Atrophy subtypes in prodromal Alzheimer's disease are associated with cognitive decline

Mara ten Kate, Ellen Dicks, Pieter Jelle Visser, Wiesje M. van der Flier, Charlotte E. Teunissen, Frederik Barkhof, Philip Scheltens, Betty M. Tijms

Supplementary Data

Supplementary Methods

MRI acquisition parameters

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. Participants in the ADCd were all scanned on a single GE Signa 3T using a three dimensional (3D) T1-weighted sagittal fast spoiled gradient echo (FSPGR) sequence (repetition time (TR) = 8, echo time (TE) = 3, inversion time (TI) = 459, flip angle (FA) = 12°, 0.98 x 0.98 x 1.00 mm voxels). Participants in the ADCv were scanned on either of two scanners: Toshiba Titan 3T scanner with 3D sagittal fast field echo (FFE) sequence (TR = 9, TE = 3, TI = 800, FA = 7°, 1.00 x 1.00 x 1.00 mm voxels) and Philips Ingenuity Time-of-Flight PET/MRI-scanner with a 3D sagittal turbo field echo (TFE) sequence (TR = 7.9 ms, TE = 4.5 ms, FA° = 8, 1.00 mm x 1.00 mm x 1.00 mm voxels). In ADNI, 3D T1-weighted scans were performed on 1.5 (ADNI-1) or 3T (ADNI-2 and ADNI-GO) scanners using previously described standardized protocol at each site, typically a sagittal MP-RAGE with a resolution around 1.2 mm isotropic[1].

Cluster analysis

Cluster analyses were performed in R (version 3.3.1). Clustering of grey matter volumes across regions of interest was performed with the R package Nonnegative Matrix Factorization (NMF, version 0.20.6)[2], separately for each of the three participant datasets. NMF is a data-driven dual-clustering approach that can be used to identify clusters of features (in our study atrophy patterns) and participants at the same time. NMF decomposes a dataset V into two non-negative matrices W and H, such that V = W*H. The original dataset V has size n x m with *n* features (in our case the 1024 regional grey matter volumes) and *m* participants. W (size n x r) is a matrix grouping the *n* ROI grey matter volumes into *r* clusters, corresponding to distinct atrophy patterns. The matrix H (size r x m) represents the clustering solution of the original dataset, in which participant specific 'loadings' of their grey matter volume patterns on each of the *r* atrophy clusters is represented. Participants 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. We performed NMF using the non-smooth NMF algorithm that enhances the sparsity of the cluster solution[3]. Within each dataset, we determined the optimal number of clusters by assessing changes in the cophenetic correlation coefficient, which represents the stability of the cluster solution[4], and changes in residual sum of squares (RSS), which represents how much of variation remains unexplained, compared to changes in RSS in random data[5]. For two to six cluster solutions, estimates of the cophenetic correlation and RSS were obtained with 30 repeats of the non-smooth NMF algorithm in the original and random data. After determining the optimal number of clusters in each dataset, NMF was run for 500 repeat runs. We characterised each atrophy cluster based on the top 10% cluster-defining features (i.e., ROIs) in each dataset. Correspondence of cluster-solutions across datasets was assessed with the Dice coefficient.

Voxel-based morphometry

Voxel-based morphometry was used to compare patterns of grey matter loss between the atrophy subtypes, and for reference with a control group of participants with normal amyloid markers and normal cognition (264 cognitively normal from ADNI; 88 subjective cognitive decline from ADC (supplementary table 4)). 3D T1 scans from control participants were segmented following procedures described above. Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL) was used to create a custom template by non-linearly aligning grey matter segmentations from participants with AD dementia and controls[6]. Native space grey matter segmentations were then spatially normalized to this template using individual flow fields. Resulting grey matter images were modulated to preserve total amount of grey matter volume and smoothed with an isotropic Gaussian filter 8 mm full-width at half-maximum (FWHM). Voxel-wise statistical comparisons between the atrophy subtypes and controls were performed using the general linear model implemented in SPM12, while correcting for TIV. Statistical maps were thresholded at a voxel-level family-wise error corrected p (p_{FWE}) < 0.05.

Amyloid pathology
In ADC, CSF markers of amyloid-beta 1-42 ($A\beta_{1-42}$) was determined using InnoTest sandwich ELISAs (Innogenetics, Fujirebio, Ghent, Belgium)[7] and a cut-off of CSF $A\beta_{1-42}$ < 640 ng/L was used to determine positivity[8]. In ADNI, CSF $A\beta_{1-42}$ measurements were performed using the Luminex multiplex platform with Innogenetics (INNO-BIO Alzbio3, Fujirebio, Ghent, Belgium) immunoassay reagents and a cut-off of CSF $A\beta_{1-42}$ < 192 ng/L was used to determine positivity.
Amyloid-PET was assessed in 160 participants in the ADC using various tracers ([^{11}C]-PiB n=41, [^{18}F]-Flutemetamol n=49, [^{18}F]-Florbetaben n=70) using routine local protocols as previously described[9,10]. Amyloid-PET scans were visually assessed by an expert nuclear physician as amyloid-positive or amyloid-negative. In ADNI, amyloid-PET was assessed in 137 participants using [^{11}C]-PiB (n=4) or [^{18}F]-Florbetapir (n=133). Cut-offs for amyloid positivity were a standardized uptake value ratio (SUVR) above 1.5 for [^{11}C]-PiB or 1.11 for [^{18}F]-AV45[11,12].

Other biomarkers
White matter hyperintensities
In the ADC, WMH were visually assessed on FLAIR images using the four point Fazekas scale (none, punctuate, early confluent, confluent)[13]. In ADNI, WMH were automatically quantified as previously described in ADNI-1[14] and ADNI-2[15].

Cerebrospinal fluid markers
In ADC, CSF t-tau and p-tau were measured using InnoTest sandwich ELISAs (Innogenetics, Fujirebio, Ghent, Belgium)[7]. In ADNI, CSF t-tau and p-tau measurements were performed using the Luminex multiplex platform with Innogenetics (INNO-BIO Alzbio3, Fujirebio, Ghent, Belgium) immunoassay reagents.

APOE genotype
In ADC, APOE genotype was determined with Light Cycler APOE mutation detection (Roche Diagnostics GmbH, Mannheim, Germany). In ADNI, APOE genotype was determined using DNC extracted by Cogenics. APOE e4 genotype was dichotomized by the presence of at least 1 APOE e4 allele.

Neuropsychological assessments
In ADC and ADNI the neuropsychological assessment covered similar cognitive domains, although the cohorts differed in the tests used[10,16]. To aid comparability between cohorts, we combined test scores into four domains: 1) Memory domain: for ADC we used the Dutch version of the Rey Auditory Verbal Learning Test (RAVLT) total immediate recall and delayed recognition and correct words of the visual association test (VAT) and for ADNI we used RAVLT total immediate recall and delayed recognition and logical memory (LM) immediate and delayed recall; 2) Language domain: for ADC we used Arizona Battery for Communication Disorders of Dementia (ABCD) naming, VAT naming and animal fluency and for ADNI we used Boston naming and animal fluency; 3) Visuospatial domain: for ADC we used number location, dot counting and fragmented letters and for ADNI we used clock drawing; 4) Attention and executive domain: for ADC we used trail making test part A (TMTA) and B (TMTB), forward Digit Span, letter fluency test (DAT), Letter Digit Substitution (LDST) test and the frontal assessment battery (FAB) and for ADNI TMTA and TMTB. We grouped the attention and executive domain, as ADNI has not enough tests available to split these domains (only TMT tests are available for all participants from ADNI-1 and ADNI-2). 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 5). 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. Imputation was repeated for 15 times. Test scores of TMT tests were inverted to have the same direction as other cognitive variables. All test-scores were z-transformed to remove measuring scale. Within each cognitive domain, z-transformed scores were averaged to obtain the composite scores.

References

1 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.
2 Gaujoux R, Seoighe C. A flexible R package for nonnegative matrix factorization. *BMC Bioinformatics* 2010; **11**: 367.
3 Pascual-Montano A, Carazo JM, Kochi K, Lehmann D, Pascual-Marqui RD. Nonsmooth nonnegative matrix factorization (nsNMF). *IEEE Trans Pattern Anal Mach Intell* 2006; **28**: 403–15.
4 Brunet J-P, Tamayo P, Golub TR, Mesirov JP. Metagenes and molecular pattern discovery using matrix factorization. *Proc Natl Acad Sci* 2004; **101**: 4164–9.

5 Frigyesi A, Höglund M. Non-negative matrix factorization for the analysis of complex gene expression data: identification of clinically relevant tumor subtypes. *Cancer Inform* 2008; **6**: 275–92.
6 Ashburner J. A fast diffeomorphic image registration algorithm. *NeuroImage* 2007; **38**: 95–113.
7 Mulder C, Verwey NA, Flier WM van der, *et al.* Amyloid-β(1–42), Total Tau, and Phosphorylated Tau as Cerebrospinal Fluid Biomarkers for the Diagnosis of Alzheimer Disease. *Clinical Chemistry* 2010; **56**: 248–53.
8 Zwan M, van Harten A, Ossenkoppele R, *et al.* Concordance between cerebrospinal fluid biomarkers and [11C]PIB PET in a memory clinic cohort. *J Alzheimers Dis* 2014; **41**: 801–7.
9 de Wilde A, van Maurik IS, Kunneman M, *et al.* Alzheimer's biomarkers in daily practice (ABIDE) project: Rationale and design. *Alzheimers Dement DADM* 2017; **6**: 143–51.
10 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.
11 Jagust WJ, Bandy D, Chen K, *et al.* The ADNI PET Core. *Alzheimers Dement* 2010; **6**: 221–9.
12 Landau S, Jagust W. Florbetapir processing methods. 2015. http://ida.loni.usc.edu.
13 Fazekas F, Chawluk JB, Alavi A, Hurtig HI, Zimmerman RA. MR Signal Abnormalities at 1.5 T in Alzheimer's Dementia and Normal Aging. *Am J Neuroradiol* 1987; **8**: 421–6.
14 Schwarz C, Fletcher E, DeCarli C, Carmichael O. Fully-Automated White Matter Hyperintensity Detection With Anatomical Prior Knowledge and Without FLAIR. *Inf Process Med Imaging* 2009; **21**: 239–51.
15 DeCarli C, Fletcher E, Ramey V, Harvey D, Jagust WJ. Anatomical mapping of white matter hyperintensities (WMH): exploring the relationships between periventricular WMH, deep WMH, and total WMH burden. *Stroke* 2005; **36**: 50–5.
16 Petersen RC, Aisen PS, Beckett LA, *et al.* Alzheimer's Disease Neuroimaging Initiative (ADNI): Clinical characterization. *Neurology* 2010; **74**: 201–9.

Supplementary Tables

Number of clusters	ADCd	ADCv	ADNI
2	0.95	0.95	0.95
3	0.82	0.81	0.85
4	0.80	0.80	0.81
5	0.76	0.73	0.76
6	0.72	0.69	0.67

Supplementary table 1: Determining the optimal number of clusters using the cophenetic correlation coefficient
The cophenetic coefficient indicates the robustness of the cluster solution for different numbers of clusters. For four clusters, there is still a high cophenetic correlation (> 0.80), after which there is a substantial decrease. For 2 to 6 cluster solutions, estimates of the cophenetic correlation for each rank were obtained with 30 repeats of the non-smooth NMF algorithm. ADCd: Amsterdam Dementia Cohort discovery dataset; ADCv: Amsterdam Dementia Cohort validation dataset; ADNI: Alzheimer's Disease Neuroimaging Initiative dataset.

Number of clusters	ADCd				ADCv				ADNI			
	Observed		Random		Observed		Random		Observed		Random	
	RSS	Change (%)	RSS	Change (%)	RSS	Change (%)	RSS	Change (%)	RSS	Change (%)	RSS	Change (%)
2	4197		18686		2926		11198		3701		15637	
3	3936	261 (6.2)	18542	144 (0.8)	2714	212 (7.2)	11065	133 (1.2)	3440	261 (7.0)	15486	151 (1.0)
4	3789	147 (3.7)	18399	143 (0.8)	2573	141 (5.2)	10931	134 (1.2)	3254	186 (5.4)	15336	150 (1.0)
5	3645	143 (3.8)	18265	134 (0.7)	2473	70 (3.9)	10815	116 (1.1)	3117	137 (4.2)	15199	137 (0.9)
6	3501	144 (3.9)	18125	140 (0.8)	2375	98 (4.0)	10696	119 (1.1)	3000	117 (3.8)	15062	137 (0.9)

Supplementary table 2: Determining the optimal number of clusters using RSS for observed and random data.
The residual sum of squares (RSS) represents how much of the variation the model did not explain. For 2 to 6 cluster solutions, estimates of the RSS were obtained for the observed data and random data with 30 repeats of the non-smooth NMF algorithm. Change is the difference in RSS between current and previous cluster number. For 3 and 4 clusters, the change in explained variation is greater in the observed data than in random data in all three datasets. ADCd: Amsterdam Dementia Cohort discovery dataset; ADCv: Amsterdam Dementia Cohort validation dataset; ADNI: Alzheimer's Disease Neuroimaging Initiative dataset.

Top n contributing features	Cluster	ADCd-ADCv	ADCd-ADNI	ADCv-ADNI
100	Cluster 1	0.63	0.40	0.52
	Cluster 2	0.29	0.35	0.33
	Cluster 3	0.47	0.58	0.46
	Cluster 4	0.54	0.46	0.72
200	Cluster 1	0.74	0.53	0.63
	Cluster 2	0.49	0.57	0.55
	Cluster 3	0.58	0.76	0.60
	Cluster 4	0.70	0.66	0.81

Supplementary table 3: Dice overlap between cluster features

The dice overlap between the most important features defining each cluster were computed between the datasets. Top part: dice overlap between 100 most important cluster-defining features. Bottom part: dice overlap between 200 most important cluster-defining features. For all clusters, overlap across datasets increased when increasing the number of cluster-defining ROIs. The selected features are visualised in Figure 1a (100 features) and Supplementary Figure 2 (200 features). ADCd: Amsterdam Dementia Cohort discovery dataset; ADCv: Amsterdam Dementia Cohort validation dataset; ADNI: Alzheimer's Disease Neuroimaging Initiative dataset.

	ADC (n=88)		ADNI (n=250)	
	measure	n missing	measure	n missing
Demographics				
Age (yr)	61 ± 7	0 (0%)	73 ± 6	0 (0%)
Sex, female	25 (28%)	0 (0%)	124 (50%)	0 (0%)
Education (yr)	12.0 ± 2.9	1 (1%)	16.5 ± 2.6	0 (0%)
Global cognition				
MMSE	28.1 ± 1.7	0 (0%)	29 ± 1.2	0 (0%)
APOE genotype				
APOE e4 carrier	35 (40%)	1 (1%)	45 (18%)	0 (0%)
CSF biomarkers †				
Aβ1-42 (Innotest)	987 ± 225	0 (0%)	-	-
Aβ1-42 (Luminex)	-	-	229 ± 37	0 (0%)
Total tau (Innotest)	280 ± 163	0 (0%)	-	-
Total tau (Luminex)	-	-	60 ± 24	0 (0%)
total tau abnormal [#]	12 (14%)	-	28 (11%)	-
p-tau (Innotest)	47 ± 20	1 (1%)	-	-
p-tau (Luminex)	-	-	29 ± 13	31 (12%)
p-tau abnormal [#]	32 (37%)	-	129 (59%)	-
Imaging biomarkers ‡				
WMH visual rating	0.6 ± 0.7	3 (3%)	-	-
WMH volume (in ml)	-	-	0.6 ± 1.7	185 (74%)

Supplementary table 4: Clinical and biomarker characteristics of cognitively normal individuals with normal biomarkers. Data are presented as count (%) or mean ± standard deviation. ADC: Amsterdam Dementia Cohort; ADNI: Alzheimer's Disease Neuroimaging Initiative dataset; APOE: Apolipoprotein E; $A\beta_{1\text{-}42}$: amyloid beta 1-42, p-tau: phosphorylated tau; MMSE: mini-mental state examination. † CSF biomarkers were measured in ADCd/ADCv using sandwich ELISAs (cut-off $A\beta_{1\text{-}42} < 640$ ng/L, t-tau ≥ 375 ng/L, p-tau ≥ 52 ng/L) and in ADNI using immunoassays (cut-off $A\beta_{1\text{-}42} < 192$ ng/L, t-tau > 93 ng/L, p-tau > 23 ng/L). [#] percentage abnormal are based on available data (excluding missings). ‡ WMH were measured in ADC using the visual Fazekas scale (range 0-3) and using automated software in ADNI (unit: ml).

	ADCd				ADCv				ADNI			
	ST1 (n=57)	**ST2 (n=81)**	**ST3 (n=111)**	**ST4 (n=50)**	**ST1 (n=35)**	**ST2 (n=49)**	**ST3 (n=59)**	**ST4 (n=38)**	**ST1 (n=44)**	**ST2 (n=68)**	**ST3 (n=77)**	**ST4 (n=38)**
Demographics												
Age (yr)	70 ± 8	63 ± 7	66 ± 7	70 ± 8	70 ± 6	65 ± 7	65 ± 6	69 ± 7	78 ± 5	76 ± 8	71 ± 8	73 ± 8
Sex, female	19 (33%)	35 (43%)	63 (57%)	32 (64%)	15 (43%)	27 (55%)	34 (58%)	19 (50%)	12 (27%)	29 (43%)	43 (56%)	15 (39%)
Education (yr)	11.5 ± 3.1	11.0 ± 2.5	11.1 ± 2.7	11.2 ± 2.7	11.5 ± 2.9	11.0 ± 2.8	11.5 ± 3.0	10.8 ± 2.7	16.1 ± 2.6	15.4 ± 3.3	15.0 ± 3.0	16.0 ± 2.6
Global cognition												
MMSE	22.1 ± 3.8	20.7 ± 3.2	22.3 ± 3.1	21.9 ±3.0	22.7 ± 2.9	21.4 ± 3.2	23.0 ± 3.4	22.1 ± 3.0	22.5 ± 1.8	23.1 ± 2.1	23.8 ± 1.8	23.0 ± 2.3
APOE genotype												
APOE e4 carrier	38 (67%)	57 (70%)	85 (77%)	34 (68%)	24 (69%)	30 (61%)	43 (73%)	26 (68%)	33 (75%)	48 (71%)	56 (73%)	30 (79%)
CSF biomarkers †												
$A\beta_{1\text{-}42}$	457 ± 88	463 ± 97	477 ± 101	453 ± 102	520 ± 89	564 ± 106	536 ± 111	489 ± 88	-	-	-	-
$A\beta_{1\text{-}42}$	-	-	-	-	-	-	-	-	131 ± 25	133 ± 19	127 ± 20	128 ± 22
total tau	556 ± 367	697 ± 481	750 ± 375	667 ± 299	660 ± 357	800 ± 389	836 ± 433	688 ± 372	-	-	-	-
total tau	-	-	-	-	-	-	-	-	101 ± 38	138 ± 64	138 ± 58	140 ± 78
p-tau	69 ± 34	93 ± 48	93 ± 35	89 ± 37	78 ± 35	90 ± 33	95 ± 36	81 ± 35	-	-	-	-
p-tau	-	-	-	-	-	-	-	-	47 ± 22	57 ± 34	65 ± 39	64 ± 30
MRI biomarkers									-	-	-	-
WMH visual ‡	1.4 ± 0.8	0.8 ± 0.7	1.0 ± 0.8	1.2 ± 0.9	1.4 ± 0.8	0.9 ± 0.9	0.9 ± 0.7	1.2 ± 0.8				
WMH volume ‡	-	-	-	-	-	-	-	-	8.4 ± 10.7	4.0 ± 8.4	5.3 ± 6.5	4.8 ± 5.1

Supplementary table 5: Comparison of clinical and biomarker between subtypes for each dataset

Data are presented as count (%) or mean ± standard deviation. p-values are based on chi-square, anova or kruskall-wallis tests when appropriate. ADCd: Amsterdam Dementia Cohort discovery dataset; ADCv: Amsterdam Dementia Cohort validation dataset; ADNI: Alzheimer's Disease Neuroimaging Initiative dataset; APOE: Apolipoprotein E; $A\beta_{1\text{-}42}$: amyoid-beta 1-42; ST1: subtype 1 (medial-temporal dominant atrophy); ST2: subtype 2 (parieto-occipital atrophy); ST3: subtype 3 (mild atrophy); ST4: subtype 4 (diffuse atrophy); MMSE: mini-mental state examination; p-tau: phosphorylated tau; t-tau: total tau; WMH: white matter hyperintensities.

†CSF biomarkers were measured in ADC using sandwich ELISAs (cut-off $A\beta_{1\text{-}42}$ < 640 ng/L, t-tau 375 ng/L, p-tau 52 ng/L) and in ADNI using immunoassays (cut-off $A\beta_{1\text{-}42}$ < 192 ng/L, t-tau 93 ng/L, p-tau 23 ng/L). ‡ WMH were measured in ADC using the visual Fazekas scale (range 0-3) and using automated software in ADNI (unit: ml).

	ADCd				ADCv				ADNI				N missing (%)		
	ST1	ST2	ST3	ST4	ST1	ST2	ST3	ST4	ST1	ST2	ST3	ST4	ADCd	ADCv	ADNI
Memory	-0.14 ± 0.72	0.10 ± 0.78	0.08 ± 0.75	-0.17 ± 0.72	-0.02 ± 0.78	-0.01 ± 0.71	-0.03 ± 0.78	0.08 ± 0.70	-0.14 ± 0.60	-0.20 ± 0.57	0.32 ± 0.85	-0.13 ± 0.70	-	-	-
RAVLT IR	20.6 ± 6.9	23.9 ± 7.6	23.1 ± 7.5	21.3 ± 7	23.2 ± 6.8	21.4 ± 7.1	21.3 ± 7.7	22.8 ± 7	21.6 ± 8.5	21.5 ± 7.0	24.9 ± 7.1	20.3 ± 7.1	18 (6)	5 (3)	2 (1)
RAVLT DR	1.6 ± 2.3	2.3 ± 2.3	1.9 ± 2.3	1.2 ± 1.8	1.5 ± 2.0	1.9 ± 1.8	1.6 ± 2.0	2.1 ± 2.1	0.5 ± 1.2	0.4 ± 1.0	1.3 ± 2.2	0.7 ± 1.4	20 (7)	6 (3)	0 (0)
VAT	5.9 ± 3.7	5.8 ± 3.7	6.8 ± 3.7	5.9 ± 4.1	6.2 ± 3.8	6.7 ± 3.8	7.1 ± 4.2	6.6 ± 3.7	-	-	-	-	24 (8)	1 (1)	-
LM IR	-	-	-	-	-	-	-	-	3.6 ± 2.7	3.4 ± 2.3	4.7 ± 2.8	3.6 ± 2.6	-	-	0 (0)
LM DR	-	-	-	-	-	-	-	-	1.1 ± 1.4	0.9 ± 1.4	1.9 ± 2.0	1.3 ± 1.7	-	-	0 (0)
Language	-0.08 ± 0.62	-0.08 ± 0.82	0.12 ± 0.76	-0.04 ± 0.69	-0.09 ± 1.01	-0.13 ± 0.69	0.21 ± 0.65	-0.06 ± 0.76	-0.18 ± 0.92	-0.15 ± 0.88	0.20 ± 0.78	0.07 ± 0.90	-	-	-
Animal fluency	12.4 ± 5.2	12.1 ± 5.0	13.6 ± 5.1	12.6 ± 5.1	14.2 ± 4.8	13.4 ± 4.1	14.2 ± 5.3	12.2 ± 4.6	11.6 ± 4.8	11.2 ± 4.8	13.6 ± 4.5	12.1 ± 5.3	27 (9)	7 (4)	0 (0)
ABCD naming	15.4 ± 2.8	16.1 ± 3.2	16.6 ± 3.0	15.4 ± 2.8	17.4 ± 2.6	16.1 ± 2.6	16.1 ± 3.6	16.6 ± 3.2	-	-	-	-	99 (33)	7 (4)	-
VAT naming	11.3 ± 0.9	11.2 ± 1.5	11.5 ± 1.1	11.5 ± 1.0	11.7 ± 0.8	11.2 ± 1.3	11.1 ± 1.3	11.5 ± 1.1	-	-	-	-	24 (8)	2 (1)	-
Boston naming	-	-	-	-	-	-	-	-	21.1 ± 6.4	21.8 ± 6.5	23.2 ± 5.2	23.3 ± 6.3	-	-	2 (1)
Visuospatial	0.01 ± 0.64	-0.29 ± 0.88	0.15 ± 0.52	0.14 ± 0.56	0.24 ± 0.47	-0.44 ± 0.91	0.20 ± 0.53	0.03 ± 0.58	-0.01 ± 0.90	-0.31 ± 0.92	0.27 ± 0.73	0.00 ± 0.77	-	-	-
Number loc	8.0 ± 2.0	7.5 ± 2.1	8.4 ± 1.6	7.9 ± 2.0	8.4 ± 1.7	7.5 ± 2.2	8.2 ± 2	7.8 ± 2.3	-	-	-	-	29 (10)	11 (6)	-
Dot count	9.5 ± 0.9	8.5 ± 1.9	9.3 ± 1.0	9.8 ± 0.5	9.4 ± 1.3	8.3 ± 2.5	9.7 ± 0.6	9.2 ± 1.1	-	-	-	-	86 (29)	7 (4)	-
Fragm letters	15.4 ± 5.7	14.7 ± 5.5	17.0 ± 3.8	16.6 ± 3.2	18 ± 2.6	14.6 ± 5.7	18.1 ± 2.3	17.8 ± 2.2	-	-	-	-	85 (28)	11 (6)	-
Clock	-	-	-	-	-	-	-	-	3.2 ± 1.3	2.9 ± 1.4	3.8 ± 1.3	3.2 ± 1.5	-	-	0 (0)
Attent/Exec	-0.08 ± 0.56	-0.20 ± 0.58	0.14 ± 0.65	0.11 ± 0.61	0.07 ± 0.60	-0.11 ± 0.53	0.18 ± 0.71	-0.21 ± 0.63	-0.10 ± 0.96	-0.26 ± 0.98	0.33 ± 0.93	-0.08 ± 1.04	-	-	-
TMT-A	-80 ± 44	-106 ± 73	-66 ± 37	-76 ± 56	-60 ± 34	-91 ± 62	-66 ± 37	-82 ± 54	-64 ± 37	-82 ± 42	-52 ± 27	-66 ± 36	12 (4)	4 (2)	4 (2)
TMT-B	-174 ± 64	-174 ± 60	-158 ± 64	-151 ± 58	-162 ± 68	-169 ± 64	-162 ± 63	-186 ± 63	-206 ± 90	-217 ± 84	-184 ± 84	-199 ± 87	124 (41)	55 (30)	23 (10)
Digit span	11.2 ± 2.9	10.6 ± 2.9	11.1 ± 2.9	11.5 ± 2.1	12.0 ± 2.9	11.2 ± 2.5	11.2 ± 2.6	10.7 ± 2.7	-	-	-	-	7 (2)	5 (3)	-
Letter fluency	23.7 ± 11.3	28.3 ± 11.7	29.3 ± 11.4	27.5 ± 9.1	31.2 ± 12.3	31.6 ± 11.1	28.2 ± 10.5	27.2 ± 11.4	-	-	-	-	50 (17)	9 (5)	-
LDST	26.8 ± 9.6	20.8 ± 11.3	30.9 ± 11.4	30.0 ± 10.7	31.1 ± 9.9	25.4 ± 13.0	31.8 ± 11.1	25.8 ± 9.6	-	-	-	-	98 (33)	30 (17)	-
FAB	12.5 ± 3.5	12.4 ± 3.3	13.5 ± 3.1	13.3 ± 3.7	14.2 ± 3.3	13.0 ± 3.2	14.0 ± 2.4	12.7 ± 3.8	-	-	-	-	70 (23)	10 (6)	-

Supplementary table 6: Neuropsychological cluster characteristics comparisons between subtypes for each dataset
ADCd: Amsterdam Dementia Cohort discovery dataset; ADCv: Amsterdam Dementia Cohort validation dataset; ADNI: Alzheimer's Disease Neuroimaging Initiative dataset.
ST1: subtype 1 (medial-temporal dominant atrophy); ST2: subtype 2 (parieto-occipital atrophy); ST3: subtype 3 (mild atrophy); ST4: subtype 4 (diffuse atrophy).

	ADC (n=160)		ADNI (n=443)	
	measure	N missing	measure	N missing
Demographics				
Age (yr)	68 ± 7	0 (0%)	73 ± 7	0 (0%)
Sex, female	76 (48%)	0 (0%)	186 (42%)	0 (0%)
Education (yr)	12 ± 3.0	2 (1%)	16 ± 2.8	0 (0%)
Global cognition				
MMSE	26.4 ± 2.5	0 (0%)	27.5 ± 1.8	0 (0%)
APOE genotype				
APOE e4 carrier	105 (66%)	13 (8%)	281 (63%)	0 (0%)
Follow-up information				
Progression to dementia (%)	99 (62%)	0 (0%)	178 (40%)	30 (7%)
Time to progression (yr)	-	-	2.6 ± 1.6	39 (9%)
Number of follow-up	3.7 ± 1.5	10 (6%)	5.5 ± 2.1	0 (0%)
Follow-up time (yr)	2.6 ± 1.5	0 (0%)	3.4 ± 1.9	0 (0%)
CSF biomarkers †				
Aβ1-42 (Innotest)	469 ± 106	0 (0%)	-	-
Aβ1-42 (Luminex)	-	-	140 ± 29	30 (7%)
Total tau (Innotest)	588 ± 365	0 (0%)	-	-
Total tau (Luminex)	-	-	108 ± 58	30 (7%)
total tau abnormal [#]	114 (71%)	-	209 (51%)	-
p-tau (Innotest)	81 ± 35	0 (0%)	-	-
p-tau (Luminex)	-	-	46 ± 24	116 (26%)
p-tau abnormal [#]	131 (82%)	-	282 (86%)	-
Imaging biomarkers ‡				
WMH visual rating	1.0 ± 0.8	0 (0%)	-	-
WMH volume (in ml)	-	-	5.1 ± 8.2	44 (10%)

Supplementary table 7: Clinical and biomarker characteristics per prodromal AD dataset. Data are presented as count (%) or mean ± standard deviation. ADC: Amsterdam Dementia Cohort; ADNI: Alzheimer's Disease Neuroimaging Initiative dataset; APOE: Apolipoprotein E; $A\beta_{1\text{-}42}$: amyloid beta 1-42, p-tau: phosphorylated tau; MMSE: mini-mental state examination.. † CSF biomarkers were measured in ADCd/ADCv using sandwich ELISAs (cut-off $A\beta_{1\text{-}42} < 640$ ng/L, t-tau ≥ 375 ng/L, p-tau ≥ 52 ng/L) and in ADNI using immunoassays (cut-off $A\beta_{1\text{-}42} < 192$ ng/L, t-tau > 93 ng/L, p-tau > 23 ng/L). [#] percentage abnormal are based on available data (excluding missings). ‡ WMH were measured in ADC using the visual Fazekas scale (range 0-3) and using automated software in ADNI (unit: ml).

	ST1: medial-temporal	ST2: parieto-occipital	ST3: mild	ST4: diffuse	ST1 vs ST2	ST1 vs ST3	ST1 vs ST4	ST2 vs ST3	ST2 vs ST4	ST3 vs ST4
MMSE										
Baseline score	27.66 ± 0.24	27.27 ± 0.17	27.47 ± 0.11	27.04 ± 0.28	0.19	0.47	0.096	0.32	0.49	0.16
Annual change	-1.30 ± 0.26	-0.93 ± 0.21	-0.88 ± 0.07	-1.15 ± 0.19	0.075	0.025	0.55	0.69	0.30	0.19
Memory										
Baseline score	0.07 ± 0.10	0.03 ± 0.07	0.05 ± 0.05	-0.11 ± 0.12	0.78	0.85	0.25	0.88	0.28	0.21
Annual change	-0.20 ± 0.03	-0.16 ± 0.02	-0.13 ± 0.01	-0.13 ± 0.03	0.19	0.022	0.14	0.22	0.60	0.82
Language										
Baseline score	-0.17 ± 0.15	0.004 ± 0.11	0.17 ± 0.07	-0.25 ± 0.18	0.36	0.049	0.74	0.20	0.22	0.031
Annual change	-0.36 ± 0.08	-0.23 ± 0.05	-0.23 ± 0.03	-0.21 ± 0.08	0.14	0.11	0.19	0.97	0.89	0.86
Visuospatial										
Baseline score	0.02 ± 0.13	-0.03 ± 0.09	0.20 ± 0.06	0.02 ± 0.15	0.73	0.21	0.98	0.034	0.77	0.27
Annual change	-0.18 ± 0.08	-0.15 ± 0.05	-0.19 ± 0.03	-0.24 ± 0.08	0.78	0.89	0.61	0.56	0.39	0.61
Executive/attention										
Baseline score	-0.11 ± 0.21	-0.14 ± 0.14	0.29 ± 0.10	-0.11 ± 0.24	0.90	0.086	0.99	0.013	0.92	0.12
Annual change	-0.35 ± 0.08	-0.38 ± 0.06	-0.24 ± 0.04	-0.27 ± 0.09	0.72	0.27	0.54	0.041	0.29	0.80

Supplementary table 8: Baseline and longitudinal cognitive score estimated beta's (se) in prodromal AD participants classified according to atrophy subtype.
Estimates from linear mixed models. ST1: subtype 1 (medial-temporal dominant atrophy); ST2: subtype 2 (parieto-occipital atrophy); ST3: subtype 3 (mild atrophy); ST4: subtype 4 (diffuse atrophy).

	ST1: medial-temporal	ST2: parieto-occipital	ST3: mild	ST4: diffuse	ST1 vs ST2	ST1 vs ST3	ST1 vs ST4	ST2 vs ST3	ST2 vs ST4	ST3 vs ST4
MMSE										
Baseline score	27.73 ± 0.25	27.27 ± 0.18	27.47 ± 0.14	27.27 ± 0.29	0.11	0.33	0.13	0.32	0.78	0.33
Annual change	-1.30 ± 0.17	-0.93 ± 0.11	-0.88 ± 0.07	-1.14 ± 0.19	0.067	0.023	0.53	0.71	0.33	0.19
Memory										
Baseline score	0.09 ± 0.10	0.002 ± 0.07	-0.01 ± 0.05	-0.08 ± 0.12	0.46	0.35	0.24	0.85	0.50	0.55
Annual change	-0.20 ± 0.03	-0.16 ± 0.02	-0.13 ± 0.01	-0.14 ± 0.03	0.19	0.021	0.13	0.23	0.61	0.82
Language										
Baseline score	-0.07 ± 0.16	0.04 ± 0.11	0.18 ± 0.09	-0.12 ± 0.18	0.56	0.15	0.84	0.26	0.44	0.17
Annual change	-0.36 ± 0.07	-0.23 ± 0.05	-0.23 ± 0.03	-0.21 ± 0.08	0.14	0.11	0.19	0.97	0.89	0.86
Visuospatial										
Baseline score	0.09 ± 0.13	-0.05 ± 0.10	0.14 ± 0.07	0.06 ± 0.16	0.38	0.73	0.89	0.087	0.52	0.65
Annual change	-0.18 ± 0.08	-0.16 ± 0.05	-0.19 ± 0.03	-0.24 ± 0.09	0.78	0.88	0.61	0.55	0.39	0.61
Executive/attention										
Baseline score	-0.05 ± 0.21	-0.17 ± 0.15	0.23 ± 0.12	-0.008 ± 0.24	0.63	0.23	0.90	0.018	0.56	0.35
Annual change	-0.34 ± 0.08	-0.38 ± 0.06	-0.24 ± 0.04	-0.27 ± 0.09	0.72	0.27	0.54	0.042	0.30	0.81

Supplementary table 9: Baseline and longitudinal cognitive scores in prodromal AD participants classified according to atrophy subtype with covariate correction. Estimates from linear mixed models. Age, gender and education were included as covariates. ST1: subtype 1 (medial-temporal dominant atrophy); ST2: subtype 2 (parieto-occipital atrophy); ST3: subtype 3 (mild atrophy); ST4: subtype 4 (diffuse atrophy).

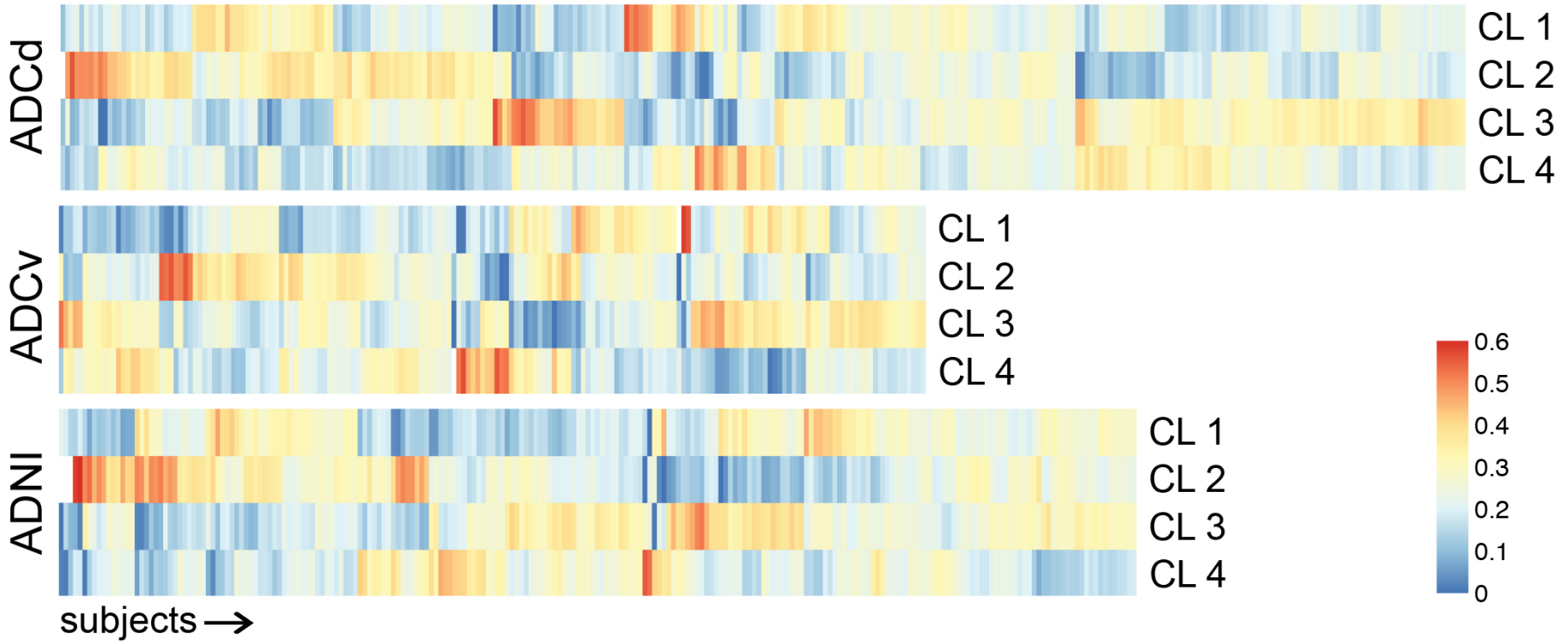
	ST1: medial-temporal	ST2: parieto-occipital	ST3: mild	ST4: diffuse	ST1 vs ST2	ST1 vs ST3	ST1 vs ST4	ST2 vs ST3	ST2 vs ST4	ST3 vs ST4
Memory										
Year 3	-0.43 ± 0.20	-0.23 ± 0.13	-0.17 ± 0.08	-0.31 ± 0.20	0.4	0.2	0.7	0.7	0.7	0.5
Language										
Year 3	-1.03 ± 0.30	-0.57 ± 0.20	-0.16 ± 0.13	-0.72 ± 0.31	0.2	**0.008**	0.5	0.08	0.7	0.1
Visuospatial										
Year 3	-0.34 ± 0.33	-0.36 ± 0.22	-0.27 ± 0.14	-0.44 ± 0.35	0.9	0.8	0.8	0.7	0.9	0.6
Executive/attention										
Year 3	-0.85 ± 0.39	-0.94 ± 0.25	0.06 ± 0.16	-0.88 ± 0.39	0.8	**0.03**	0.9	**0.001**	0.9	**0.03**

Supplementary table 10: 3 year follow-up cognitive scores in prodromal AD participants classified according to atrophy subtype. Data are presented as estimate ± se normalized values (z-scores). p-values are based Anova tests. Number of subjects available at year 3 = 352. ST1: subtype 1 (medial-temporal dominant atrophy); ST2: subtype 2 (parieto-occipital atrophy); ST3: subtype 3 (mild atrophy); ST4: subtype 4 (diffuse atrophy).

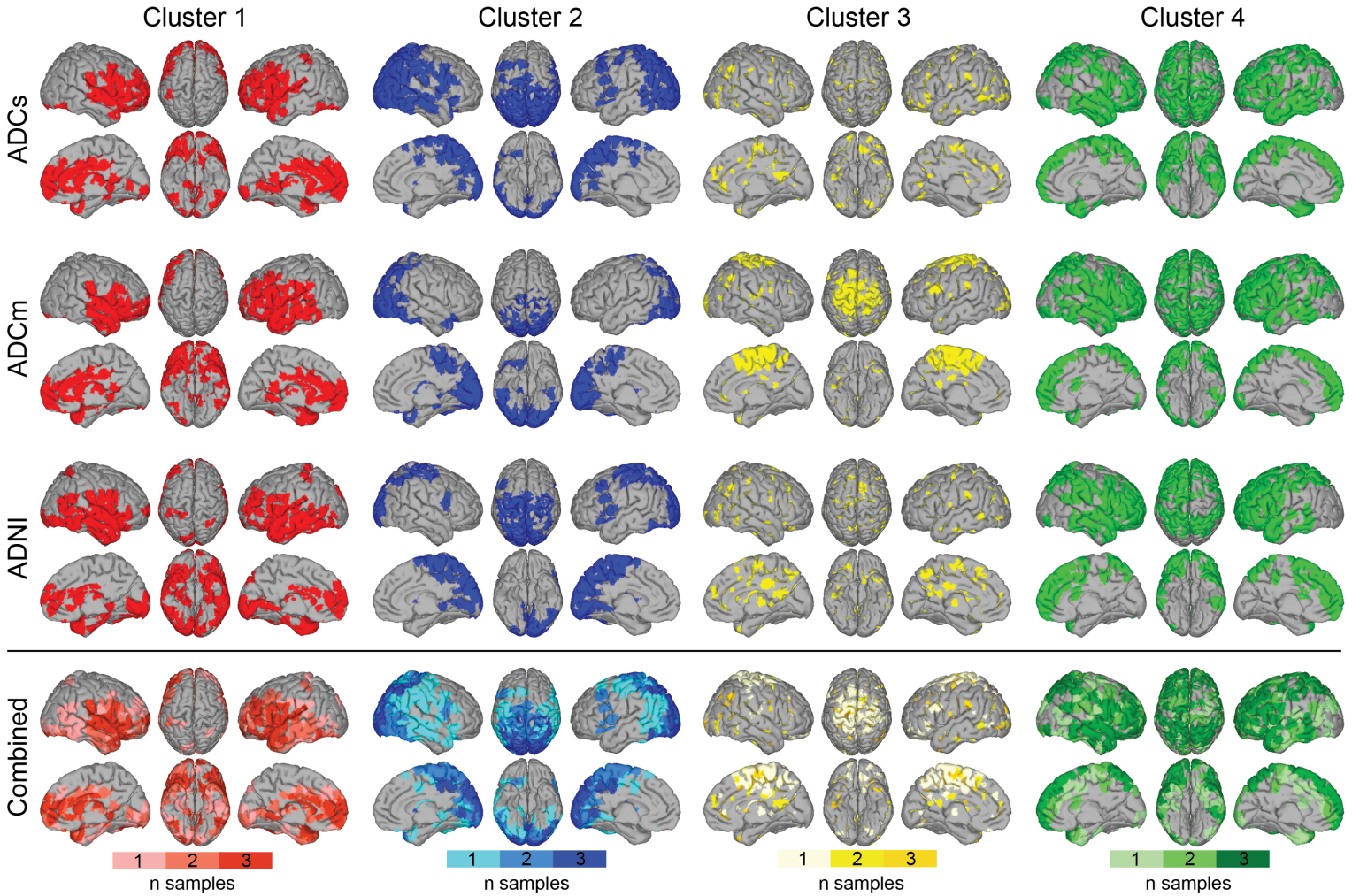
	< 2 years follow-up (n=245)	≥ 2 years follow-up (n=358)	p-value
Demographics			
Age (yr)	72.7 ± 7.6	71.4 ± 7.4	0.030
Sex, female	104 (42%)	158 (44%)	0.74
Education	14.9 ± 3.4	14.8 ± 3.5	0.79
Global cognition			
MMSE	26.8 ± 2.1	27.5 ± 2.1	0.0003
APOE genotype			
APOE e4 carrier	164 (67%)	222 (62%)	0.025
CSF biomarkers			
$A\beta_{1-42}$	-0.061 ± 0.84	0.036 ± 1.08	0.26
t-tau	0.190 ± 1.18	-0.114 ± 0.85	0.0004
p-tau	0.192 ± 1.11	-0.099 ± 0.92	0.002
MRI biomarkers			
WMH	0.040 ± 1.02	-0.029 ± 0.98	0.42

Supplementary table 10: Difference between prodromal AD participants with and without more than 2 years follow-up.

Supplementary Figures



Supplementary Figure 1: Cluster membership probabilities. Participants were assigned to one of the four atrophy clusters based on the fit of their regional grey matter volumes to each of the clusters. In each dataset, each column represents one participant. The warmer the colour, the better the fit of participants' regional grey matter volumes to the regional grey matter volumes of that cluster. ADCd: Amsterdam Dementia Cohort discovery dataset; ADCv: Amsterdam Dementia Cohort validation dataset; ADNI: Alzheimer's Disease Neuroimaging Initiative dataset. CL1: cluster 1 (medial-temporal dominant atrophy); CL2: cluster 2 (parieto-occipital atrophy); CL3: cluster 3 (mild atrophy); CL4: cluster 4 (diffuse atrophy).



Supplementary Figure 2: Top 200 cluster features across datasets. In each dataset we visualized the top 200 most important cluster-defining features. The lower row represents the combined important cluster features across datasets: colour bars indicate whether the top 200 cluster defining features were observed in 1/3, 2/3 or 3/3 datasets. Right hemisphere is displayed on the left side and vice versa. ADCd: Amsterdam Dementia Cohort discovery dataset; ADCv: Amsterdam Dementia Cohort validation dataset; ADNI: Alzheimer's Disease Neuroimaging Initiative dataset.