Atrophy subtypes in Alzheimer's disease are detectable in prodromal stage and associated with different cognitive trajectories

Mara ten Kate, MD[1], Ellen Dicks MSc[1], Wiesje M. van der Flier, PhD[1,2], Charlotte E. Teunissen, PhD[3], Philip Scheltens, PhD[1], Frederik Barkhof, PhD[4,5], Pieter Jelle Visser, PhD[1,6], Betty M. Tijms, PhD[1];
Alzheimer's Disease Neuroimaging Initiative* and Amsterdam Dementia Cohort.

1. Alzheimer Center & Department of Neurology, Neuroscience Campus Amsterdam, VU University Medical Center, Amsterdam, the Netherlands.
2. Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, the Netherlands.
3. Neurochemistry Laboratory and Biobank, Department of Clinical Chemistry, VU University Medical Center, Neuroscience Amsterdam, the Netherlands.
4. Department of Radiology and Nuclear Medicine, Neuroscience Campus Amsterdam, VU University Medical Center, Amsterdam, the Netherlands.
5. Institutes of Neurology and Healthcare Engineering, University College London, London, United Kingdom.
6. Department of Psychiatry & Neuropsychology, School for Mental Health and Neuroscience, Maastricht University, Maastricht, the Netherlands.

* Data used in preparation for 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 analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Corresponding author:
Mara ten Kate
Alzheimer Center & department of Neurology, VU University Medical Center
PO Box 7057
1007 MB Amsterdam, the Netherlands
e-mail: m.tenkate1@vumc.nl
tel: + 31 (0) 20 444 33 37

Manuscript details:

Word count abstract:	303
Word count main text:	3890
Number of references	30
Number of tables:	4
Number of figures:	4
Supplementary data:	methods, 5 tables, 2 figures

Keywords Alzheimer's disease, structural MRI, clustering, nonnegative matrix factorization, biomarkers; mild cognitive impairment

ABSTRACT

Background: Alzheimer's disease (AD) is a clinically and biologically heterogeneous disease. Understanding biological factors that underlie heterogeneity is important for patient care management, and development of personalised medicine. Using an unbiased approach we determined and validated atrophy subtypes in AD-dementia patients in three large independent samples. Taking these subtypes as a starting point we further tested whether these were present in prodromal stages of AD and could explain inter-individual differences in cognitive decline.

Methods: We studied T1 weighted structural MRI from subjects with AD-dementia and prodromal AD from the Amsterdam Dementia Cohort and the Alzheimer's Disease Neuroimaging Initiative. Atrophy subtypes were identified with nonnegative matrix factorization in three independent samples of subjects with dementia (a mono-scanner mono-center sample (n=299), a multi-scanner mono-center sample (n=181) and a multi-scanner multi-center sample n=227). Subtypes were compared on clinical and biological characteristics. Next, prodromal AD subjects (n=603) were classified according to the best match of their atrophy pattern with one of the subtypes, and compared for clinical and biological characteristics, as well as trajectories of cognitive decline.

Findings: In all dementia samples we identified four atrophy subtypes: a mild atrophy subtype, a parieto-occipital atrophy subtype, a medial-temporal predominant atrophy subtype and a diffuse cortical atrophy subtype. Subtypes showed distinct impairment in cognitive domains, as well as pathological characteristics in terms of CSF tau levels and amount of vascular lesions. When classifying subjects with prodromal AD into the atrophy subtypes, we observed similar subtype profiles as in the dementia stages. Importantly, atrophy subtype was predictive for the type of cognitive decline prodromal subjects showed.

Interpretation: These findings show that clinical heterogeneity in AD can partly be explained by atrophy subtypes. These atrophy subtypes are already detectable in the prodromal stage, and can inform on expected trajectories of cognitive decline.

Funding: EU/EFPIA Innovative Medicines Initiative Joint Undertaking and ZonMW.

INTRODUCTION

Alzheimer's disease (AD) is a heterogeneous neurodegenerative disorder. Patients differ in age of dementia onset, genetic risk factors, clinical presentation, and speed of cognitive decline over time[1]. Heterogeneity in clinical presentation and trajectories of cognitive decline can already be seen in the prodromal stage of AD, i.e. subjects with mild cognitive impairment (MCI) and biomarker evidence of Alzheimer's pathology[2]. Prodromal AD subjects have a high likelihood of developing dementia, but show considerable variability in their rates of clinical progression[3] and may differ in cognitive domains that will become affected. Understanding mechanisms underlying heterogeneity in cognitive symptoms and trajectories of decline is crucial to help clinicians in prognosis and an important step towards precision medicine in AD.

Cognitive subtypes of AD-dementia have been observed in several independent patient cohorts and also in autosomal dominant AD, which suggests that clinical heterogeneity has a biological basis[4,5]. Post-mortem studies support the existence of biological heterogeneity by showing neuropathological subtypes, consisting of typical AD, limbic predominant and hippocampal sparing, which were associated with differences in antemortem clinical presentation[6]. Studying differences in brain atrophy patterns provides an *in vivo* opportunity to examine biological heterogeneity[7]. Imaging studies classifying subjects on a priori definitions of heterogeneity have identified atrophy subtypes in patients with AD-dementia that can explain some of the variability in clinical and cognitive characteristics, including a medial-temporal or typical AD atrophy variant with predominant memory dysfunction and a subtype with relative hippocampal sparing with young age of onset with more pronounced non-memory presentation[8,9]. Data-driven methods provide an unbiased approach to detect atrophy subtypes, and may be more informative for capturing heterogeneity in AD[10–14]. So far, data-driven imaging studies have examined atrophy subtypes in subjects with clinical AD-type dementia, which may not accurately reflect underlying pathology and could therefore include subtypes trivially representing subjects with non-Alzheimer pathology[10–14].

Furthermore, it remains unclear when during disease development heterogeneity in atrophy patterns arises. If heterogeneity in atrophy patterns reflects true pathophysiological subtypes of AD, subtypes may already manifest in prodromal stage of the disease when cognitive dysfunction becomes apparent and brain atrophy diverges from normal aging[15,16]. We

hypothesize that atrophy subtypes as detected in AD-dementia subjects may already be present in prodromal AD, and will be associated with decline in specific cognitive domains.

In the present study, we used a data-driven clustering approach to identify atrophy subtypes in patients with AD-type dementia and biomarker evidence of amyloid pathology, and subsequently tested for the presence of such subtypes in the prodromal stage of AD. First, we externally validated the identified subtypes in dementia patients by repeating analyses in three independent dementia patient samples: in a mono-center mono-scanner sample, a mono-center multi-scanner sample and a multi-center sample. We compared subjects of the different atrophy subtypes on clinical and cognitive characteristics. We further investigated the potential role of biological factors contributing to variability in atrophy patterns by comparing the subtypes on biomarkers (CSF markers of total tau (t-tau) and phosphorylated tau (p-tau) and white matter hyperintensities (WMH) on MRI) and Apolipoprotein E (APOE) e4 genotype[6,17,18]. Second, we classified subjects with prodromal AD into these atrophy subtypes based on the best fit of their regional grey matter volumes to the identified subtypes. We examined whether the atrophy subtypes showed similar clinical and biomarker profiles in this earlier disease stage, and whether membership of an atrophy subtype was associated with longitudinal decline in different cognitive domains.

MATERIALS AND METHODS

Sample

We included subjects from two large cohorts: the Amsterdam Dementia Cohort (ADC)[19] and the Alzheimer's Disease Neuroimaging Initiative (ADNI) (www.adni-info.org). The ADC consists of subjects attending the memory clinic of the VU University Medical Centre Amsterdam. All subjects in this cohort underwent a routine dementia screening, usually including physical and neurological examination, extensive neuropsychological testing, APOE genotyping, brain MRI scanning and lumbar puncture (unless contra-indication or patient refusal). Baseline clinical diagnosis was established during a consensus meeting from a multidisciplinary team. The ADC study protocol was approved by the VU University Medical Centre institutional review board. All subjects gave written informed consent for their clinical data to be used for research purposes. Within ADNI, subjects were recruited

over 50 sites throughout the U.S. and Canada. Standard measures include physical and neurological examination, extensive neuropsychological evaluation, APOE genotyping, brain MRI, lumbar puncture and PET. We used data of baseline or screening visits from ADNI phase-1 and phase-2. This data was obtained from the ADNI database (adni.loni.usc.edu). The ADNI study was approved by the Institutional Review Boards of all of the participating institutions. All patients gave written informed consent.

From these cohorts, subjects with a good quality 3D T1-weighted structural MRI were selected based on research criteria for AD-type dementia, i.e. having a clinical diagnosis of AD-type dementia and evidence for amyloid pathology in CSF or on amyloid PET (see supplementary methods), or prodromal AD, i.e. clinical diagnosis of MCI and evidence of amyloid pathology[2]. Patients with severe dementia (i.e., mini-mental state examination (MMSE) score < 16) were excluded. The majority of these subjects had data available on CSF biomarkers amyloid-beta 1-42 ($A\beta_{1-42}$), t-tau and p-tau, as well as WMH and APOE genotype (see Table 1 and Supplementary Methods for details). APOE e4 genotype was dichotomized by the presence of at least 1 APOE e4 allele.

Subjects with AD-dementia were divided into three samples to investigate influence of setting and scanner differences on cluster results: mono-center single-scanner (subjects from ADC scanned on single scanner; ADCs; n=299); mono-center multi-scanner (subjects from ADC scanned on either of two different scanners; ADCm; n=181); multi-center multi-scanner (ADNI; n=227). Subjects with prodromal AD from ADC (n=160) and ADNI (n=443) were grouped into one sample.

Neuropsychological assessments

In ADC and ADNI the neuropsychological assessment covered similar cognitive domains, although the cohorts differed in the tests used[19,20]. To aid comparability between cohorts, we combined test scores into four domains: memory, language, visuospatial, and attention/executive. We grouped the attention and executive domain, as ADNI has not enough tests available to split these domains. Details of the neuropsychological tests in each domain are presented in supplementary methods. The percentage of missing values in any neuropsychological test ranged from 1% to 41% in ADC and from 0 to 10% in ADNI (Supplementary Table 4). Before combining test scores into domains, missing values were

estimated through multiple imputation as implemented in SPSS (version 22; IBM) to obtain unbiased estimates of cognition. Age, sex, MMSE and education were included as predictors.

MRI acquisition and processing

For the ADC, anatomical 3D T1-weighted images were acquired as part of regular patient care on three different MRI 3T scanners using an 8-channel head coil. Subjects in the ADCs were all scanned on a single GE Signa scanner and subjects in the ADCm were scanned on either of two scanners: Toshiba Titan 3T or Philips Ingenuity PET/MRI. Details on acquisition parameters are provided in supplementary methods. In ADNI, 3D T1-weighted scans were performed on 1.5 (ADNI-1) or 3T (ADNI-2) scanners using previously described standardized protocol at each site[21].

Structural 3D T1 images were segmented into grey matter, white matter and cerebrospinal fluid using Statistical Parametric Mapping 12 (SPM12) software (Wellcome Trust Centre for Neuroimaging, University College London, UK) running in MATLAB 2011a (MathWorks Inc., Natick, MA, USA). The quality of all segmentations was visually inspected and five had to be excluded due to erroneous segmentation. In the native space grey matter segmentation of each subject, 1024 cortical and subcortical anatomical areas were defined using a brain parcellation that was generated by randomly subdividing the automated anatomical labelling atlas[22] into equally-sized regions[23]. This atlas was warped from standard space to subject space using inverted parameters that were calculated by non-linear normalization of subject images to standard MNI space. For each region, grey matter volume was defined as the sum of grey matter estimates across the voxels multiplied by voxel volume, and all volumes were normalised by total grey matter volume.

Cluster analysis

To examine atrophy subtypes, we used Nonnegative Matrix Factorization (NMF) in R (version 3.3.1, NMF version 0.20.6)[24], separately for each of the three dementia patient samples. NMF is a data-driven dual-clustering approach that identifies clusters of features (in our study atrophy patterns) and subjects at the same time (Fig. 1). Subjects are grouped into a subtype based on the best fit of their data on the identified atrophy clusters. As NMF is designed to focus on positive values, regional grey matter values were inverted to cluster regions of atrophy, rather than regions of more grey matter to facilitate interpretation. Further

details are described in the supplementary methods. We characterized each atrophy cluster based on the top 10% cluster-defining features (i.e., ROIs) in each sample. Correspondence of atrophy cluster solutions across samples was assessed with the Dice coefficient.

Classification of prodromal AD subjects

For each of the identified atrophy clusters, we made a cluster-signature by computing the average grey matter in the cluster-defining ROIs across all AD-dementia subjects classified into that subtype (Fig. 1). We then classified subjects with prodromal AD based on the lowest absolute minimal distance between their own regional grey matter values and each of the cluster-signatures. Prodromal AD subjects were next compared according to subtype classification on clinical, demographic and biomarker measures. Additionally, we examined whether prodromal AD subjects showed cognitive decline over time in subtype specific cognitive domains. To this end, baseline neuropsychological test scores of prodromal AD subjects were z-transformed and follow-up z-scores were determined relative to baseline in each sample. Per time point, the z-transformed scores were averaged across tests to obtain domain scores. In all samples, time to dementia was defined as the time between baseline visit and date of dementia diagnosis.

Statistical analysis

Subtypes were compared on demographic, clinical, neuropsychological, genetic and biomarker measures with ANOVA, Kruskal Wallis or chi-square tests when appropriate. We performed comparisons between subtypes for each sample separately, and pooled across samples. Prior to pooling, variables that were measured at different scales (i.e., level of education, CSF biomarkers, WMH) were z-transformed. For composite neuropsychological scores, results were pooled over imputed datasets using Rubins's rules as implemented in the R package MICE[25]. Subtypes were compared on atrophy patterns using voxel-based morphometry (see Supplementary Methods). In prodromal AD subjects, we used a Cox proportional hazard model to assess differences in time to progression to dementia (dependent variable) for subjects classified in each subtype (predictor variable) using the survival package (version 2.41-3) in R. Linear mixed models were used to evaluate baseline cognition and decline over time in the cognitive domains of prodromal AD subjects classified in the subtypes using the lme4 package (version 1.1-12) in R. Time from baseline (years), subtype,

and their interaction were included as fixed effects. Subject intercepts and slopes were modelled as random effects.

RESULTS

Sample characteristics

Characteristics of the three dementia patient samples are summarized in Table 1. On average, patients were 69 ± 8 years old, with the ADCs and ADCm subjects being approximately 10 years younger than ADNI subjects ($p < 0.001$). There were no significant sex differences between the samples (overall $p = 0.17$). All patients had mild-to moderate AD-dementia, with an average MMSE of 22.4 (range 16-30). Subjects in ADNI had slightly higher MMSE scores than subjects in ADCs and ADCm ($p < 0.001$ for ADCs and $p = 0.008$ for ADCm).

Subtype identification in AD-dementia subjects with NMF

In each of the samples, four clusters showed an optimal fit of the data, with high stability of the cluster solutions (Supplementary Table 1) and explaining more variance in the data than random partitions (Supplementary Table 2).

Figure 2 shows the top 10% cluster-defining ROIs (i.e. the ROIs that contribute the most to that cluster) for each of the four clusters for each subject sample. In each sample, subjects were grouped into subtypes based on the correspondence of their regional grey matter values to one of the four atrophy clusters (Supplementary Fig. 1). On average 34% (range 33%-37%) of subjects were classified as subtype 1 which showed a diffuse pattern of cluster defining features, amongst which the motor cortex. On average 28% (range 27%-30%) of subjects were classified as subtype 2, which showed parieto-occipital dominant cluster features. On average 19% (range 19%-19%) of subjects were classified in subtype 3, which showed temporal-dominant cluster features, also including the insula. The fourth subtype, including on average 18% (range 16%-21%) of subjects, showed most distinctly from the other subtypes involvement of the lateral and medial frontal lobes and lateral temporal cortex

The Dice overlap of cluster defining features was reasonable to good for most of the atrophy clusters and samples (Supplementary Table 3). Cluster 4 showed the highest Dice overlap

across samples between the 100 top cluster-defining ROIs (mean 0.57, range 0.46 - 0.72), making this the most robust cluster. The lowest Dice overlap was seen for cluster 2 (mean 0.32, range 0.29 - 0.33). For all clusters, overlap across samples increased when increasing the number of cluster-defining ROIs. Since the atrophy clusters showed similar features across samples, we pooled subjects of each subtype across samples for further analysis.

Atrophy characterization of subtypes

Compared to cognitively normal controls, all subtypes showed widespread atrophy patterns including the medial-temporal lobe, which is consistent with AD-type dementia (Fig. 3). Pairwise comparisons of atrophy patterns between subtypes revealed that the cluster-defining features largely represent areas in which subjects classified in these subtypes have most prominent atrophy, with the exception of subtype 1, which had relatively mild atrophy (Supplementary Fig. 2). The second subtype has most predominant parieto-occipital atrophy. The third subtype had most pronounced medial-temporal atrophy. The fourth subtype had diffuse cortical atrophy.

Cognitive, genetic and biological characterization of atrophy subtypes

We compared the AD-dementia atrophy subtypes on demographic, clinical, genetic and biomarker data (table 2). Subjects of the mild atrophy and parieto-occipital subtypes were younger compared to the other subtypes. Subjects of the medial-temporal subtype were the oldest and more often male. In terms of cognition, subjects of the mild atrophy subtype performed best on global cognition (highest MMSE) and in all cognitive domains. Subjects of the parieto-occipital subtype scored lowest on the MMSE and on visuospatial and executive/attention domains. Subjects of the medial-temporal subtype had low memory and language scores. Subjects of the diffuse cortical subtype had somewhat lower memory scores and average scores on the other domains. In terms of CSF biomarkers, the mild atrophy subtype had the highest t-tau and p-tau, followed by the parieto-occipital subtype. The medial-temporal subtype had the lowest t-tau and p-tau. In terms of WMH, the medial-temporal subtype had the highest amount of WMH, and the parieto-occipital the lowest. The diffuse cortical subtype had intermediate levels of t-tau, p-tau and WMH. The subtypes showed similar distributions of APOE e4 genotype. The sample specific comparisons between subtypes are presented in supplementary tables 4 and 5.

Classification of prodromal AD subjects

We further studied to what extent the atrophy subtypes are detectable in prodromal AD. Subjects with prodromal AD were classified as being one of the four AD-dementia atrophy subtypes, based on the best fit of their regional grey matter volumes to the AD cluster-signatures. The majority of prodromal AD subjects were classified into the mild atrophy subtype (table 3). When comparing subtypes in prodromal AD, several subtype differences were found that showed a similar direction as observed in dementia subjects: the mild and parieto-occipital subtypes were younger than those in the medial-temporal and diffuse cortical subtypes. Furthermore, the medial-temporal subtype showed the lowest p-tau and t-tau levels in CSF and the highest amount of WMH.

Atrophy subtypes predict which cognitive domain most prominently declines in prodromal AD

At baseline, prodromal AD subjects in the medial-temporal and diffuse cortical subtypes had significantly lower scores in the language domain, compared to prodromal AD subjects in the mild atrophy subtype (Table 4). Prodromal AD subjects in the parieto-occipital subtype had significantly worse baseline scores in the executive/attention and visuospatial domains compared to subjects in the mild atrophy subtype. There were no significant baseline group differences in global cognition (MMSE) or in the memory domain.

Prodromal AD subjects of the diffuse cortical subtype had a higher probability of progressing to dementia compared to prodromal subjects of the mild atrophy subtype (HR 1.50, $p = 0.046$) within a mean follow-up period of 2.6 ± 1.6 years (range 0.4-10.0) (Fig. 4). On average, all subtypes showed longitudinal decline in global cognition (MMSE) and in each of the cognitive domains (table 4). Prodromal AD subjects in the medial-temporal subtype showed steeper longitudinal decline on the MMSE and in the memory domain compared to subjects in the mild atrophy subtype, even though there were no significant differences at baseline. Prodromal AD subjects in the parieto-occipital atrophy subtype showed the most severe longitudinal decline on the executive/attention domain.

DISCUSSION

In this study we robustly identified four atrophy subtypes in subjects with AD-dementia using an unbiased data-driven clustering approach. These subtypes were detected independently in three samples that differed in patient populations (e.g. memory clinic vs multi-center research cohort, geographical location) and imaging acquisition protocols (single scanner vs multiple scanners). Atrophy subtypes showed distinct clinical, neuropsychological and biomarker characteristics. Moreover, atrophy subtypes were already apparent in the prodromal stage of AD, and were associated with decline in subtype specific cognitive domains, differences that were less evident at baseline. Our results support that heterogeneity in AD has a biological basis.

In line with previous studies, we found a medial-temporal dominant subtype with worst memory and language performance, a parieto-occipital subtype which was relatively young and had poor visuo-spatial and executive functioning and a diffuse cortical atrophy subtype, with intermediate cognitive scores[4,10–14,18]. A limitation of previous studies is that they were almost exclusively performed using data from ADNI, limiting the generalizability of their findings. In this study, we show that atrophy subtypes can be robustly identified in three independent patient samples, supporting the notion that these reflect true pathophysiological subtypes of AD. In line with this, we found distinct clinical and biomarker profiles between the atrophy subtypes. The medial-temporal subtype had the lowest CSF tau concentrations and highest amount of WMH, a sign commonly thought to reflect small vessel disease. This suggests that heterogeneity in atrophy patterns may in part reflect differences between AD-dementia patients in terms of the presence of co-morbidities, an issue that becomes more relevant with advancing age[26,27]. WMH have previously been associated with differences in regional atrophy patterns[28,29]. The parieto-occipital cluster had relatively high CSF tau levels, which has also been reported in a previous study[13]. Although we did not find any significant differences in the proportion of APOE e4 carriers between the subtypes, there was a trend towards less APOE e4 carriers in the parieto-occipital subtype, which is in the expected direction of previous studies that found that a young age of onset and non-memory presentation are associated with the absence of APOE e4[18]. It is likely that other genetic factors play a role in heterogeneity in atrophy patterns, as has been suggested by one study

which found differences in three AD associated single nucleotide polymorphisms between atrophy subtypes[13].

In addition to these three subtypes, we found a fourth subtype that was characterized by medial temporal and cortical atrophy compared to control subjects, but this atrophy was relatively mild compared to the other AD subtypes. One other study also identified a mild atrophy subtype, but that solution may have reflected disease stage or non-Alzheimer pathology as they included subjects with MCI without evidence of amyloid pathology in their clustering analyses[10]. Since we only included AD-type dementia subjects with biomarker evidence of underlying amyloid pathology, we minimized the possibility that subtypes reflect non-AD dementia, and we are able to say that this mild atrophy cluster is a true subtype of AD-dementia. Unexpectedly, this mild atrophy cluster showed the highest CSF p-tau and t-tau levels. CSF tau proteins and atrophy on MRI are both considered markers of neurodegeneration, which are sometimes used interchangeably[3,30]. Our results might suggest that atrophy and CSF t-tau or p-tau reflect different processes, and so caution is warranted when using CSF t-tau and p-tau in an exchangeable way as atrophy measures. Further research should aim at understanding how CSF tau concentrations are related to AD atrophy subtypes.

A particularly innovative finding of our study is that we show that the different atrophy subtypes can already be detected in subjects with prodromal AD, and that these atrophy subtypes are associated with time to progression to dementia and the cognitive domains that will become most affected with progression of the disease. Prodromal AD subjects in the medial-temporal subtype performed poorly in the language domain at baseline. These subjects declined fastest in the memory domain over time, even though there was not yet a baseline difference in memory score. Both the language and memory domain were most prominently affected in AD-dementia patients of the medial-temporal subtype. Similarly, prodromal subjects classified in the parieto-occipital subtype performed poorest on visuospatial and attention/executive domain at baseline and declined fastest in the attention/executive domain. Both these domains were also the worst affected in subjects with AD-dementia of the parieto-occipital subtype. These findings show that atrophy subtypes can be identified before the dementia stage, and are predictive for the clinical symptoms that arise at a later point in time. Recognizing this phenotypic variation in atrophy subtypes in subjects with prodromal AD could be used to improve patient care, as it provides a first step towards improving individual

prognosis. Our results may also be of consequence for clinical trials in subjects with prodromal AD. The existence of atrophy subtypes with different trajectories of cognitive decline may require subtype specific outcome measures tailored to the expected rate of decline in different cognitive domains.

A potential limitation of our study is that we did not have pathological data available. However, we did have biomarker evidence of amyloid pathology, which provides strong *in vivo* evidence of underlying Alzheimer's pathology. Moreover, it could be argued that our atrophy clusters represent differences in disease stage, rather than subtypes as we observed differences in MMSE scores. However, if that were the case then the subtype showing the lowest MMSE scores would be expected to show the most severe atrophy, as well as highest tau levels and worse performance in all cognitive domains. In contrast, our subtypes showed distinct demographics, biomarker profiles and cognitive symptoms, suggesting that these atrophy subtypes exist in parallel and do not reflect staging. In autopsy cases, which usually represent advanced disease stages, different regional patterns of neurofibrillary tangle deposits have also been identified, demonstrating the existence of heterogeneity separate from disease staging[6]. Future research in which both *in vivo* MRI and post-mortem data is available could examine how our atrophy subtypes differ in neuropathological characteristics.

In conclusion, we have robustly identified four different atrophy subtypes amongst patients with AD-dementia using a data-driven clustering approach. These subtypes had distinct demographic, cognitive and biomarker profiles. Our results provide further evidence for the existence of heterogeneity in AD patients. Furthermore, we were able to classify subjects in the prodromal stage of AD into these subtypes, which were also associated with different trajectories of cognitive decline. Heterogeneity in atrophy patterns might reflect different pathophysiological processes, which warrants further investigation. Understanding the causes of this heterogeneity is an important step towards precision medicine for AD. In the meantime, our results may inform clinical trials in prodromal AD, where atrophy subtypes could be used for subject stratification to better establish treatment efficacy.

Acknowledgements

Research of the VUmc Alzheimer Centre is part of the neurodegeneration research program of Neuroscience Amsterdam. The VUmc Alzheimer Centre is supported by Stichting Alzheimer

Nederland and Stichting VUmc fonds. The clinical database structure was developed with funding from Stichting Dioraphte. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Funding

This work has received support from the EU/EFPIA Innovative Medicines Initiative Joint Undertaking (EMIF grant: 115372) (MtK and PJV) and ZonMW Memorabel grant programme (73305056 and 733050824) (BMT and PJV). FB was supported by the NIHR biomedical research centre at UCLH. Funding sources were not involved in data collection, data analysis, interpretation, or writing of the manuscript.

Supplementary material

Supplementary material is available online at xxx.

References

1 Lam B, Masellis M, Freedman M, Stuss DT, Black SE. Clinical, imaging, and pathological heterogeneity of the Alzheimer's disease syndrome. *Alzheimers Res Ther* 2013; **5**: 1.

2 Dubois B, Feldman HH, Jacova C, *et al.* Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *The Lancet Neurology* 2014; **13**: 614–29.

3 Vos SJB, Verhey F, Frölich L, *et al.* Prevalence and prognosis of Alzheimer's disease at the mild cognitive impairment stage. *Brain* 2015; **138**: 1327–38.

4 Scheltens NME, Tijms BM, Koene T, *et al.* Cognitive subtypes of probable Alzheimer's disease robustly identified in four cohorts. *Alzheimers Dement* 2017; **13**: 1226–36.

5 Tang M, Ryman DC, McDade E, *et al.* Neurological manifestations of autosomal dominant familial Alzheimer's disease: a comparison of the published literature with the Dominantly Inherited Alzheimer Network observational study (DIAN-OBS). *Lancet Neurol* 2016; **15**: 1317–25.

6 Murray ME, Graff-Radford NR, Ross OA, Petersen RC, Duara R, Dickson DW. Neuropathologically defined subtypes of Alzheimer's disease with distinct clinical characteristics: a retrospective study. *Lancet Neurol* 2011; **10**: 785–96.

7 Whitwell JL, Dickson DW, Murray ME, *et al.* Neuroimaging correlates of pathologically defined subtypes of Alzheimer's disease: a case-control study. *Lancet Neurol* 2012; **11**: 868–77.

8 Risacher SL, Anderson WH, Charil A, *et al.* Alzheimer disease brain atrophy subtypes are associated with cognition and rate of decline. *Neurology* 2017; **89**: 2176–86.

9 Möller C, Vrenken H, Jiskoot L, *et al.* Different patterns of gray matter atrophy in early- and late-onset Alzheimer's disease. *Neurobiology of Aging* 2013; **34**: 2014–22.

10 Dong A, Toledo JB, Honnorat N, *et al.* Heterogeneity of neuroanatomical patterns in prodromal Alzheimer's disease: links to cognition, progression and biomarkers. *Brain* 2017; **140**: 735–47.

11 Hwang J, Kim CM, Jeon S, *et al.* Prediction of Alzheimer's disease pathophysiology based on cortical thickness patterns. *Alzheimers Dement* 2016; **2**: 58–67.

12 Noh Y, Jeon S, Lee JM, *et al.* Anatomical heterogeneity of Alzheimer disease Based on cortical thickness on MRIs. *Neurology* 2014; **83**: 1936–44.

13 Varol E, Sotiras A, Davatzikos C, Alzheimer's Disease Neuroimaging Initiative. HYDRA: Revealing heterogeneity of imaging and genetic patterns through a multiple max-margin discriminative analysis framework. *Neuroimage* 2017; **145**: 346–64.

14 Zhang X, Mormino EC, Sun N, *et al.* Bayesian model reveals latent atrophy factors with dissociable cognitive trajectories in Alzheimer's disease. *Proc Natl Acad Sci* 2016; **113**: E6535–44.

15 Leung KK, Bartlett JW, Barnes J, *et al.* Cerebral atrophy in mild cognitive impairment and Alzheimer disease Rates and acceleration. *Neurology* 2013; **80**: 648–54.

16 Schuff N, Tosun D, Insel PS, *et al.* Nonlinear time course of brain volume loss in cognitively normal and impaired elders. *Neurobiology of Aging* 2012; **33**: 845–55.

17 Swartz RH, Stuss DT, Gao F, Black SE. Independent Cognitive Effects of Atrophy and Diffuse Subcortical and Thalamico-Cortical Cerebrovascular Disease in Dementia. *Stroke* 2008; **39**: 822–30.

18 van der Flier WM, Pijnenburg YA, Fox NC, Scheltens P. Early-onset versus late-onset Alzheimer's disease: the case of the missing APOE ε4 allele. *The Lancet Neurology* 2011; **10**: 280–8.

19 van der Flier WM, Pijnenburg YAL, Prins N, *et al.* Optimizing Patient Care and Research: The Amsterdam Dementia Cohort. *Journal of Alzheimer's Disease* 2014; **41**: 313–27.

20 Petersen RC, Aisen PS, Beckett LA, *et al.* Alzheimer's Disease Neuroimaging Initiative (ADNI): Clinical characterization. *Neurology* 2010; **74**: 201–9.

21 Jack CR, Bernstein MA, Fox NC, *et al.* The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *J Magn Reson Imaging* 2008; **27**: 685–91.

22 Tzourio-Mazoyer N, Landeau B, Papathanassiou D, *et al.* Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 2002; **15**: 273–89.

23 Zalesky A, Fornito A, Harding IH, *et al.* Whole-brain anatomical networks: does the choice of nodes matter? *Neuroimage* 2010; **50**: 970–83.

24 Gaujoux R, Seoighe C. A flexible R package for nonnegative matrix factorization. *BMC Bioinformatics* 2010; **11**: 367.

25 Buuren S van, Groothuis-Oudshoorn CGM. mice: Multivariate Imputation by Chained Equations in R. *J Stat Softw* 2011; **45**. https://research.utwente.nl/en/publications/mice-multivariate-imputation-by-chained-equations-in-r (accessed Oct 25, 2017).

26 Tanskanen M, Mäkelä M, Notkola I-L, *et al.* Population-based analysis of pathological correlates of dementia in the oldest old. *Ann Clin Transl Neurol* 2017; **4**: 154–65.

27 Wang BW, Lu E, Mackenzie IRA, *et al.* Multiple pathologies are common in Alzheimer patients in clinical trials. *Can J Neurol Sci* 2012; **39**: 592–9.

28 Habes M, Erus G, Toledo JB, *et al.* White matter hyperintensities and imaging patterns of brain ageing in the general population. *Brain* 2016; **139**: 1164–79.

29 Lambert C, Benjamin P, Zeestraten E, Lawrence AJ, Barrick TR, Markus HS. Longitudinal patterns of leukoaraiosis and brain atrophy in symptomatic small vessel disease. *Brain* 2016; **139**: 1136–51.

30 Jack CR, Bennett DA, Blennow K, *et al.* A/T/N: An unbiased descriptive classification scheme for Alzheimer disease biomarkers. *Neurology* 2016; **87**: 539–47.

	ADCs (n=299)		ADCm (n=181)		ADNI (n=227)	
	measure	n missing	measure	n missing	measure	n missing
Demographics *						
Age (year)	67 ± 8	0 (0%)	66 ± 7	0 (0%)	74 ± 8	0 (0%)
Sex, female	149 (50%)	0 (0%)	95 (52%)	0 (0%)	99 (44%)	0 (0%)
Education (Verhage)	5.1 ± 1.2	1 (0%)	5.1 ± 1.2	0 (0%)	-	-
Education (year)	-	-	-	-	15.5 ± 3.0	0 (0%)
Global cognition						
MMSE	21.8 ± 3.3	0 0%)	22.3 ± 3.2	0 (0%)	23.3 ± 2.0	0 (0%)
APOE genotype						
APOE e4 carrier	214 (73%)	6 (2%)	123 (69%)	2 (1%)	167 (75%)	3 (1%)
CSF biomarkers †						
$A\beta_{1-42}$ (ELISA)	465 ± 97	8 (3%)	532 ± 104	23 (13%)	-	-
$A\beta_{1-42}$ (immunoassay)	-		-		130 ± 21	23 (10%)
total tau (ELISA)	684 ± 398	10 (3%)	764 ± 398	23 (13%)	-	-
total tau (immunoassay)	-	-	-		131 ± 61	23 (10%)
p-tau (ELISA)	87 ± 40	8 (3%)	88 ± 35	23 (13%)	-	-
p-tau (immunoassay)	-	-	-		59 ± 34	95 (42%)
Imaging biomarkers ‡						
WMH visual rating	1.1 ± 0.8	3 (1%)	1.1 ± 0.8	2 (1%)	-	-
WMH volume (in ml)	-	-	-	-	5.4 ± 8.0	0 (0%)

Table 1: Clinical and biomarker characteristics per dementia patient sample

Data are presented as count (%) or mean ± standard deviation. ADCs: Amsterdam Dementia Cohort single scanner; ADCm: Amsterdam Dementia Cohort multiple scanners; APOE: Apolipoprotein E; $A\beta_{1-42}$: amyloid beta 1-42, p-tau: phosphorylated tau; MMSE: mini-mental state examination.

* Education is presented according to the Verhage scale (1-7, resp low-high education) for ADCs/ADCm and in years of education for ADNI. †CSF biomarkers were measured in ADCs/ADCm using sandwich ELISAs (cut-off $A\beta_{1-42}$ < 640 ng/L, t-tau 375 ng/L, p-tau 52 ng/L) and in ADNI using immunoassays (cut-off $A\beta_{1-42}$ < 192 ng/L, t-tau 93 ng/L, p-tau 23 ng/L). ‡ WMH were measured in ADCs/ADCm using the visual Fazekas scale (range 0-3) and using automated software in ADNI (unit: ml).

	ST1 (n=247)	ST2 (n=198)	ST3 (n=136)	ST4 (n=126)	p-overall	ST1 vs ST2	ST1 vs ST3	ST1 vs ST4	ST2 vs ST3	ST2 vs ST4	ST3 vs ST4
Demographics											
Age (yr)	67 ± 8	68 ± 9	73 ± 8	70 ± 8	< 0.001	0.93	< 0.001	0.01	< 0.001	0.06	0.07
Sex, female	140 (57)	91 (46)	46 (34)	66 (52)	< 0.001	0.03	< 0.001	0.50	0.04	0.31	0.004
Education	-0.05 ± 0.98	-0.06 ± 1.04	0.15 ± 1.00	0.03 ± 0.97	0.32						
Global cognition											
MMSE	22.9 ± 2.9	21.7 ± 3.1	22.4 ± 3.1	22.3 ± 2.8	< 0.001	< 0.001	0.12	0.03	0.03	0.18	0.51
APOE genotype											
APOE e4 carrier	184 (74)	135 (68)	95 (70)	90 (71)	0.44						
CSF biomarkers											
$A\beta_{1-42}$	0.03 ± 1.01	0.13 ± 0.98	-0.05 ± 0.99	-0.20 ± 0.99	0.05						
t-tau	0.16 ± 0.97	0.07 ± 1.09	-0.36 ± 0.83	-0.03 ± 0.97	< 0.001	0.16	< 0.001	0.09	< 0.001	0.63	0.001
p-tau	0.16 ± 0.99	0.07 ± 1.08	-0.39 ± 0.85	0.001 ± 0.94	< 0.001	0.19	< 0.001	0.21	< 0.001	0.92	< 0.001
MRI biomarkers											
WMH	-0.07 ± 0.9	-0.22 ± 1.0	0.37 ± 1.1	0.08 ± 0.9	< 0.001	0.03	< 0.001	0.21	< 0.001	0.003	0.05
Neuropsychology											
Memory	0.13 ± 0.80	-0.03 ± 0.71	-0.11 ± 0.70	-0.08 ± 0.72	0.008	0.03	0.003	0.01	0.34	0.54	0.78
Language	0.16 ± 0.74	-0.12 ± 0.82	-0.12 ± 0.84	-0.01 ± 0.78	< 0.001	< 0.001	< 0.001	0.04	0.99	0.25	0.29
Visuospatial	0.20 ± 0.60	-0.33 ± 0.90	0.06 ± 0.71	0.06 ± 0.64	< 0.001	< 0.001	0.09	0.09	< 0.001	< 0.001	0.99
Executive/attention	0.21 ± 0.77	-0.20 ± 0.74	-0.05 ± 0.72	-0.04 ± 0.78	< 0.001	< 0.001	0.001	0.002	0.07	0.08	0.93

Table 2: Clinical, biomarker and neuropsychological characteristic of dementia subtypes in pooled sample

Data are presented as count (%) or mean ± standard deviation. p-values are based on chi-square or kruskall-wallis tests when appropriate. Normalized values are given for education, CSF biomarkers and MRI biomarkers. For neuropsychology, composite scores are presented. APOE: Apolipoprotein E; $A\beta_{1-42}$: amyoid-beta 1-42; ST1: subtype 1 (mild atrophy); CL2: subtype 2 (parieto-occipital atrophy); ST3: subtype 3 (medial-temporal dominant atrophy); ST4: subtype 4 (diffuse atrophy); CSF: cerebrospinal fluid; MMSE: mini-mental state examination; p-tau: phosphorylated tau; t-tau: total tau; WMH: white matter hyperintensities.

	ST1 (n=329)	ST2 (n=155)	ST3 (n=68)	ST4 (n=51)	p-overall	ST1 vs ST2	ST1 vs ST3	ST1 vs ST4	ST2 vs ST3	ST2 vs ST4	ST3 vs ST4
Demographics											
Age (yr)	70 ± 8	73 ± 7	76 ± 6	74 ± 7	< 0.001	0.009	< 0.001	0.002	0.006	0.483	0.533
Sex, female	142 (48%)	45 (38%)	18 (29%)	20 (44%)	0.003	0.05	< 0.001	0.3	0.09	0.9	0.1
Education	-0.04 ± 0.96	0.14 ± 1.01	-0.03 ± 1.15	-0.07 ± 1.02	0.625						
Global cognition											
MMSE	27.3 ± 2	27.5 ± 1.9	27.4 ± 1.8	27.3 ± 1.8	0.854						
APOE genotype											
APOE e4 carrier	218 (66%)	101 (65%)	38 (56%)	29 (57%)	0.137						
CSF biomarkers											
$A\beta_{1-42}$	0.028 ± 1.02	-0.022 ± 0.98	-0.035 ± 1.15	-0.067 ± 0.71	0.894						
t-tau	0.004 ± 0.94	0.17 ± 1.22	-0.465 ± 0.64	0.053 ± 0.83	< 0.001	0.33	0.004	0.99	< 0.001	0.89	0.03
p-tau	-0.002 ± 1.03	0.145 ± 1.02	-0.276 ± 0.91	-0.05 ± 0.77	0.068						
MRI biomarkers											
WMH	-0.06 ± 0.92	-0.038 ± 1.09	0.343 ± 1.16	0.028 ± 0.92	0.02	0.50	0.003	0.15	0.02	0.34	0.36

Table 3: Classification of prodromal AD subjects

Data are presented as count (%) or mean ± standard deviation. p-values are based on chi-square or kruskall-wallis tests when appropriate. Normalized values are given for education, CSF biomarkers and MRI biomarkers. APOE: Apolipoprotein E; $A\beta_{1-42}$: amyoid-beta 1-42; ST1: subtype 1 (mild atrophy); ST2: subtype 2 (parieto-occipital atrophy); ST3: subtype 3 (medial-temporal dominant atrophy); ST4: subtype 4 (diffuse atrophy). CSF: cerebrospinal fluid; MMSE: mini-mental state examination; p-tau: phosphorylated tau; t-tau: total tau; WMH: white matter hyperintensities.

Domain	Measure	Contrast	Estimate	Standard Error	p-value
MMSE	Baseline score	ST1 reference	27.467	0.113	0.000
	Difference at baseline	ST1 vs ST2	-0.200	0.202	0.324
		ST1 vs ST3	0.196	0.272	0.471
		ST1 vs ST4	-0.429	0.305	0.159
	Annual change	ST1 reference	-0.877	0.074	0.000
	Difference in annual change	ST1 vs ST2	-0.054	0.136	0.692
		ST1 vs ST3	**-0.420**	**0.187**	**0.025**
		ST1 vs ST4	-0.268	0.202	0.185
Memory	Baseline score	ST1 reference	0.045	0.047	0.334
	Difference at baseline	ST1 vs ST2	-0.013	0.083	0.880
		ST1 vs ST3	0.022	0.113	0.847
		ST1 vs ST4	-0.159	0.127	0.211
	Annual change	ST1 reference	-0.127	0.013	0.000
	Difference in annual change	ST1 vs ST2	-0.029	0.024	0.221
		ST1 vs ST3	**-0.076**	**0.033**	**0.022**
		ST1 vs ST4	-0.008	0.036	0.822
Language	Baseline score	ST1 reference	0.170	0.072	0.019
	Difference at baseline	ST1 vs ST2	-0.166	0.130	0.201
		ST1 vs ST3	**-0.341**	**0.173**	**0.049**
		ST1 vs ST4	**-0.421**	**0.195**	**0.031**
	Annual change	ST1 reference	-0.227	0.033	0.000
	Difference in annual change	ST1 vs ST2	0.002	0.060	0.971
		ST1 vs ST3	-0.131	0.082	0.111
		ST1 vs ST4	0.016	0.089	0.859
Visuospatial	Baseline score	ST1 reference	0.199	0.061	0.001
	Difference at baseline	**ST1 vs ST2**	**-0.232**	**0.109**	**0.034**
		ST1 vs ST3	-0.177	0.141	0.210
		ST1 vs ST4	-0.182	0.166	0.272
	Annual change	ST1 reference	-0.191	0.034	0.000
	Difference in annual change	ST1 vs ST2	0.037	0.063	0.556
		ST1 vs ST3	0.011	0.083	0.894
		ST1 vs ST4	-0.047	0.092	0.608
Executive / attention	Baseline score	ST1 reference	0.292	0.098	0.003
	Difference at baseline	**ST1 vs ST2**	**-0.429**	**0.172**	**0.013**
		ST1 vs ST3	-0.398	0.231	0.086
		ST1 vs ST4	-0.402	0.262	0.125
	Annual change	ST1 reference	-0.242	0.038	0.000
	Difference in annual change	**ST1 vs ST2**	**-0.140**	**0.069**	**0.041**
		ST1 vs ST3	-0.103	0.093	0.268
		ST1 vs ST4	-0.025	0.100	0.804

Table 4: Baseline and longitudinal cognitive scores in prodromal AD subjects classified according to atrophy subtype. Estimates from linear mixed models. The mild atrophy cluster (cluster 1) was used as reference. ST1: subtype 1 (mild atrophy); ST2: subtype 2 (parieto-occipital atrophy); ST3: subtype 3 (medial-temporal dominant atrophy); ST4: subtype 4 (diffuse atrophy).

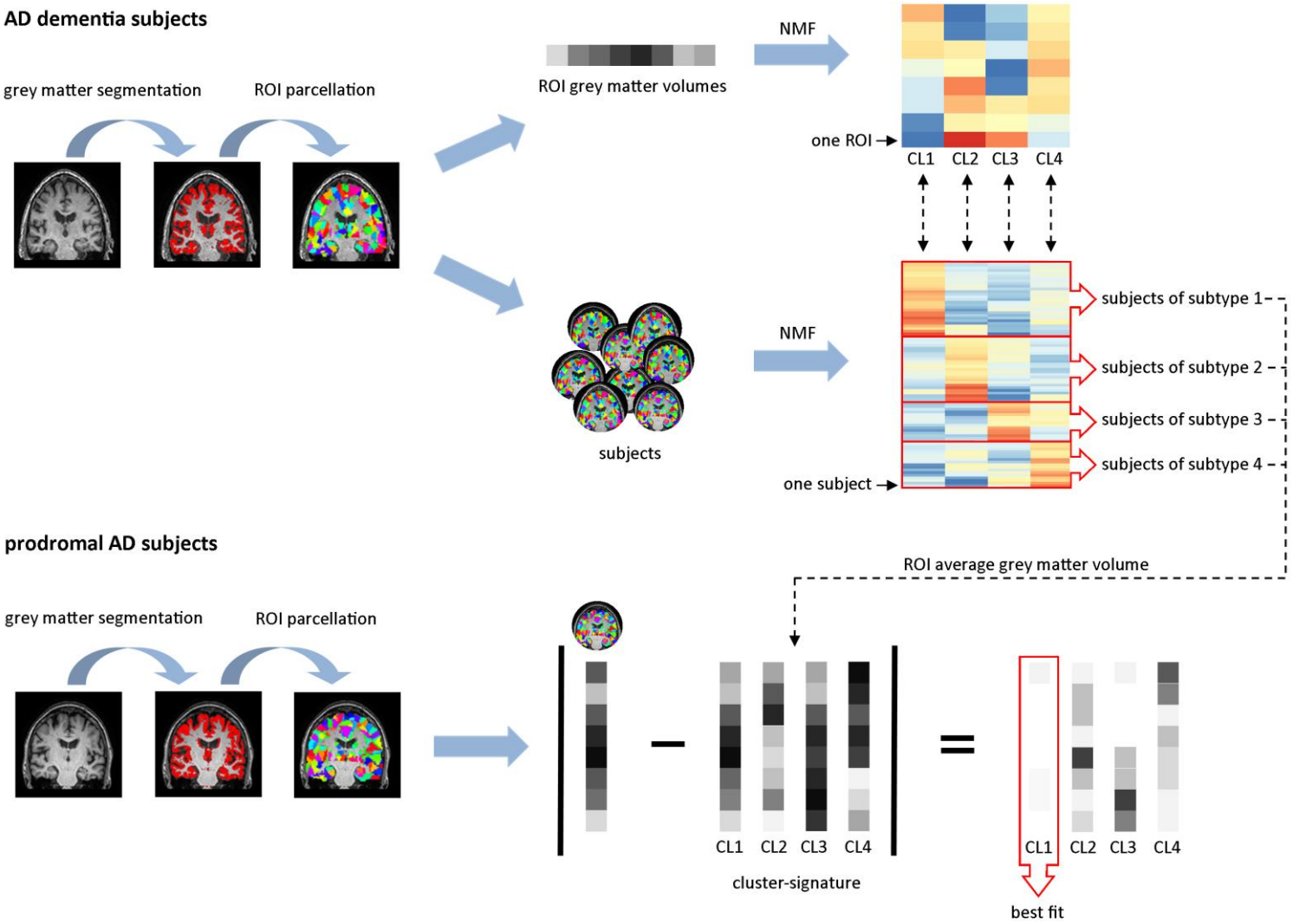

Figure 1: Non-negative matrix factorization in AD-dementia patients and classification of prodromal subjects. Grey matter segmentations were extracted from structural MRI and parcellated into 1024 equally sized ROIs, from which regional grey matter volumes were derived (for illustrative purposes only 8 ROIs are shown). In AD-dementia patients, NMF, a dual-clustering approach, was used to identify clusters of features (in this case atrophy patterns) and subjects at the same time. The ROIs are clustered into distinct atrophy patterns, illustrated in the top right part of the figure. Each row represents an ROI and each column an atrophy cluster. The warmer the colour, the more that ROI contributes to the atrophy cluster. Subjects are grouped into subtypes based on the best fit of their ROI volumes to each of the atrophy clusters as can be seen in the middle of the figure. Here, each row represents one subject and the warmer the colour, the better the fit of that subjects' ROI volumes to the ROI volumes of that atrophy cluster. For each of the atrophy clusters, we made a cluster-signature by computing the average volume in each of the top cluster-defining ROIs across all AD-dementia subjects classified as that atrophy subtype. We classified prodromal AD subjects based on the lowest mean absolute minimal distance between their own ROI volumes and that of the cluster-signatures. AD: Alzheimer's disease; CL: cluster; NMF: non-negative matrix factorization; ROI: region of interest.

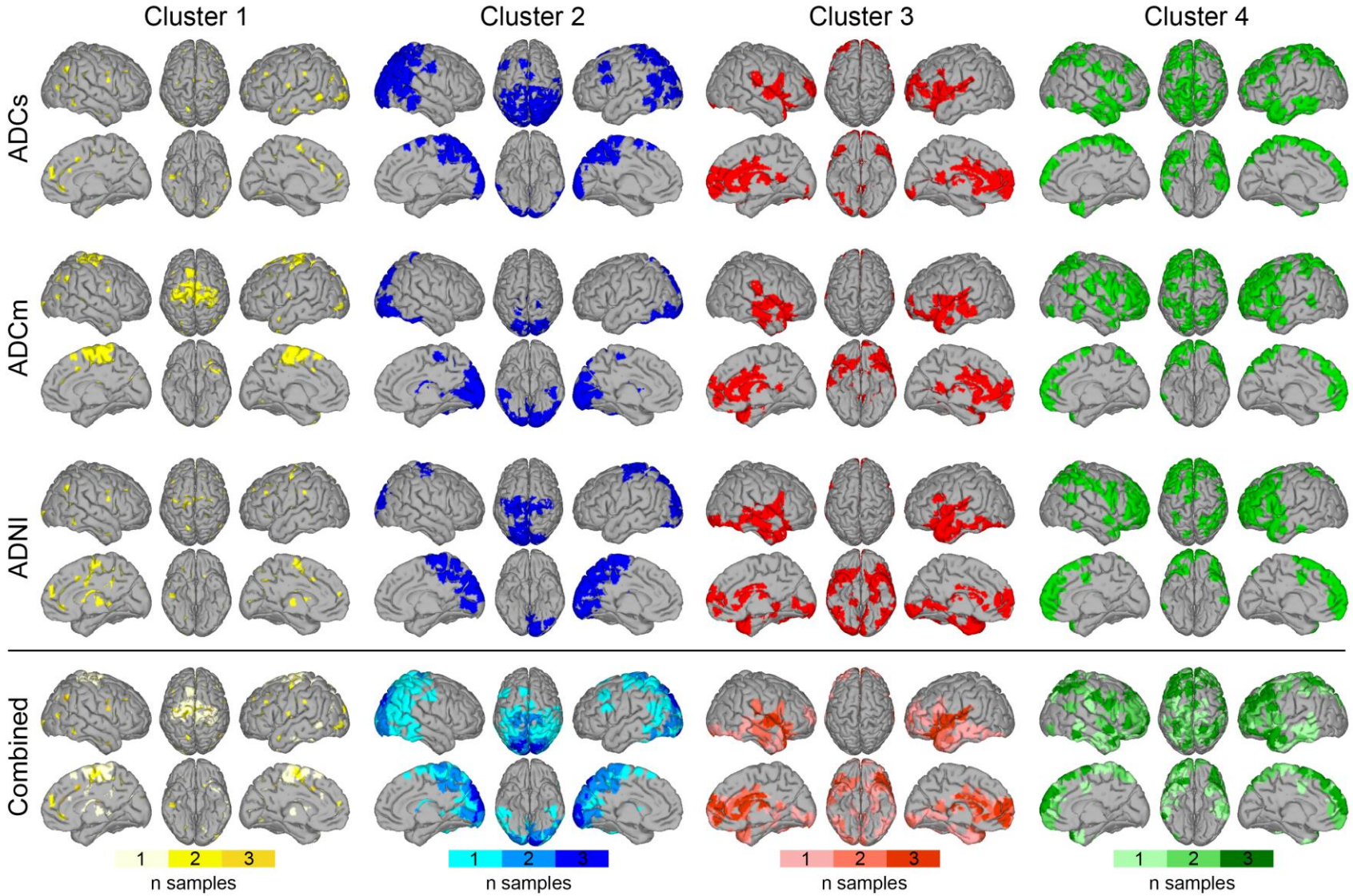

Figure 2: Cluster features across samples. In each sample we visualized the top 10% most important cluster-defining features. The lower row represents the combined important cluster features across samples: color bars indicate whether features are amongst 100 most important in 1/3, 2/3 or 3/3 samples. ADCs: Amsterdam Dementia Cohort single scanner; ADCm: Amsterdam Dementia Cohort multiple scanners; ADNI: Alzheimer's Disease Neuroimaging Initiative. Right hemisphere is displayed on the left side and vice versa.

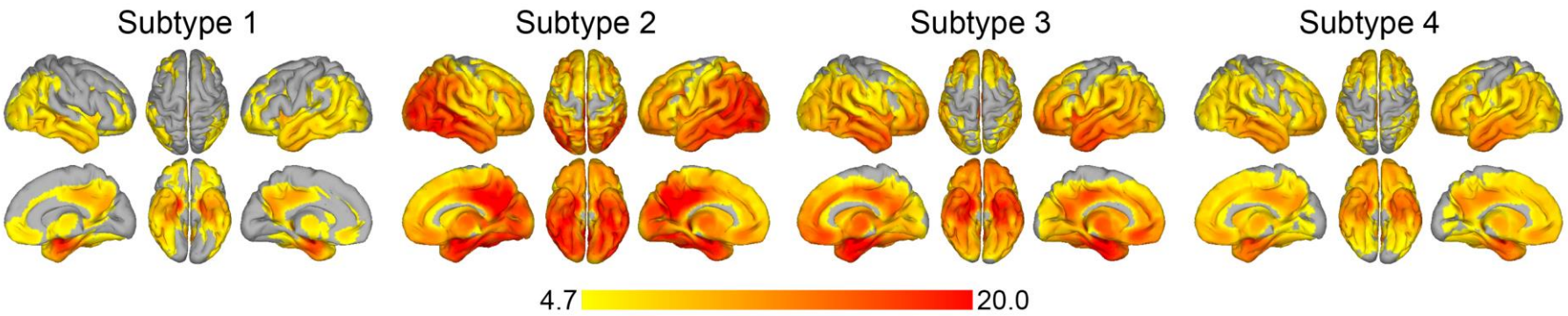

Figure 3: Voxel-based morphometry comparison between atrophy subtypes and control subjects.
Subjects in each subtype were compared to cognitively normal, amyloid negative subjects.
Colour bar represents t-statistic. Data are presented at voxel-level $p_{FWE} < 0.05$. Right hemisphere is displayed on the left side and vice versa.

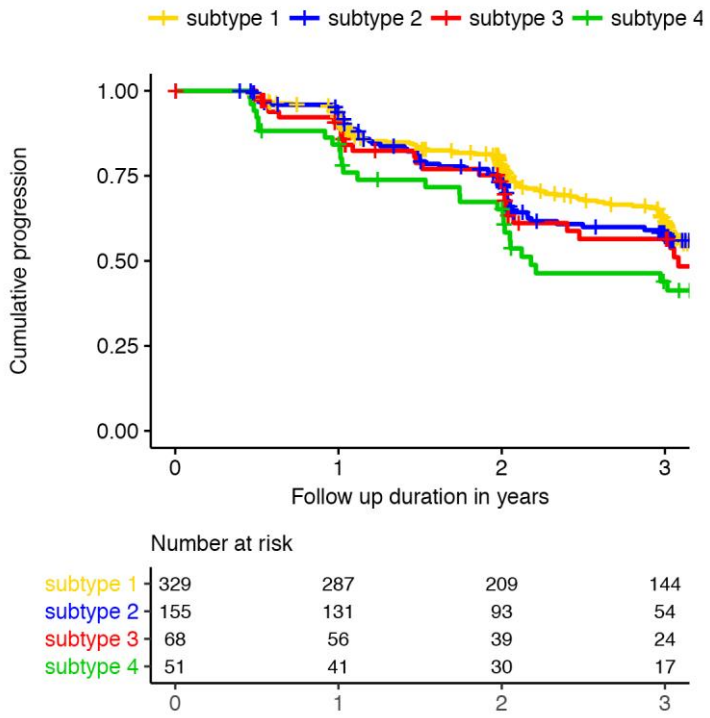

Figure 4: Progression curves for time to dementia onset for subjects with prodromal AD classified in each of the four atrophy subtypes.
Subtype 1: mild atrophy; subtype 2: parieto-occipital atrophy; subtype 3: medial-temporal atrophy; subtype 4: diffuse cortical atrophy.