

# SCIENTIFIC REPORTS



OPEN

## Distinct subtypes of Alzheimer's disease based on patterns of brain atrophy: longitudinal trajectories and clinical applications

Received: 06 January 2017

Accepted: 08 March 2017

Published: 18 April 2017

Daniel Ferreira<sup>1</sup>, Chloë Verhagen<sup>1,2</sup>, Juan Andrés Hernández-Cabrera<sup>3</sup>, Lena Cavallin<sup>4,5</sup>, Chun-Jie Guo<sup>1,6</sup>, Urban Ekman<sup>1</sup>, J-Sebastian Muehlboeck<sup>1</sup>, Andrew Simmons<sup>1,7,8,9</sup>, José Barroso<sup>3</sup>, Lars-Olof Wahlund<sup>1</sup> & Eric Westman<sup>1,7</sup>

Atrophy patterns on MRI can reliably predict three neuropathological subtypes of Alzheimer's disease (AD): typical, limbic-predominant, or hippocampal-sparing. A method to enable their investigation in the clinical routine is still lacking. We aimed to (1) validate the combined use of visual rating scales for identification of AD subtypes; (2) characterise these subtypes at baseline and over two years; and (3) investigate how atrophy patterns and non-memory cognitive domains contribute to memory impairment. AD patients were classified as either typical AD (n = 100), limbic-predominant (n = 33), or hippocampal-sparing (n = 35) by using the Scheltens' scale for medial temporal lobe atrophy (MTA), the Koedam's scale for posterior atrophy (PA), and the Pasquier's global cortical atrophy scale for frontal atrophy (GCA-F). A fourth group with no atrophy was also identified (n = 30). 230 healthy controls were also included. There was great overlap among subtypes in demographic, clinical, and cognitive variables. Memory performance was more dependent on non-memory cognitive functions in hippocampal-sparing and the no atrophy group. Hippocampal-sparing and the no atrophy group showed less aggressive disease progression. Visual rating scales can be used to identify distinct AD subtypes. Recognizing AD heterogeneity is important and visual rating scales may facilitate investigation of AD heterogeneity in clinical routine.

Alzheimer's disease (AD) is a heterogeneous disease<sup>1–5</sup>. Current diagnostic criteria recognize this heterogeneity in the form of different cognitive presentations<sup>6–8</sup>. However, there is also neuropathological and structural heterogeneity<sup>4,9</sup>. Whitwell *et al.*<sup>10</sup> grouped AD patients into amnesic and non-amnesic types. Amnesic patients evidenced atrophy in the medial temporal lobe, while non-amnesic patients showed atrophy in lateral regions of the parietal, temporal, and frontal lobes with relative sparing of the medial temporal lobes<sup>10</sup>. Subtyping based on the spread of neurofibrillary tangles (NFT) revealed fairly corresponding groups<sup>4</sup>. The amnesic form was highly represented on both the typical AD subtype (balanced NFT counts in the hippocampus and the associative cortex, i.e. lateral parietal, temporal, and frontal regions) and the limbic-predominant subtype (NFT counts predominantly in the hippocampus). The non-amnesic syndromes were more frequent in the atypical hippocampal-sparing AD subtype (NFT counts predominantly in the associative cortex). In a subsequent study, patterns of atrophy in MRI reliably tracked the distribution of NFT pathology at autopsy<sup>9</sup>. Hence, evidence suggests a connection between patterns of NFT spread, brain atrophy, and the cognitive presentation.

<sup>1</sup>Department of Neurobiology, Care Sciences and Society, Centre for Alzheimer Research, Division of Clinical Geriatrics, Karolinska Institutet, Stockholm, Sweden. <sup>2</sup>Department of Psychology, Faculty of Social and Behavioural Sciences, Utrecht University, Utrecht, The Netherlands. <sup>3</sup>Faculty of Psychology, University of La Laguna, Tenerife, Spain. <sup>4</sup>Department of Clinical Science, Intervention and Technology, Division of Medical Imaging and Technology, Karolinska Institutet, Stockholm, Sweden. <sup>5</sup>Department of Radiology, Karolinska University Hospital in Huddinge, Huddinge, Sweden. <sup>6</sup>Department of Radiology, The First Hospital of Jilin University, Jilin, China. <sup>7</sup>Department of Neuroimaging, Centre for Neuroimaging Sciences, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK. <sup>8</sup>NIHR Biomedical Research Centre for Mental Health, London, UK. <sup>9</sup>NIHR Biomedical Research Unit for Dementia, London, UK. Correspondence and requests for materials should be addressed to D.F. (email: daniel.ferreira.padilla@ki.se)

	Healthy controls (n = 230)	Alzheimer's disease subtypes				p-value
		Typical AD (n = 100, 50.5%)	Limbic-predominant (n = 33, 16.7%)	Hippocampal-sparing (n = 35, 17.7%)	No atrophy group (n = 30, 15.2%)	
Age	75.9 (5.0)	77.7 (6.4)	72.3 (7.1) <sup>a,b</sup>	78.2 (8.2) <sup>c</sup>	69.2 (6.9) <sup>a,b,d</sup>	<0.001
Gender, % female	48.7	37.0	57.6	48.6	73.3 <sup>b</sup>	0.016
Years of education	16.0 (2.9)	15.0 (3.4) <sup>a</sup>	14.6 (2.2)	14.9 (3.1)	13.6 (3.7) <sup>a</sup>	<0.001
MMSE	29.1 (1.0)	22.9 (2.1) <sup>a</sup>	23.2 (1.9) <sup>a</sup>	23.7 (2.2) <sup>a</sup>	24.2 (1.5) <sup>ab</sup>	<0.001
CDR total	0 (0)	0.8 (0.2) <sup>a</sup>	0.7 (0.2) <sup>ab</sup>	0.7 (0.3) <sup>a</sup>	0.7 (0.2) <sup>ab</sup>	<0.001
Age at onset <sup>e</sup>	—	73.6 (7.2)	68.6 (6.9) <sup>b</sup>	76.3 (7.6) <sup>c</sup>	66.8 (7.1) <sup>bd</sup>	<0.001
Disease duration <sup>e</sup>	—	3.8 (2.6)	3.3 (2.2)	3.2 (2.7)	2.6 (1.4)	0.189
ApoE status, % $\epsilon 4$ allele	26.5	69.0 <sup>a</sup>	69.7 <sup>a</sup>	51.4 <sup>a</sup>	73.3 <sup>a</sup>	<0.001
CSF A $\beta_{1-42}$ , % abnormal <sup>f</sup>	40.0	96.0 <sup>a</sup>	90.0 <sup>a</sup>	85.0 <sup>a</sup>	91.7 <sup>a</sup>	<0.001
CSF T-tau, % abnormal <sup>g</sup>	21.2	56.0 <sup>a</sup>	75.0 <sup>a</sup>	75.0 <sup>a</sup>	66.7 <sup>a</sup>	<0.001

**Table 1. Characteristics of the AD subtypes and healthy controls.** The table shows mean (SD) except for Gender, APOE status and the CSF biomarkers, where percentage is reported. <sup>a</sup>Significantly different from Healthy controls; <sup>b</sup>Significantly different from Typical AD; <sup>c</sup>Significantly different from Limbic-predominant; <sup>d</sup>Significantly different from Hippocampal-sparing. <sup>e</sup>N = 154; <sup>f</sup>N = 217; <sup>g</sup>N = 215. AD = Alzheimer's disease; CDR = clinical dementia rating; MMSE = mini-mental state examination; APOE = apolipoprotein E; CSF = cerebrospinal fluid; A $\beta_{1-42}$  = amyloid- $\beta$ -peptide 1–42; T-tau = total level of tau protein.

Recently, Byun *et al.*<sup>11</sup> investigated these three subtypes as well as a fourth AD group with no atrophy by studying brain atrophy patterns on MRI data from the Alzheimer's Disease Neuroimaging Initiative (ADNI-1<sup>12,13</sup>). Further, longitudinal progression over two years was studied. Limbic-predominant AD and the group with no atrophy showed slower progression than typical AD and hippocampal-sparing AD<sup>11</sup>. Data-driven approaches using MRI data have largely confirmed these pathologically defined subtypes<sup>1,2,14,15</sup>. Other authors have also applied data-driven approaches to cognitive data but the resulting subtypes differ noticeably from study to study<sup>3,5,16,17</sup>. However, data-driven approaches rely on group analysis and sophisticated methods that make them difficult to translate into clinical practice at present. Still, MRI is in a privileged position for studying AD heterogeneity because impairment in a given cognitive function may emerge from heterogeneous underlying neuropathology and atrophy patterns<sup>8–10,18</sup>.

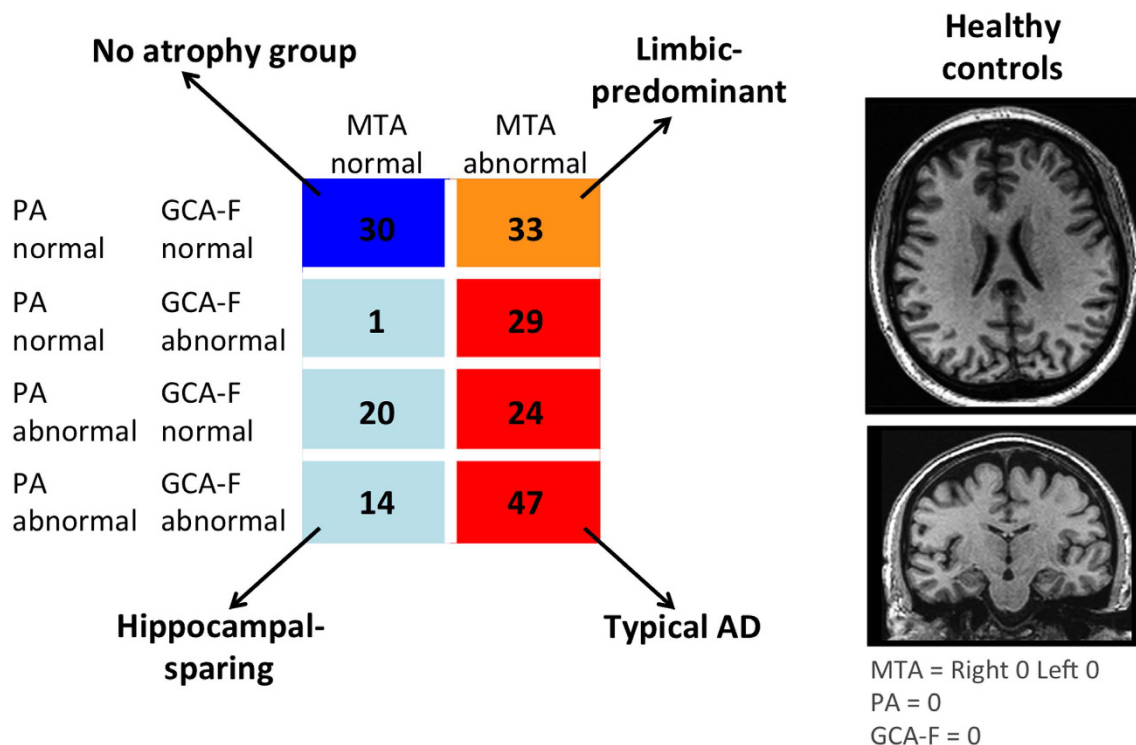
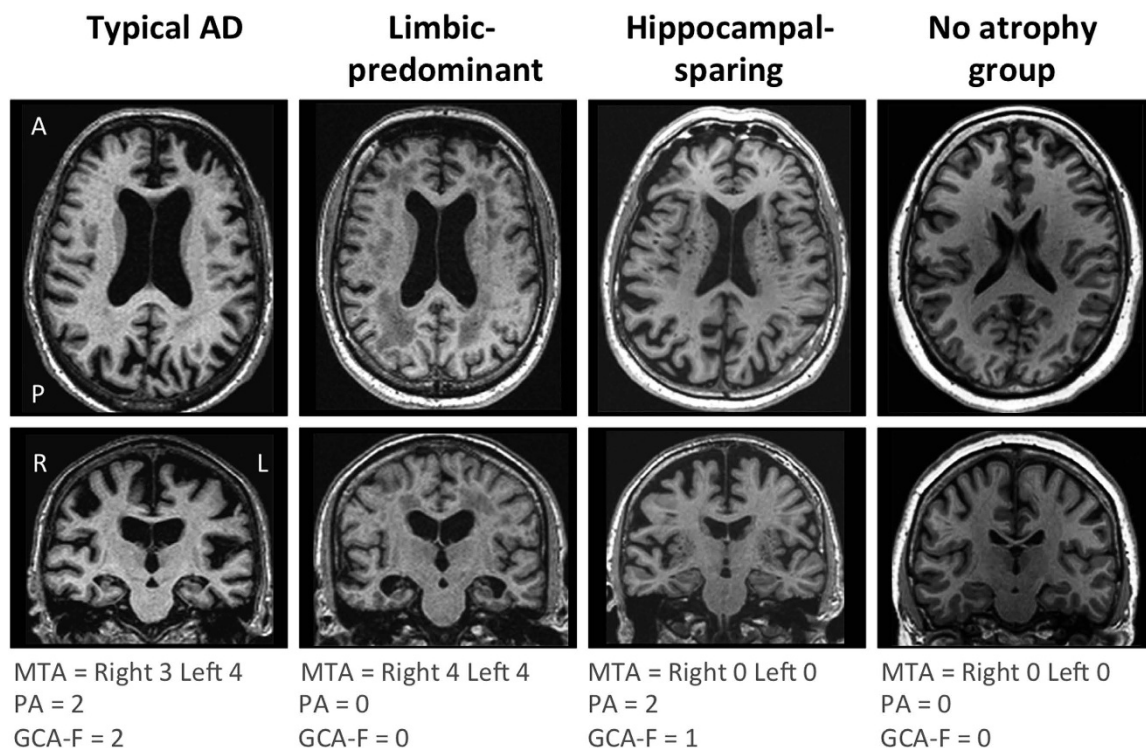
We investigated whether visual rating scales of brain atrophy in MRI might be useful to capture the above-mentioned AD subtypes. Visual rating scales are quick and easy to use, and are the primary method for assessing brain structural changes in clinical settings<sup>18–21</sup>. However, visual rating scales are often used individually. Applying them in combination increases their diagnostic capacity and enables the study of patterns of brain atrophy<sup>18,19</sup>. We propose a way to easily identify patterns of atrophy using three visual rating scales covering the medial temporal, frontal and posterior cortices. We aimed to (1) validate the combined use of visual rating scales for identification of AD subtypes; (2) characterize the resulting subtypes at baseline and longitudinally over two years; and (3) since all the AD patients in our sample were amnesic, we also investigated how atrophy patterns and non-memory cognitive domains contribute to memory impairment, a relevant question not yet investigated in different AD subtypes. Thus, the three aims were addressed to facilitate investigation of the different AD subtypes in the clinical routine using already at-place and widely used clinical diagnostic tools.

## Results

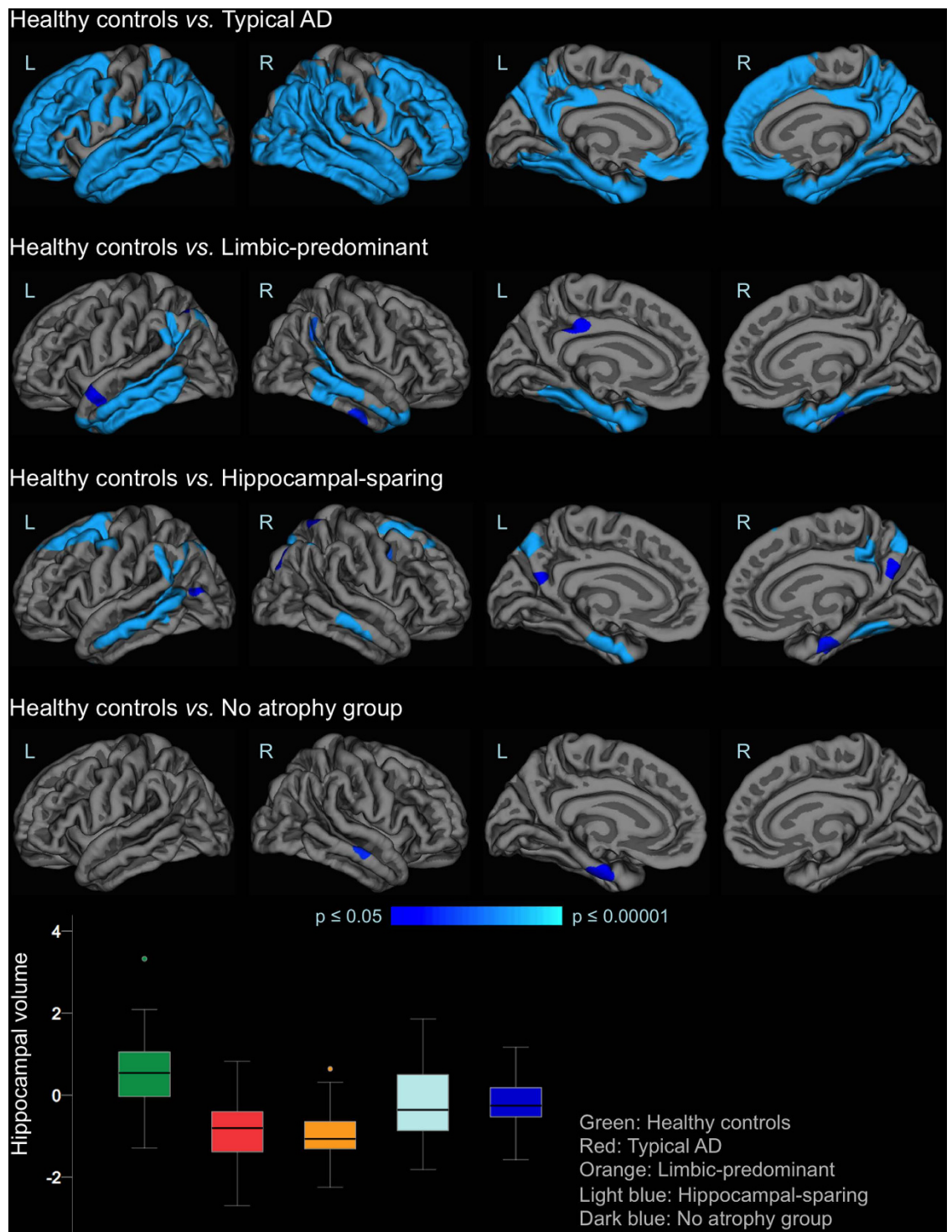
**Clinical and cognitive characterization of the AD subtypes.** Table 1 shows the main demographic and clinical characteristics of the study groups. Visual examples for each group are shown in Fig. 1. The largest group was typical AD (n = 100), as expected, present in 50.5% of the AD patients. The atypical subtypes were less prevalent and showed similar frequency: hippocampal-sparing (n = 35, 17.7%), limbic-predominant (n = 33, 16.7%), and no atrophy group (n = 30, 15.2%). Maps of cortical thickness as well as hippocampal volumes are displayed in Fig. 2.

Three random forest models were conducted to characterize the study groups according to (1) demographic-clinical variables, (2) memory variables, and (3) non-memory cognitive variables (see Table 2 for a list of variables included in each analysis as well as summary of results). Results showed great overlap (Fig. 3). Healthy controls and typical AD patients were correctly classified in the three models. Both resulted as the least and the most severe groups, respectively. However, any combination of the different sets of variables did not allow classifying limbic-predominant, hippocampal-sparing, and the no atrophy group better than chance. To note, the four AD subtypes were comparable on disease duration, CSF biomarkers and APOE  $\epsilon 4$  distribution.

A follow-up mixed ANOVA was performed to further investigate the interaction between five memory components (within-subjects factor: total learning, interference, immediate free recall, delayed free recall, and recognition) and AD subtype (between-subjects factor). This interaction was statistically significant ( $F_{(5, 324)} = 3.419$ ;  $p = 0.004$ ). Impairment in the learning component was greater in typical AD than in both hippocampal-sparing ( $t_{(86)} = -3.427$ ;  $p = 0.006$ ) and the no atrophy group ( $t_{(117)} = -3.181$ ;  $p = 0.010$ ), but this effect was not observed for the other components (Fig. 3). These results hold after controlling for age, Mini-Mental state examination (MMSE), Clinical Dementia Rating (CDR), and disease duration (data not shown).



**Figure 1. Subtypes of AD based on patterns of brain atrophy from visual rating scales.** Regional atrophy was measured with the MTA, PA and GCA-F visual rating scales based only on T1-weighted images. In the three visual rating scales, a score of zero denotes no atrophy, whereas scores from one to three (PA and GCA-F) or four (MTA) indicate an increasing degree of atrophy. The typical AD subtype was defined as abnormal MTA together with abnormal PA and/or abnormal GCA-F. The limbic-predominant subtype was defined as abnormal MTA alone with normal PA and GCA-F. The hippocampal-sparing subtype included abnormal PA and/or abnormal GCA-F, but normal MTA. The no atrophy group was defined as normal scores in MTA, PA, and GCA-F. The figure shows examples for each subtype as well as descriptive analysis on the different study groups. AD = Alzheimer’s disease; MTA = medial temporal atrophy scale; PA = posterior atrophy scale; GCA-F = global cortical atrophy scale – frontal subscale; A = anterior part of the brain; P = posterior part of the brain; R = right; L = left.



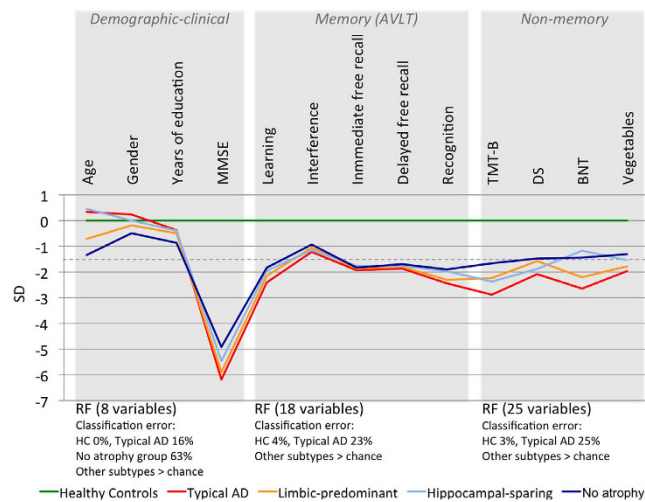
**Figure 2. Cortical thickness and hippocampal volumes.** The brain images show cortical maps of differences in thickness when comparing the different AD subtypes with the healthy controls. A general linear model was fitted at each vertex. Study group was entered as independent variable and TIV as a covariate. Z Monte Carlo simulations were conducted for cluster-forming with a threshold of  $p \leq 0.001$  (two-sided), yielding clusters corrected for multiple comparisons. Only vertexes belonging to clusters surviving this correction are displayed. Significant clusters were mapped on standard templates depicted in lateral (first two images on each row) and medial (last two images on each row) views, both for left (L) and right (R) hemispheres. The coloured bar illustrates the significance level of the differences (i.e. less cortical thickness in the AD patients) from dark blue ( $p \leq 0.05$ ) to light blue ( $p \leq 0.00001$ ). All these results stand after controlling for age, gender, years of education, and APOE  $\epsilon 4$  status (data not shown). The boxplot represents the averaged hippocampal volume controlling for TIV, age, years of education, and APOE  $\epsilon 4$  status. Hence, the y-axis represents adjusted and standardized values. Box values represent median and confidence intervals. The groups' sizes are specified in Fig. 1 and Table 1. AD = Alzheimer's disease.

(A) Description of variables included in the models		
Model 1	Model 2	Model 3
Demographic-clinical variables	Memory variables	Non-memory cognitive variables
Age	AVLT Learning trial 1	WAIS-R Digit Symbol
Gender	AVLT Learning trial 2	WMS-R Digit forward span
Years of education	AVLT Learning trial 3	WMS-R Digit forward length
APOE status	AVLT Learning trial 4	WMS-R Digit backward span
GDS	AVLT Learning trial 5	WMS-R Digit backward length
FAQ	AVLT Learning trial 1 errors	TMT-A correct answers
CDR	AVLT Learning trial 2 errors	TMT-A commission errors
MMSE	AVLT Learning trial 3 errors	TMT-A omission errors
CSF A $\beta_{1-42}$ <sup>a</sup>	AVLT Learning trial 4 errors	TMT-B correct answers
CSF Total tau <sup>a</sup>	AVLT Learning trial 5 errors	TMT-B commission errors
CSF p-tau <sup>a</sup>	AVLT B (interference)	TMT-B omission errors
Disease duration <sup>a</sup>	AVLT B (interference) errors	Verbal Fluency – vegetables
Age of onset <sup>a</sup>	AVLT immediate free recall	Verbal Fluency – vegetables intrusions
	AVLT immediate free recall errors	Verbal Fluency – vegetables perseverations
	AVLT delayed free recall	Verbal Fluency – animals
	AVLT delayed free recall errors	Verbal Fluency – animals intrusions
	AVLT recognition	Verbal Fluency – animals perseverations
	AVLT recognition errors	BNT total
		BNT spontaneous
		BNT number semantic cues
		BNT spontaneous + semantic
		BNT number phonetic cues
		BNT spontaneous + semantic + phonetic
		Clock Test drawing
		Clock Test copy
(B) Summary of results		
Model 1	Model 2	Model 3
N = 415	N = 408	N = 398
Error by chance = 80%	Error by chance = 80%	Error by chance = 80%
Classification error:	Classification error:	Classification error:
Healthy controls: 0%	Healthy controls: 4%	Healthy controls: 3%
Typical AD: 16%	Typical AD: 23%	Typical AD: 25%
Limbic-predominant: 88%	Limbic-predominant: 100%	Limbic-predominant: 100%
Hippocampal-sparing: 82%	Hippocampal-sparing: 100%	Hippocampal-sparing: 97%
No atrophy group: 63%	No atrophy group: 100%	No atrophy group: 100%

**Table 2. Random Forest models.** <sup>a</sup>The same model was repeated including CSF biomarkers in all the sample (model 1<sup>o</sup>), and CSF biomarkers + disease duration + age of onset only in Alzheimer's disease patients (model 1<sup>o</sup>). Results obtained were the same as in model 1. Results presented in the manuscript Fig. 3 are those from model 1 (8 variables). APOE = apolipoprotein E; GDS = geriatric depression scale; FAQ = functional activities questionnaire; CDR = clinical dementia rating; MMSE = mini-mental state examination; CSF = cerebrospinal fluid; A $\beta_{1-42}$  = amyloid- $\beta$ -peptide 1–42; T-tau = total level of tau protein; AVLT = Auditory Verbal Learning test; WAIS-R = Wechsler Adult Intelligence Scale – Revised; WMS-R = Wechsler Memory Scale – Revised; TMT-B = Trail Making Test part B; BNT = Boston Naming Test; AD = Alzheimer's disease.

We then investigated whether these findings in memory could have any clinical relevance by classifying performance as normal or clinically impaired ( $-1.5$  SD) using the healthy controls as reference group. Descriptive analyses confirmed that clinical impairment in learning was more frequent in typical AD than in the other subtypes (Table 3). Clinical impairment in delayed recall and recognition was more frequent in typical AD and limbic-predominant. The hippocampal-sparing group was the subtype having more interference effect. The gain variable (recognition minus delayed free recall) showed that hippocampal-sparing and specially the no atrophy group get more benefit from cues (recognition) than typical AD and limbic-predominant.

**Contribution of non-memory functions to memory performance.** The results above suggest that despite memory impairment looks similar across AD subtypes, the nature of this memory impairment might be different, i.e. non-memory cognitive functions may be contributing differently to performance in memory across subtypes. To test for this we performed separate multiple linear regression models to investigate which non-memory components contributed the most to performance in memory across AD subtypes, and in comparison with the healthy controls. Due to the small sample size for some AD subtypes ( $n \approx 30$ ), results that were not



**Figure 3. Demographic-clinical variables and cognitive profile.** A selection of variables is reported for each random forest (RF) analysis. Classification error is reported only for the correctly classified study groups. The complete list of variables as well as full results from these models are detailed in Table 2. The dashed line shows the clinical cut-off of  $-1.5$  standard deviations (SD). The random forest models reveal great overlap among subtypes with high comparability in demographic, clinical, and cognitive variables at baseline, including cerebrospinal fluid biomarkers, and APOE  $\epsilon 4$  distribution. The groups' sizes are specified in Fig. 1 and Table 1. AD = Alzheimer's disease; HC = healthy controls; MMSE = mini-mental state examination; AVLT = Auditory Verbal Learning test; TMT-B = Trail Making Test part B; DS = digit symbol from the Wechsler Adult Intelligence Scale - Revised; BNT = Boston Naming Test.

	Typical AD (n = 100)	Limbic- predominant (n = 33)	Hippocampal-sparing (n = 35)	No atrophy group (n = 30)
Learning <sup>a</sup>	88.0	75.8	81.8	74.1
Interference <sup>b</sup>	36.6	33.3	42.6	29.6
Immediate <sup>c</sup>	78.0	66.7	63.6	66.7
Delayed <sup>d</sup>	87.1	87.9	82.4	74.1
Recognition <sup>d</sup>	74.2	78.8	64.7	51.9
Gain <sup>d</sup>	12.9	9.1	17.7	22.2

**Table 3. Percentage of AD patients with clinical impairment ( $-1.5$  SD) across memory components.** Values in the table represent percentage. The healthy control group was used as reference for calculating the clinical cut-off of  $-1.5$  SD. Interference was based on the AVLT list B item in order to estimate interference effects during learning. A higher percentage of individuals showing clinical impairment ( $-1.5$  SD) in this variable reflects greater interference effect, which is frequently interpreted as distortions of existing memories possibly due to source-monitoring deficits. Gain was calculated by subtracting the recognition percentage from the delayed percentage in order to quantify benefit from additional help when retrieving stored information; higher gain values represent more retrieving problems. <sup>a</sup>N = 411; <sup>b</sup>N = 412; <sup>c</sup>N = 408; <sup>d</sup>N = 414. AD = Alzheimer's disease.

stable in the non-parametric multiple regression model (i.e. dominance analysis) were rejected and not reported (Table 4). Only learning, delayed free recall and recognition were tested for simplicity.

In the healthy controls, learning was significantly associated with semantic abilities ( $\beta = 0.292$ ) and processing speed ( $\beta = 0.229$ ). The same pattern was obtained for typical AD and limbic-predominant, but for typical AD, lexical access was a significant predictor of learning as well ( $\beta = 0.219$ ).

Delayed free recall was significantly predicted by learning alone in healthy controls ( $\beta = 0.744$ ), typical AD ( $\beta = 0.560$ ), and limbic-predominant ( $\beta = 0.483$ ); and by both learning ( $\beta = 0.725$ ) and attention/processing speed ( $\beta = -0.336$ ) in the no atrophy group. In hippocampal-sparing, delayed free recall was associated to executive functioning ( $\beta = 0.548$ ) and semantic abilities ( $\beta = 0.455$ ), but not learning.

Finally, recognition was significantly predicted by learning alone in healthy controls ( $\beta = 0.469$ ) and typical AD ( $\beta = 0.486$ ); and by learning and lexical access in the no atrophy group ( $\beta = 0.445$  and  $\beta = 0.357$ , respectively). Dominance analysis showed that semantic abilities are also important for recognition in the no atrophy group.

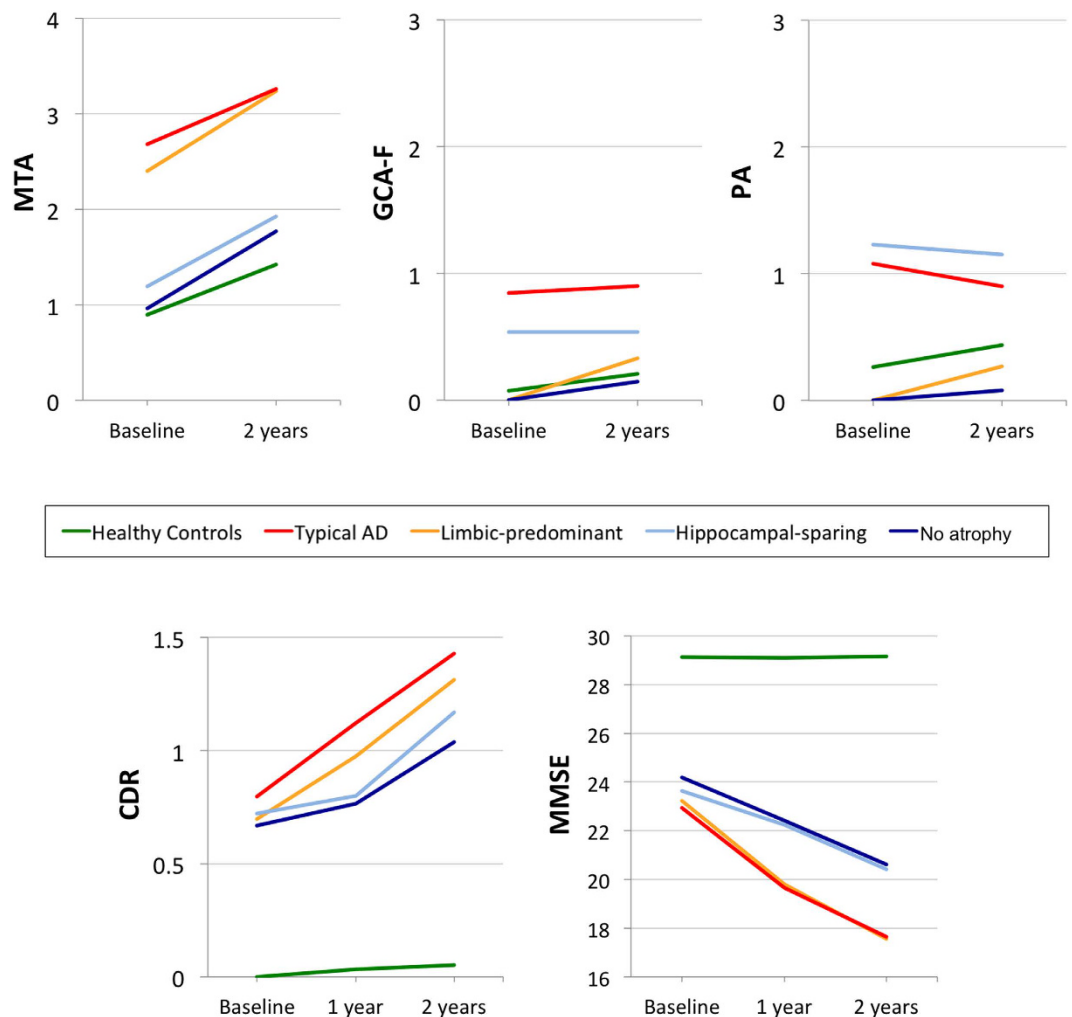
**Disease progression over two years.** Figure 4 shows the longitudinal progression in the visual rating scores across study groups.

		TMT-B	DS	BNT	Vegetables	Learning
<b>Learning</b>						
	<b>Multiple regression model</b>					
Healthy controls	$F_{(2, 222)} = 22.065; p < 0.001; R^2_{adj.} = 16\%$		<b>0.229</b>		<b>0.292</b>	—
Typical AD	$F_{(3, 80)} = 15.880; p < 0.001; R^2_{adj.} = 35\%$		<b>0.264</b>	<b>0.219</b>	<b>0.363</b>	—
Limbic-predominant	$F_{(2, 30)} = 13.526; p < 0.001; R^2_{adj.} = 44\%$		<b>0.340</b>		<b>0.510</b>	—
Hippocampal-sparing	<i>non significant</i>					—
No atrophy group	<i>non significant</i>					—
	<b>Dominance analysis (bootstrapping)</b>					
Healthy controls	Explained variance = 18%	1%	<b>6%</b>	1%	<b>10%</b>	—
Typical AD	Explained variance = 38%	4%	<b>8%</b>	<b>9%</b>	<b>17%</b>	—
Limbic-predominant	Explained variance = 49%	10%	<b>11%</b>	5%	<b>23%</b>	—
<b>Delayed free recall</b>						
	<b>Multiple regression model</b>					
Healthy controls	$F_{(1, 223)} = 276.979; p < 0.001; R^2_{adj.} = 55\%$					<b>0.744</b>
Typical AD	$F_{(2, 81)} = 15.080; p < 0.001; R^2_{adj.} = 25\%$		<b>-0.249*</b>			<b>0.560</b>
Limbic-predominant	$F_{(1, 31)} = 9.435; p = 0.008; R^2_{adj.} = 21\%$					<b>0.483</b>
Hippocampal-sparing	$F_{(2, 27)} = 5.225; p = 0.012; R^2_{adj.} = 23\%$	<b>0.548</b>			<b>0.455</b>	
No atrophy group	$F_{(2, 22)} = 17.815; p < 0.001; R^2_{adj.} = 58\%$		<b>-0.336</b>			<b>0.725</b>
	<b>Dominance analysis (bootstrapping)</b>					
Healthy controls	Explained variance = 56%	0%	3%	1%	3%	<b>49%</b>
Typical AD	Explained variance = 28%	0%	<b>2%*</b>	1%	2%	<b>23%</b>
Limbic-predominant	Explained variance = 26%	1%	2%	0%	2%	<b>21%</b>
Hippocampal-sparing	Explained variance = 33%	<b>14%</b>	5%	3%	<b>10%</b>	1%
No atrophy group	Explained variance = 66%	6%	<b>8%</b>	2%	6%	<b>44%</b>
<b>Recognition</b>						
	<b>Multiple regression model</b>					
Healthy controls	$F_{(1, 223)} = 62.858; p < 0.001; R^2_{adj.} = 22\%$					<b>0.469</b>
Typical AD	$F_{(2, 81)} = 11.366; p < 0.001; R^2_{adj.} = 20\%$	<b>0.253*</b>				<b>0.486</b>
Limbic-predominant	<i>non significant</i>					
Hippocampal-sparing	<i>non significant</i>					
No atrophy group	$F_{(2, 22)} = 5.890; p = 0.027; R^2_{adj.} = 29\%$			<b>0.357</b>		<b>0.445</b>
	<b>Dominance analysis (bootstrapping)</b>					
Healthy controls	Explained variance = 24%	1%	1%	1%	1%	<b>20%</b>
Typical AD	Explained variance = 24%	<b>3%*</b>	1%	2%	1%	<b>17%</b>
No atrophy group	Explained variance = 37%	1%	1%	<b>8%</b>	10%	<b>17%</b>

**Table 4. Contribution of non-memory cognitive functions to different memory measurements.** Values in the table represent standardized beta (for multiple regression model) and the fraction of the variance explained by a predictor (dominance analysis). In addition to multiple regression analyses, confirmatory dominance analyses were performed due to the small sample size in some of the AD subtypes. The asterisk (\*) in the table indicates when a significant predictor in the regression models (in bold) was rejected due to lack of stability in the dominance analysis (i.e. the percentage in bold is low, or equal to the one in not-bold). All p-values are corrected with Hochberg's correction for multiple comparisons. AD = Alzheimer's disease; adj. = adjusted; TMT-B = Trail Making Test part B; DS = digit symbol from the Wechsler Adult Intelligence Scale - Revised; BNT = Boston Naming Test.

The mixed effects model showed that clinical progression (CDR) over 2 years was faster in all the AD subtypes (estimates between 0.14 and 0.29) than in the healthy controls (estimate = 0.03) ( $F_{(4, 690)} = 46.896; p < 0.001$ ) (Fig. 4). Moreover, the slope was greater in typical AD (estimate = 0.29) than in hippocampal-sparing (estimate = 0.16;  $t_{(728)} = -2.998; p = 0.015$ ) and the no atrophy group (estimate = 0.14;  $t_{(680)} = -3.443; p < 0.001$ ); and in limbic-predominant (estimate = 0.28) than in hippocampal-sparing ( $t_{(708)} = -2.374$ ; trend for significance:  $p = 0.054$ ) and the no atrophy group ( $t_{(672)} = -2.746; p = 0.024$ ). No significant effects were obtained for the quadratic model, indicating that clinical progression in all the subtypes is linear rather than quadratic.

Regarding global cognitive decline (MMSE), all the AD subtypes had faster decline over 2 years (estimates between -3.08 and -1.48) than the healthy controls (estimate = -0.01) ( $F_{(4, 676)} = 57.432; p < 0.001$ ) (Fig. 4). Moreover, the slope was greater in typical AD (estimate = -2.69) than in hippocampal-sparing (estimate = -1.48;  $t_{(704)} = 3.152; p = 0.010$ ); and in limbic-predominant (estimate = -3.08) than in hippocampal-sparing ( $t_{(688)} = 3.551; p < 0.001$ ) and the no atrophy group ( $t_{(663)} = 2.850; p = 0.020$ ). No significant effects were obtained for the quadratic model, indicating that cognitive decline in all the subtypes is linear rather than quadratic. All these results stand after controlling for age, gender, and years of education (data not shown).



**Figure 4. Disease progression over two years.** Longitudinal scores for the visual rating scales (i.e. MTA, GCA-F, and PA) were available for 190 cases (110 healthy controls, 39 typical AD, 15 limbic-predominant, 13 hippocampal-sparing, 13 no atrophy group). Scores from MTA left and MTA right were averaged for simpler representation since longitudinal progression was similar in the two of them. Longitudinal values for CDR were available for 229 cases (147 healthy controls, 41 typical AD, 16 limbic-predominant, 12 hippocampal-sparing, 13 no atrophy group). Mixed effects analysis uses all data available at each time point. MTA = medial temporal atrophy visual rating scale; posterior atrophy visual rating scale; global cortical atrophy visual rating scale - frontal subscale; AD = Alzheimer's disease; HC = healthy controls; CDR = clinical dementia rating; MMSE = mini-mental state examination.

## Discussion

The AD subtypes investigated here have consistently been identified in previous studies. However, a method to enable their investigation in the clinical routine has been lacking so far. In the current study, visual rating scales were used as subtyping method because they are the primary method for assessing brain structural changes in clinical settings<sup>18–21</sup>. Although grouped data is reported, the visual rating scales can be applied at the individual level for clinical diagnosis. The aims in the current study were to (1) validate the combined use of visual rating scales for identification of AD subtypes; (2) characterize the resulting subtypes; and (3) investigate how atrophy patterns and non-memory cognitive domains contribute to memory impairment. The findings show that visual rating scales help to identify distinct AD subtypes with different disease progression. However, routine cognitive and clinical evaluations, CSF biomarkers and APOE  $\epsilon$ 4 genotype did not allow such discrimination, which may limit their clinical use for subtypes identification. Typical AD was the most impaired subtype at baseline and together with limbic predominant AD had worse clinical progression. Identification of patients with hippocampal-sparing and no atrophy is also clinically relevant since they showed less aggressive disease progression.

Visual rating scales are feasible and reproducible<sup>19,21–25</sup>. Intra-rater values (weighted  $\kappa$ ) are usually around 0.90 for MTA<sup>19,23</sup>, and between 0.70 and 0.90 for GCA and PA<sup>19,23,25</sup>. Inter-rater values (weighted  $\kappa$ ) are usually around 0.85 for MTA<sup>23</sup>, and between 0.60 and 0.80 for GCA and PA<sup>23,25</sup>. These weighted  $\kappa$  values correspond to substantial and almost perfect agreement<sup>26</sup>, thus proving their reproducibility.



The subtypes in this study were defined a-priori based on converging evidence suggesting three AD subtypes<sup>1–5,9–11,14,15</sup>. Vertex analyses in the current study showed that visual rating scales can successfully identify patterns of atrophy similar to those depicted by previous sophisticated MRI studies<sup>1,2,11,14,15</sup>, and coherent with those tracking the spread of NFT<sup>9</sup>. The prevalence of different subtypes obtained in this study are very similar to those previously reported, where typical AD usually includes 50–75% of the AD patients, limbic-predominant is referred in around 15–35%, and hippocampal-sparing in around 10–25%<sup>1,4,9,11,15</sup>. In the only study identifying a no atrophy group, the prevalence was 10%<sup>11</sup>. Since visual rating scales can be easily applied in the clinical routine<sup>18</sup>, this finding may have significant impact for current diagnosis and management of AD patients in clinical settings. The fact that the four subtypes were rather comparable on cognition, CSF biomarkers and APOE ε4 distribution at baseline further supports the use of MRI to identify clinically relevant subtypes. Different longitudinal progression of these subtypes supports this approach as well.

There was great overlap in the clinical and cognitive profiles, especially among the atypical AD subtypes, which highlights the difficulty in detecting these in routine clinical evaluations by only using clinical or cognitive measurements. Despite this, when comparing our results with other studies, the characteristics of our AD subtypes are largely comparable with what has previously been reported. The typical AD subtype has previously been found to be among the oldest<sup>1,2,4,9,11,15</sup>, have later onset<sup>1,4,9,11,15</sup>, include a higher frequency of males<sup>2,11</sup>, and have similar disease duration to limbic-predominant and hippocampal-sparing<sup>1,9,11</sup>. Contrary to our finding, hippocampal-predominant AD has been found to be among the oldest groups<sup>1,4,9,15</sup> and to have shorter disease duration<sup>14</sup>. An explanation could be that atrophy in the medial temporal lobe is frequent in normal aging, hence clinical cut-offs for the medial temporal atrophy (MTA) scale are age-corrected<sup>19</sup>. This could lead to a younger hippocampal-predominant subtype in the present study. Age-corrections were not performed in previous studies except in Byun *et al.*<sup>11</sup> and Varol *et al.*<sup>2</sup>, who indeed showed consistent results with our findings. Also contrary to our finding, hippocampal-sparing has been related with younger age<sup>1,4,9,15</sup>, earlier onset<sup>1,4,9,14,15</sup>, and shorter disease duration<sup>4</sup>. These results in previous studies could be explained by higher prevalence of early-onset AD in their samples, known to display predominant posterior atrophy and a more aggressive presentation<sup>9,27</sup>. However, early-onset AD is not common in ADNI-1, our sample, where posterior atrophy possibly results from a different process perhaps related or amplified by increased age.

To our knowledge, only our study and the one by Byun *et al.*<sup>11</sup> have investigated a group of AD patients with no atrophy. Although data from ADNI-1 was used in both studies, we included 198 AD patients while Byun *et al.*<sup>11</sup> included 163 AD patients. Most of the drops in Byun *et al.*<sup>11</sup> corresponds to the hippocampal-sparing and the no atrophy group groups. Visual inspection of the demographic and clinical characteristics reveals that both groups are largely comparable, although our no atrophy group is younger and has higher frequency of females. An age correction was performed in both studies but the gender correction was not performed in our study because scores in our scales of posterior and frontal atrophy are not influenced by gender, and scores in MTA are only marginally influenced by gender<sup>19</sup>. Thus, the gender correction and the fact that most of the dropped cases in the no atrophy group from Byun *et al.*<sup>11</sup> were young females, could be the explanation for these differences.

Interactions between memory components and contribution of non-memory cognitive functions were investigated for the first time in a study of this kind. Learning capacity was compromised in the four subtypes, but typical AD and limbic-predominant evidenced more consolidation problems, while hippocampal-sparing and the no atrophy group showed more problems in free retrieval of information. In addition, hippocampal-sparing was the group showing greater vulnerability to interference. Similar results were obtained in another study using ADNI-1 data<sup>5</sup>. The cognitive profiles discussed above were further confirmed in the regression/dominance analysis, where delayed recall and recognition were more dependent on non-memory cognitive functions in hippocampal-sparing and the no atrophy group than in typical AD and limbic-predominant. These profiles are coherent with the underlying pattern of brain atrophy. Noh *et al.*<sup>1</sup> found prominent memory impairment in their parietal predominant AD subtype (analogous to hippocampal-sparing). The authors suggested that memory deficits might be associated with attention and working memory dysfunction in their parietal predominant subtype<sup>1</sup>. Hence, despite great overlap among AD subtypes, the nature of memory impairment seems to be different<sup>2</sup>. In this regard, Whitwell *et al.*<sup>9</sup> made an important observation: a patient with Alzheimer's dementia dominated by memory impairment can have any of the three subtypes of AD. This is exactly what we demonstrate in the current study, since all our AD patients have memory impairment, and we further demonstrate that the determinants of this memory impairment are varied and correspond with different atrophy patterns. This finding may have clinical utility, both for diagnosis/prognosis and cognitive interventions. Likewise, it is very important to note that these cognitive profiles were obtained using multivariate statistical methods on grouped data. Such information can be very difficult to find at the individual level in clinical routine, not to mention with often used cognitive tests. Thus, the use of visual rating scales seems to be extremely beneficial in this context.

One limitation in the current study is that ADNI-1 is a quite homogeneous sample. All AD patients fulfil the amnesic criteria at entry and aspects such as vascular pathology are excluded. Our results should thus be replicated in a more heterogeneous clinical sample that also includes non-amnesic AD presentations. It could also be argued that the different subtypes identified here reflect AD patients at different stages of the disease (e.g. typical AD being a later stage of limbic-predominant AD), rather than truly distinct AD subtypes. However, it has been demonstrated that these subtypes result from differential spread of NFT<sup>4</sup>. Further, no differences were observed on scales that stage the disease such as CDR, neither on disease duration. Related to this, it could be argued that the no atrophy group might be an initial stage of the disease but, again, no differences on CDR or disease duration were obtained. Alternatively, it could be argued that visual rating scales might not be sensitive enough to capture subtle atrophy in this no atrophy group. However, automated MRI methods confirmed this lack of atrophy in our vertex analysis, and a similar result was obtained using voxel-based morphometry in another ADNI-1 study<sup>11</sup>. Finally, we performed group analyses in order to characterize the different AD subtypes. Due to large

within-group variability, there was great overlapping between subtypes and some conclusions especially on the non-imaging variables may be difficult to translate to the individual level, where clinical diagnosis takes place.

In conclusion, visual rating scales can be used to identify distinct and clinically relevant AD subtypes. To note, these subtypes could not be properly discriminated based on other common clinical tools such as cognitive tests, CSF biomarkers or APOE  $\epsilon$ 4 distribution. It has previously been argued that heterogeneity in AD complicates overcoming the two main challenges in the field at present, namely, discovery of disease-modifying treatments and achievement of accurate diagnosis and clinical prognosis<sup>1–3,10,28</sup>. It has also been discussed that advancing in the knowledge on different AD subtypes could shed some light to recently failed clinical trials by enabling tailored treatments in more homogeneous subgroups of patients<sup>2,28</sup>. The same would also help to better discriminate between highly overlapping clinical phenotypes such as AD with significant frontal involvement and frontotemporal lobe dementia<sup>4,10,18,29</sup>. Findings from the current study suggest that visual rating scales may facilitate investigation of AD heterogeneity in clinical routine. Implementing AD subtypes based on visual rating scales in the clinical routine should be easy and it is warranted to do as soon as possible in order to test its potential clinical impact. Whether using this method clinically may improve diagnosis and prognosis in “real world” AD patients stills needs to be determined.

## Methods

**Participants.** AD patients and healthy controls from the ADNI-1 (adni.loni.usc.edu, PI Michael M. Weiner) with longitudinal data available at 1 and 2 years follow-up were selected for the current study, giving a total of 198 AD patients and 230 healthy controls. ADNI was launched in 2003 by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, the Food and Drug Administration, private pharmaceutical companies, and non-profit organizations<sup>12</sup>. The project was established to develop standardized imaging techniques and biomarkers in AD research. The AD patients and healthy controls were clinically diagnosed following standard procedures as detailed before<sup>19</sup>. Of note, memory impairment was required for all the AD patients based on Logical Memory II (Wechsler Memory Scale – Revised, WMS-R<sup>30</sup>). All diagnoses were made without the use of MRI scans. The study was approved by the institutional review boards of all participating ADNI centres. Written informed consent was obtained from all participants or authorized representatives after extensive description of the ADNI according to the Declaration of Helsinki. All methods were performed in accordance with the relevant guidelines and regulations.

**Magnetic resonance imaging, automated image processing, and visual rating scales.** A 3D T1-weighted magnetization-prepared rapid gradient-echo (MPRAGE) sequence was acquired on 1.5 T MRI scanners (voxel size  $1.1 \times 1.1 \times 1.2 \text{ mm}^3$ )<sup>12</sup>. TheHiveDB Database system<sup>31</sup> was used to automatically process the images with FreeSurfer 5.3.0, following previously described standard procedures<sup>32</sup>. This procedure provides measurements of cortical thickness at the vertex level, as well as a measurement of the total intracranial volume (TIV).

Regional atrophy was measured with visual rating scales based only the T1-weighted images as detailed elsewhere<sup>19</sup>. Briefly, atrophy in the medial temporal lobe was evaluated with the MTA scale<sup>33</sup>; 2) atrophy in the posterior cortex was evaluated with the posterior atrophy (PA) scale<sup>23</sup>; and 3) atrophy in the frontal lobe was evaluated with the global cortical atrophy scale – frontal subscale (GCA-F)<sup>25</sup>. The MTA scale scores the degree of atrophy from zero to four in the hippocampus, parahippocampal gyrus, entorhinal cortex and the surrounding cerebrospinal fluid spaces. The PA scale scores the degree of atrophy from zero to three in the posterior cingulate sulcus, precuneus, parieto-occipital sulcus and the parietal cortex. The GCA-F scale scores the degree of atrophy from zero to three in the frontal lobe as delimited by the central sulcus, the frontal bone and the fissure of Sylvius. Therefore, the three scales primarily reflect cortical atrophy. Nonetheless, these scales also indirectly reflect ventricular enlargement since the inferior lateral ventricles are considered in MTA ratings, and GCA and PA ratings are based on widening of sulci, which is frequently correlated with ventricular enlargement. This is relevant because AD patients have more brain atrophy and larger ventricles than controls in the ADNI-1 cohort<sup>34</sup>. On the other hand, these scales are not designed to capture white matter hyperintensities (WMH), which are commonly regarded as markers of cerebrovascular disease. However, WMH burden in ADNI-1 is minimal relative to other cohorts<sup>35</sup> due to exclusion of individuals with a Hachinski Ischemic Score<sup>36</sup> of more than or equal to 5.

In the three visual rating scales, a score of zero denotes no atrophy, whereas scores from one to four indicate an increasing degree of atrophy. MTA analysis was based on coronal reconstructions, GCA-F on axial reconstructions and PA on reconstructions from all three planes. The images were rated both at baseline and at two years follow-up. Subtyping is based on baseline ratings, while longitudinal ratings were used to study disease progression over two years.

All cases were rated by an experienced radiologist (L.C.). Intra-rater reliability in 120 random cases achieved a weighted  $\kappa$  of 0.94 and 0.89 for MTA in left and right hemispheres, respectively, 0.88 for PA, and 0.83 for GCA-F. The same 120 random cases were also rated by a newly trained radiologist (C.-J.G.) for inter-rater analysis: weighted  $\kappa$  of 0.71 and 0.70 for MTA in left and right hemispheres, respectively, 0.88 for PA, and 0.79 for GCA-F. Both raters were blind to any information about the participants.

**AD subtypes based on patterns of brain atrophy.** Deviation from normality was established following a recently proposed list of practical cut-offs<sup>19</sup>. The MTA scores  $\geq 1.5$ ,  $\geq 1.5$ ,  $\geq 2$ ,  $\geq 2.5$  were considered abnormal for the respective age ranges 45–64, 65–74, 75–84, and 85–94 years. A previous study using the same dataset as here demonstrated that an age-correction does not improve PA and GCA-F diagnostic performance<sup>19</sup>. Therefore, the same cut-off was used for PA and GCA-F. A score  $\geq 1$  was considered abnormal irrespectively of the age range<sup>19</sup>. The three AD subtypes identified in previous literature<sup>4,9</sup> were defined based on the combination of MTA,

PA, and GCA-F as follows (see also Fig. 1). The typical AD subtype was defined as atrophy in the medial temporal lobe (abnormal MTA) together with atrophy in the posterior cortex (abnormal PA) and/or frontal cortex (abnormal GCA-F). The limbic-predominant subtype was defined as atrophy in the medial temporal lobe alone (abnormal MTA with normal PA and GCA-F). The hippocampal-sparing subtype included atrophy in the posterior cortex (abnormal PA) and/or frontal cortex (abnormal GCA-F), but not in the medial temporal lobe (normal MTA). A group with no atrophy was also identified as in Byun *et al.*<sup>11</sup> when AD patients displayed normal scores in MTA, PA, and GCA-F.

**Demographic and clinical variables.** Age, gender, and years of education were included as demographic variables. Clinical severity was assessed with the CDR<sup>37</sup> scale and global cognition with the MMSE<sup>38</sup>. FAQ<sup>39</sup> was used to measure functional activities of daily living and GDS<sup>40</sup> to measure depressive symptomatology. Age at disease onset, disease duration, and APOE  $\epsilon$ 4 status were also measured. Memory was assessed with the Auditory Verbal Learning test (AVLT)<sup>41</sup>. The different AVLT items were used to investigate different memory components. In particular, the sum of the five learning trials of the list A reflects learning capacity. Performance in the list B served as an estimate of interference effects during learning. Interference effects are frequently interpreted as distortions of existing memories possibly due to source-monitoring deficits. Free recall of the list A right after recall of the list B measures immediate recall. Free recall of the list A 30 minutes after learning measures delayed recall. Recognition reflects the ability to identify previously learned items from the list A. Finally, gain was calculated by subtracting the recognition percentage from the delayed percentage in order to quantify benefit from additional help when retrieving stored information. Executive functions (Trail Making Test part B, TMT-B)<sup>42</sup>, attention/processing speed (Digit Symbol, DS, from the Wechsler Adult Intelligence Scale – Revised, WAIS-R)<sup>43</sup>, language (Boston Naming Test, BNT)<sup>44</sup>, and semantic abilities (semantic fluency, vegetables)<sup>45</sup> were also assessed. Digit span from the WMS-R<sup>30</sup> and the Clock Test<sup>46</sup> were further included for the random forest models (see statistical analysis). CSF samples were available for 102 AD patients and 115 healthy controls. Complete procedure descriptions are available at [www.adni-info.org](http://www.adni-info.org).

**Statistical analysis.** Mixed effects models (fixed and random effects) and mixed ANOVA/ANCOVA (split plot) were used to analyse the interaction between a between-subjects factor (study group) and a within-subjects factor (memory component and time). In the mixed effects models, the fixed-effect factors were study group, time, and the study group-by-time interaction. The random effect factor was the participants. When time was included in the model, both linear and quadratic effects were tested in order to investigate whether for example disease progression is linear or gets accelerated/decelerated after a certain time point. Multiple linear regression (backwards) was performed to analyse the contribution of non-memory cognitive functions to different memory components. Confirmatory dominance analyses, an extension of multiple regression, were performed with a non-parametric test based on bootstrapping (1000 iterations), which is less vulnerable to small sample sizes. Random forest analysis (500 trees) was also used to investigate differences between groups in multiple variables while avoiding multiple testing. P-values in all principal and post-hoc analyses were adjusted with the Benjamini-Hochberg's<sup>47</sup> correction for multiple comparisons. Model assumptions were tested in all the cases by visual inspection of residuals and data distribution, as well as by inspecting the pertinent statistical parameters. Results were considered significant when  $p \leq 0.05$  (two-tailed).

Image analyses based on the vertex across the cortical mantle were carried out using FreeSurfer software as detailed elsewhere<sup>48</sup>. Briefly, maps were smoothed using a circularly symmetric Gaussian kernel across the surface with a full width at half maximum (FWHM) of 10 mm. A general linear model was fitted at each vertex. Study group was entered as independent variable and TIV as a covariate. Z Monte Carlo simulations were used with a cluster-forming threshold of  $p \leq 0.001$  (two-sided), yielding results corrected for multiple comparisons.

## References

- Noh, Y. *et al.* Anatomical heterogeneity of Alzheimer disease: based on cortical thickness on MRIs. *Neurology* **83**, 1936–1944 (2014).
- Varol, E., Sotiras, A. & Davatzikos, C. HYDRA: revealing heterogeneity of imaging and genetic patterns through a multiple max-margin discriminative analysis framework. *Neuroimage* pii: S1053-8119(16)00150-6. doi: 10.1016/j.neuroimage.2016.02.041 (2016).
- Scheltens, N. M. E. *et al.* The identification of cognitive subtypes in Alzheimer's disease dementia using latent class analysis. *J. Neurol. Neurosurg. Psychiatry* **87**, 235–243 (2016).
- Murray, M. E. *et al.* Neuropathologically defined subtypes of Alzheimer's disease with distinct clinical characteristics: a retrospective study. *Lancet. Neurol.* **10**, 785–96 (2011).
- Peter, J. *et al.* Subgroups of Alzheimer's disease: stability of empirical clusters over time. *J. Alzheimers. Dis.* **42**, 651–661 (2014).
- McKhann, G. M. *et al.* The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers. Dement.* **7**, 263–269 (2011).
- Morris, J. C. *et al.* Harmonized diagnostic criteria for Alzheimer's disease: recommendations. *J. Intern. Med.* **275**, 204–213 (2014).
- Dubois, B. *et al.* Revising the definition of Alzheimer's disease: a new lexicon. *Lancet Neurol.* **9**, 1118–1127 (2010).
- Whitwell, J. L. *et al.* Neuroimaging correlates of pathologically defined subtypes of Alzheimer's disease: a case-control study. *Lancet. Neurol.* **11**, 868–877 (2012).
- Whitwell, J. L. *et al.* Temporoparietal atrophy: a marker of AD pathology independent of clinical diagnosis. *Neurobiol. Aging* **32**, 1531–1541 (2011).
- Byun, M. S. *et al.* Heterogeneity of regional brain atrophy patterns associated with distinct progression rates in Alzheimer's disease. *PLoS One* **10**, e0142756 (2015).
- Mueller, S. G. *et al.* The Alzheimer's disease neuroimaging initiative. *Neuroimaging Clin. N. Am.* **15**, 869–877 (2005).
- Jack, C. R. *et al.* The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *J. Magn. Reson. Imaging* **27**, 685–991 (2008).
- Shiino, A. *et al.* Four subgroups of Alzheimer's disease based on patterns of atrophy using VBM and a unique pattern for early onset disease. *Neuroimage* **33**, 17–26 (2006).

15. Hwang, J. *et al.* Prediction of Alzheimer's disease pathophysiology based on cortical thickness patterns. *Alzheimer's Dement.* (Amst) **2**, 58–67 (2015).
16. Davidson, J. E. *et al.* An exploration of cognitive subgroups in Alzheimer's disease. *J. Int. Neuropsychol. Soc.* **16**, 233–243 (2010).
17. Stopford, C. L., Snowden, J. S., Thompson, J. C. & Neary, D. Variability in cognitive presentation of Alzheimer's disease. *Cortex* **44**, 185–195 (2008).
18. Harper, L. *et al.* MRI visual rating scales in the diagnosis of dementia: evaluation in 184 post-mortem confirmed cases. *Brain* **139**, 1211–1225 (2016).
19. Ferreira, D. *et al.* Practical cut-offs for visual rating scales of medial temporal, frontal, and posterior atrophy in Alzheimer's disease and mild cognitive impairment. *J. Intern. Med.* **278**, 277–290 (2015).
20. Ferreira, D. *et al.* Electroencephalography is a good complement to currently established dementia biomarkers. *Dement. Geriatr. Cogn. Disord.* **42**, 80–92 (2016).
21. van der Flier, W. M. *et al.* Optimizing patient care and research: the Amsterdam Dementia Cohort. *J. Alzheimers. Dis.* **41**, 313–327 (2014).
22. van de Pol, L. A. & Scheltens, P. Medial temporal lobe atrophy scores translated to clinical practice: editorial comment on 'influence of age, disease onset and ApoE4 on visual medial temporal lobe atrophy cut-offs'. *J. Intern. Med.* **275**, 331–333 (2014).
23. Koedam, E. L. *et al.* Visual assessment of posterior atrophy development of a MRI rating scale. *Eur. Radiol.* **21**, 2618–2625 (2011).
24. Scheltens, P., Pasquier, F., Weerts, J., Barkhof, F. & Leys, D. Qualitative assessment of cerebral atrophy on MRI: inter- and intra-observer reproducibility in dementia and normal aging. *Eur. Neurol.* **37**, 95–99 (1997).
25. Ferreira, D. *et al.* Quantitative validation of a visual rating scale for frontal atrophy: associations with clinical status, APOE e4, CSF biomarkers and cognition. *Eur. Radiol.* **26**, 2597–2610 (2015).
26. Landis, J. & Koch, G. The measurement of observer agreement for categorical data. *Biometrics* **33**, 159–174 (1977).
27. van der Flier, W. M., Pijnenburg, Y. Al, Fox, N. C. & Scheltens, P. Early-onset versus late-onset Alzheimer's disease: the case of the missing APOE e4 allele. *Lancet Neurol.* **10**, 280–288 (2011).
28. Kovacs, G. G. Clinical stratification of subtypes of Alzheimer's disease. *Lancet. Neurol.* **11**, 839–841 (2012).
29. Rivero-Santana, A. *et al.* Cerebrospinal fluid biomarkers for the differential diagnosis between Alzheimer's disease and frontotemporal lobar degeneration: systematic review, HSROC analysis, and confounding factors. *J. Alzheimer's Dis.* **55**, 625–644 (2016).
30. Wechsler, D. *Wechsler Memory Scale - Third Edition. Technical Manual.* (The Psychological Corporation, 1997).
31. Muehlboeck, J.-S., Westman, E. & Simmons, A. TheHiveDB image data management and analysis framework. *Front. Neuroinform.* **7**, 49 (2014).
32. Ferreira, D. *et al.* Cognitive decline is mediated by gray matter changes during middle age. *Neurobiol. Aging* **35**, 1086–1094 (2014).
33. Scheltens, P. *et al.* Atrophy of medial temporal lobes on MRI in 'probable' Alzheimer's disease and normal ageing: diagnostic value and neuropsychological correlates. *J. Neurol. Neurosurg. Psychiatry* **55**, 967–972 (1992).
34. Orellana, C. *et al.* Measuring global brain atrophy with the brain volume/cerebrospinal fluid index: normative values, cut-offs and clinical associations. *Neurodegener. Dis.* **16**, 77–86 (2016).
35. Ramirez, J., McNeely, A. A., Scott, C. J. M., Masellis, M. & Black, S. E. White matter hyperintensity burden in elderly cohort studies. The Sunnybrook Dementia Study, Alzheimer's Disease Neuroimaging Initiative, and Three-City Study. *Alzheimers Dement.* **12**, 203–210 (2016).
36. Hachinski, V. C. *et al.* Cerebral blood flow in dementia. *Arch Neurol* **32**, 632–637 (1975).
37. Morris, J. C. The Clinical Dementia Rating (CDR): current vision and scoring rules *Neurology* **43**, 2412–2414 (1993).
38. Folstein, M. F., Folstein, S. E. & McHugh, P. R. Mini-mental- state. A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* **12**, 189–198 (1975).
39. Pfeffer, R. I., Kurosaki, T. T., Harrah, C. H., Chance, J. M. & Filos, S. Measurement of functional activities in older adults in the community. *J. Gerontol.* **37**, 323–329 (1982).
40. Sheikh, J. & Yesavage, J. Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. *Clin Gerontol* **5**, 165–173 (1986).
41. Schmidt, M. *Rey Auditory Verbal Learning Test: A Handbook.* (Western Psychological Services, 1996).
42. Reitan, R. Validity of the trail making test as an indicator of organic brain damage. *Percept. Mot. Skills* **8**, 271–276 (1958).
43. Wechsler, D. *Wechsler Adult Intelligence Scale - Administration and Scoring Manual, Third Edition.* (The Psychological Corporation, San Antonio, TX., 1997).
44. Kaplan, E., Goodglass, H. & Weintraub, S. *Boston Naming Test.* (Lea & Febiger, 1983).
45. Benton, A., Hamsher, K. & Sivan, A. *Multilingual Aphasia Examination- Second Edition.* (AJA Associates, University of Iowa., 1989).
46. Goodglass, H. & Kaplan, E. *The Assessment Of Aphasia And Related Disorders.* (Lea & Febiger., 1983).
47. Hochberg, Y. & Benjamin, Y. More powerful procedures for multiple significance testing. *Stat. Med.* **9**, 811–818 (1990).
48. Ferreira, D. *et al.* Different reserve proxies confer overlapping and unique endurance to cortical thinning in healthy middle-aged adults. *Behav. Brain Res.* **311**, 375–383 (2016).

## Acknowledgements

This project is financially supported by the Swedish Foundation for Strategic Research (SSF). The authors also thank the Strategic Research Programme in Neuroscience at Karolinska Institutet (StratNeuro), Åke Wiberg foundation, Hjärnfonden, the Swedish research council (VR) Alzheimerfonden and Birgitta och Sten Westerberg for additional financial support. Data collection and sharing for the ADNI 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.; 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](http://www.fnih.org)). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. Contributors from the ADNI

consortium are detailed in Supplementary Table 1. The above-mentioned funding sources had no involvement in the study design, collection, analysis and interpretation of the data, writing of the report and the decision to submit the article for publication. 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 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](http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf).

### Author Contributions

The study design and concept was done by D.F., C.V., J.A.H.C. and E.W. Acquisition, analysis, and interpretation of data was done by D.F., C.V., J.A.H.C., L.C., C.J.G., U.E., J.S.M., A.S., J.B., L.O.W. and E.W. Drafting of the manuscript was done by D.F. and E.W. Critical revision of the manuscript for important intellectual content was done by D.F., C.V., J.A.H.C., L.C., C.J.G., U.E., J.S.M., A.S., J.B., L.O.W. and E.W. Statistical analyses were carried out by D.F., C.V. and J.A.H.C. L.O.W. and E.W. obtained funding. D.F., C.V., J.A.H.C., L.C., C.J.G., U.E., J.S.M., A.S., J.B., L.O.W. and E.W. provided admin.

### Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing Interests:** The authors declare no competing financial interests.

**How to cite this article:** Ferreira, D. *et al.* Distinct subtypes of Alzheimer's disease based on patterns of brain atrophy: longitudinal trajectories and clinical applications. *Sci. Rep.* **7**, 46263; doi: 10.1038/srep46263 (2017).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017