi:10.1126/science.1566067.

Jack Jr., C.R., Knopman, D.S., Jagust, W.J., Shaw, L.M., Aisen, P.S., Weiner, M.W., Petersen, R.C., Trojanowski, J.Q., 2010. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *Lancet Neurol.* 9 (1), 119–128. doi:10.1016/S1474-4422(09)70299-6.

Johnson, W.E., Li, C., Rabinovic, A., 2007. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8 (1), 118–127. doi:10.1093/biostatistics/kxj037.

Kia, S.M., Huijsdens, H., Dinga, R., Wolfers, T., Mennes, M., Andreassen, O.A., Westlye, L.T., Beckmann, C.F., Marquand, A.F., Martel, A.L., Abolmaesumi, P., Stoyanov, D., Mateus, D., Zuluaga, M.A., Zhou, S.K., Racoceanu, D., Joskovic, L., 2020. Proceedings of the Hierarchical Bayesian Regression for Multi-site Normative Modeling of Neuroimaging Data BT - Medical Image Computing and Computer Assisted Intervention – MICCAI 2020, pp. 699–709 Springer International Publishing.

Koval, I., Bône, A., Louis, M., Lartigue, T., Bottani, S., Marcoux, A., Samper-González, J., Burgos, N., Charlier, B., Bertrand, A., Epelbaum, S., Colliot, O., Allassonnière, S., Durleman, S., 2021. AD course map charts Alzheimer's disease progression. *Sci. Rep.* 11 (1), 8020. doi:10.1038/s41598-021-87434-1.

Kunkle, B.W., Grenier-Boley, B., Sims, R., Bis, J.C., Damotte, V., Naj, A.C., Boland, A., Vronska, M., van der Lee, S.J., Amlie-Wolf, A., Bellenguez, C., Frizzati, A., Chouraki, V., Martin, E.R., Sleegers, K., Badarinarayan, N., Jakobsdottir, J., Hamilton-Nelson, K.L., Moreno-Grau, S., ... Pericak-Vance, M.A., 2019. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates A β , tau, immunity and lipid processing. *Nat. Genet.* 51 (3), 414–430. doi:10.1038/s41588-019-0358-2.

Laforce, R., Tosun, D., Ghosh, P., Lehmann, M., Madison, C.M., Weiner, M.W., Miller, B.L., Jagust, W.J., Rabinovici, G.D., 2014. Parallel ICA of FDG-PET and PiB-PET in three conditions with underlying Alzheimer's pathology. *Neuroimage Clin.* 4, 508–516. doi:10.1016/j.nicl.2014.03.005.

- Lane, C.A., Barnes, J., Nicholas, J.M., Sudre, C.H., Cash, D.M., Parker, T.D., Malone, I.B., Lu, K., James, S.-N., Keshavan, A., Murray-Smith, H., Wong, A., Buchanan, S.M., Keuss, S.E., Gordon, E., Coath, W., Barnes, A., Dickson, J., Modat, M., ... Schott, J.M., 2019. Associations between blood pressure across adulthood and late-life brain structure and pathology in the neuroscience substudy of the 1946 British birth cohort (Insight 46): an epidemiological study. *Lancet Neurol.* 18 (10), 942–952. doi:10.1016/S1474-4422(19)30228-5.
- Livingston, G., Huntley, J., Sommerlad, A., Ames, D., Ballard, C., Banerjee, S., Brayne, C., Burns, A., Cohen-Mansfield, J., Cooper, C., Costafreda, S.G., Dias, A., Fox, N., Gitlin, L.N., Howard, R., Kales, H.C., Kivimäki, M., Larson, E.B., Ogunniyi, A., Mukadam, N., 2020. Dementia prevention, intervention, and care: 2020 report of the Lancet commission. *Lancet* 396 (10248), 413–446. doi:10.1016/S0140-6736(20)30367-6.
- Lorenzi, M., Filippone, M., Frisoni, G.B., Alexander, D.C., Ourselin, S., 2019. Probabilistic disease progression modeling to characterize diagnostic uncertainty: application to staging and prediction in Alzheimer's disease. *Neuroimage* 190, 56–68. doi:10.1016/j.neuroimage.2017.08.059.
- Lourida, I., Hannon, E., Littlejohns, T.J., Langa, K.M., Hyppönen, E., Kuzma, E., Llewellyn, D.J., 2019. Association of lifestyle and genetic risk with incidence of dementia. *JAMA* 322 (5), 430–437. doi:10.1001/jama.2019.9879.
- Ma, Y., Wolters, F.J., Chibnik, L.B., Licher, S., Ikram, M.A., Hofman, A., Ikram, M.K., 2019. Variation in blood pressure and long-term risk of dementia: a population-based cohort study. *PLoS Med.* 16 (11), e1002933. doi:10.1371/journal.pmed.1002933.
- Nettiksimmons, J., Harvey, D., Brewer, J., Carmichael, O., DeCarli, C., Jack, C.R.J., Petersen, R., Shaw, L.M., Trojanowski, J.Q., Weiner, M.W., Beckett, L., 2010. Subtypes based on cerebrospinal fluid and magnetic resonance imaging markers in normal elderly predict cognitive decline. *Neurobiol. Aging* 31 (8), 1419–1428. doi:10.1016/j.neurobiolaging.2010.04.025.
- Ngandu, T., Lehtisalo, J., Solomon, A., Levälähti, E., Ahtiluoto, S., Antikainen, R., Bäckman, L., Hänninen, T., Jula, A., Laatikainen, T., Lindström, J., Mangialasche, F., Pajanan, T., Pajala, S., Peltonen, M., Rauramaa, R., Stigsdotter-Neely, A., Strandberg, T., Tuomilehto, J., Kivipelto, M., 2015. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. *Lancet* 385 (9984), 2255–2263. doi:10.1016/S0140-6736(15)60461-5.
- Ossenkoppele, R., Cohn-Sheehy, B.L., La Joie, R., Vogel, J.W., Möller, C., Lehmann, M., van Berckel, B.N.M., Seeley, W.W., Pijnenburg, Y.A., Gorno-Tempini, M.L., Kramer, J.H., Barkhof, F., Rosen, H.J., van der Flier, W.M., Jagust, W.J., Miller, B.L., Scheltens, P., Rabinovici, G.D., 2015. Atrophy patterns in early clinical stages across distinct phenotypes of Alzheimer's disease. *Hum. Brain Mapp.* 36 (11), 4421–4437. doi:10.1002/hbm.22927.
- Ossenkoppele, R., Schonhaut, D.R., Schöll, M., Lockhart, S.N., Ayakta, N., Baker, S.L., O'Neil, J.P., Janabi, M., Lazaris, A., Cantwell, A., Vogel, J., Santos, M., Miller, Z.A., Bettcher, B.M., Vessel, K.A., Kramer, J.H., Gorno-Tempini, M.L., Miller, B.L., Jagust, W.J., Rabinovici, G.D., 2016. Tau PET patterns mirror clinical and neuroanatomical variability in Alzheimer's disease. *Brain J. Neurol.* 139 (Pt 5), 1551–1567. doi:10.1093/brain/aww027.
- Power, M.C., Schneider, A.L.C., Wruck, L., Griswold, M., Coker, L.H., Alonso, A., Jack Jr., C.R., Knopman, D., Mosley, T.H., Gottesman, R.F., 2016. Life-course blood pressure in relation to brain volumes. *Alzheimer's Dement. J. Alzheimer's Assoc.* 12 (8), 890–899. doi:10.1016/j.jalz.2016.03.012.
- Schaare, H.L., Kharabian Masouleh, S., Beyer, F., Kumral, D., Uhlig, M., Reinelt, J.D., Reiter, A.M.F., Lampe, L., Babayan, A., Erbey, M., Roebbig, J., Schroeter, M.L., Okon-Singer, H., Müller, K., Mendes, N., Margulies, D.S., Witte, A.V., Gaebler, M., Villringer, A., 2019. Association of peripheral blood pressure with gray matter volume in 19- to 40-year-old adults. *Neurology* 92 (8). doi:10.1212/WNL.0000000000006947, e758 LP-e773.
- Scheltens, N.M.E., Tijms, B.M., Koene, T., Barkhof, F., Teunissen, C.E., Wolfsgruber, S., Wagner, M., Kornhuber, J., Peters, O., Cohn-Sheehy, B.L., Rabinovici, G.D., Miller, B.L., Kramer, J.H., Scheltens, P., van der Flier, W.M., Initiative, A.D.N., Network, G.D.C., Cohort, A.D.Center, U. of C. S. F. M. and A., 2017. Cognitive subtypes of probable Alzheimer's disease robustly identified in four cohorts. *Alzheimer's Dement. J. Alzheimer's Assoc.* 13 (11), 1226–1236. doi:10.1016/j.jalz.2017.03.002.
- Selkoe, D.J., Hardy, J., 2016. The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Mol. Med.* 8 (6), 595–608. doi:10.15252/emmm.201606210.
- Sommerlad, A., Sabia, S., Singh-Manoux, A., Lewis, G., Livingston, G., 2019. Association of social contact with dementia and cognition: 28-year follow-up of the Whitehall II cohort study. *PLoS Med.* 16 (8), e1002862. doi:10.1371/journal.pmed.1002862.
- Tijms, B.M., Gobom, J., Reus, L., Jansen, I., Hong, S., Dobricic, V., Kilpert, F., ten Kate, M., Barkhof, F., Tsolaki, M., Verhey, F.R.J., Popp, J., Martinez-Lage, P., Vandenberghe, R., Lleó, A., Molinuevo, J.L., Engelborghs, S., Bertram, L., Lovestone, S., Visser, P.J., 2020. Pathophysiological subtypes of Alzheimer's disease based on cerebrospinal fluid proteomics. *Brain* 143 (12), 3776–3792. doi:10.1093/brain/awaa325.
- Veitch, D.P., Weiner, M.W., Aisen, P.S., Beckett, L.A., DeCarli, C., Green, R.C., Harvey, D., Jack, C.R., Jagust, W., Landau, S.M., Morris, J.C., Okonkwo, O., Perrin, R.J., Petersen, R.C., Rivera-Mindt, M., Saykin, A.J., Shaw, L.M., Toga, A.W., Tosun, D., Trojanowski, J.Q., 2021. Using the Alzheimer's disease neuroimaging initiative to improve early detection, diagnosis, and treatment of Alzheimer's disease. *Alzheimer's Dement.* 2021. doi:10.1002/alz.12422.
- Venkatraghavan, V., Bron, E. E., Niessen, W. J., Klein, S., Alzheimer's Disease Neuroimaging Initiative., 2019. Disease progression timeline estimation for Alzheimer's disease using discriminative event based modeling. *NeuroImage* 186, 518–532.
- Vogel, J.W., Young, A.L., Oxtoby, N.P., Smith, R., Ossenkoppele, R., Strandberg, O.T., La Joie, R., Aksam, L.M., Grothe, M.J., Iturria-Molina, Y., Weiner, M., Aisen, P., Petersen, R., Jack, C.R., Jagust, W., Trojanowski, J.Q., Toga, A.W., Beckett, L., Green, R.C. Initiative, the A. D. N., 2021. Four distinct trajectories of tau deposition identified in Alzheimer's disease. *Nat. Med.* 27 (5), 871–881. doi:10.1038/s41591-021-01309-6.
- Whitwell, J.L., Przybelski, S.A., Weigand, S.D., Ivnik, R.J., Vemuri, P., Gunter, J.L., Senjem, M.L., Shiung, M.M., Boeve, B.F., Knopman, D.S., Parisi, J.E., Dickson, D.W., Petersen, R.C., Jack, C.R.J., Josephs, K.A., 2009. Distinct anatomical subtypes of the behavioural variant of frontotemporal dementia: a cluster analysis study. *Brain J. Neurol* 132 (Pt 11), 2932–2946. doi:10.1093/brain/awp232.
- Yates, L.A., Ziser, S., Spector, A., Orrell, M., 2016. Cognitive leisure activities and future risk of cognitive impairment and dementia: systematic review and meta-analysis. *Int. Psychogeriatr.* 28 (11), 1791–1806. doi:10.1017/S1041610216001137.
- Young, A.L., Bocchetta, M., Russell, L.L., Convery, R.S., Peakman, G., Todd, E., Cash, D.M., Greaves, C.V., van Swieten, J., Jiskoot, L., Seelaar, H., Moreno, F., Sanchez-Valle, R., Borroni, B., Laforce, R., Masellis, M., Tartaglia, M.C., Graff, C., Galimberti, D., Rohrer, J.D., 2021. Characterizing the clinical features and atrophy patterns of τ- and MAPT-related frontotemporal dementia with disease progression modeling. *Neurology* doi:10.1212/WNL.00000000000012410.
- Young, A.L., Bragman, F.J.S., Rangelov, B., Han, M.K., Galbán, C.J., Lynch, D.A., Hawkes, D.J., Alexander, D.C., Hurst, J.R., 2020. Disease progression modeling in chronic obstructive pulmonary disease. *Am. J. Respir. Crit. Care Med.* 201 (3), 294–302. doi:10.1164/rccm.201908-16000C.
- Young, A.L., Marinescu, R.V., Oxtoby, N.P., Bocchetta, M., Yong, K., Firth, N.C., Cash, D.M., Thomas, D.L., Dick, K.M., Cardoso, J., van Swieten, J., Borroni, B., Galimberti, D., Masellis, M., Tartaglia, M.C., Rowe, J.B., Graff, C., Tagliavini, F., Frisoni, G.B. (ADNI), T. A. D. N. I., 2018. Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with Subtype and Stage Inference. *Nat. Commun.* 9 (1), 4273. doi:10.1038/s41467-018-05892-0.
- Young, A.L., Oxtoby, N.P., Daga, P., Cash, D.M., Fox, N.C., Ourselin, S., Schott, J.M., Alexander, D.C., 2014. A data-driven model of biomarker changes in sporadic Alzheimer's disease. *Brain J. Neurol.* 137 (Pt 9), 2564–2577. doi:10.1093/brain/awu176.