Journal of Alzheimer's Disease 42 (2014) 651–661
DOI 10.3233/JAD-140261
IOS Press

Subgroups of Alzheimer's Disease: Stability of Empirical Clusters Over Time

Jessica Peter[a,b,c,*], Ahmed Abdulkadir[b,d], Christoph Kaller[a,b], Dorothee Kümmerer[a,b], Michael Hüll[e], Werner Vach[f] and Stefan Klöppel[a,b,g] for the Alzheimer's Disease Neuroimaging Initiative[1]
[a]*Department of Neurology, University Medical Centre Freiburg, Germany*
[b]*Freiburg Brain Imaging, University Medical Centre Freiburg, Germany*
[c]*Laboratory for Biological and Personality Psychology, Department of Psychology, University of Freiburg, Germany*
[d]*Group of Pattern Recognition and Image Processing, Department of Computer Science, University of Freiburg, Germany*
[e]*Centre for Geriatric Medicine and Gerontology, University Medical Centre, Freiburg, Germany*
[f]*Department of Medical Biometry and Medical Informatics, University of Freiburg, Germany*
[g]*Department of Psychiatry and Psychotherapy, Section of Gerontopsychiatry and Neuropsychology, University Medical Centre Freiburg, Germany*

Handling Associate Editor: Stefan Teipel

Accepted 17 April 2014

AUTHOR COPY

Abstract. Although episodic memory impairment is usually the earliest sign of Alzheimer's disease (AD), there are up to 15% of patients presenting with early impairment in non-memory cognitive functions (i.e., atypical AD). Stratifying patients with AD may aid clinical trials. Previous studies divided patients by cognitive profile, focusing on cross-sectional analyses without testing stability of clusters over time. We used principal component analysis followed by cluster analyses in 127 patients with AD based on 24 cognitive scores at 0, 6, 12, and 24 months follow-up. We investigated the definition of clusters and their stability over time as well as interactions of cluster assignment and disease severity. At each time point, six distinct factors and four distinct clusters were extracted that did not differ substantially between time points. Clusters were defined by cognitive profile rather than disease severity. 85% of patients fell into the same cluster twice, 42% three times, and 17% four times. Subjects with focal semantic impairment progressed significantly faster than the other cluster. Longitudinally, focal deficits increased relatively rather than tending toward average disease severity. The observed similar cluster definitions at each time point indicate the validity of the approach. Cluster-specific longitudinal increases of focal impairments and significant between-cluster differences in disease progression make this approach useful for stratified inclusion into clinical trials.

Keywords: Alzheimer's disease, atypical AD, cluster analysis, clustering, subgroups, subtypes, typical AD

INTRODUCTION

Alzheimer's disease (AD) usually presents with early episodic memory impairment, followed by deficits in attention, visuospatial abilities, and language [1, 2]. Although episodic memory deficits are typically considered the earliest symptoms of AD [3, 4], it is suggested that up to 15% of patients with AD in dementia centers present with prominent 'atypical'

[1]Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in 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

*Correspondence to: Jessica Peter, Freiburg Brain Imaging, Department of Neurology, University Medical Centre Freiburg, Breisacherstraße 64, 79106 Freiburg, Germany. Tel./Fax: +49 761 270 54797/54160; E-mail: jessica.peter@uniklinik-freiburg.de.

ISSN 1387-2877/14/$27.50 © 2014 – IOS Press and the authors. All rights reserved

cognitive features [5, 6]. One of these focal presentations is described as 'posterior cortical atrophy' [6] with patients tending to be younger than those with typical AD [7]. Other subtypes include 'frontal' [1, 8] and 'logopenic' [9] variants, characterized by distinct clinical and pathological features. Further subtypes have been identified based on the relative density of neurofibrillary tangles in the hippocampus and association cortices: typical AD, hippocampal-sparing AD, and limbic-predominant AD which may have a corresponding clinical phenotype [10, 11].

There are a number of studies aiming to characterize the heterogeneity in the cognitive presentation of AD, using cluster or principal component analysis (PCA) alone or in combination. One study using PCA reported a single common factor (referred to as dementia severity) to explain 60% variability of the cognitive profile [5]. Stopford and colleagues [12] identified 13 different clusters with 4 of them representing focal impairments in spatial perception, executive functions, praxis, or language abilities. Using latent class analysis, Davidson et al. [13] also found 4 distinct clusters. Two of them were labelled 'mild' or 'severe' and could therefore reflect a single common factor 'disease severity' at different levels. The other two clusters showed distinct patterns of predominant impairment in attention or memory, respectively. Methodological differences (i.e., choice of neuropsychological test instruments, the cognitive domains assessed, or the statistical approach used) or a highly time-dependent variability of the identified clusters could explain these discordant results. Indeed, from the existing literature, it is difficult to deduct if clinically defined clusters remain stable over time as most studies focused on a single time point.

When looking at cluster stability it is important to define stability. If clusters reflect cognitive domains most severely affected, patients sharing a similar cognitive profile would remain together in a cluster over time and cluster characteristics would change with diseases progression (i.e., a similar profile but at an increasingly lower level of cognitive performance). Alternatively, if clusters reflect disease severity, the majority of subjects in each cluster would have a relatively 'typical AD' pattern but at different average levels. Individuals would then change clusters during the disease course. In the first scenario, the absence of stability could mean that patients would show different domain specific impairments over time. The clustering would therefore be relatively arbitrary. If clusters represent disease severity, a high instability would occur if cognitive deterioration is very heterogeneous with some patients being able to maintain performance and others showing relatively rapid decline. Of note, changes observed over time are likely to be continuous while clustering is categorical and small changes in cluster definitions could alter the grouping of many subjects. However, should clusters reflect a specific cognitive domain or should clusters differ by rate of cognitive decline, they may aid a stratified inclusion into clinical trials and help to inform patients and relatives more specifically about the individual course of the disease.

Using a sample of AD cases who received cognitive testing at baseline and after 6, 12, and 24 months, we performed cluster analyses separately for each time point and characterized the resulting clusters. In complementary analyses, we used the cluster definition from baseline to describe how subjects would group in these clusters with their follow-up cognitive scores. We aimed to identify two types of stability:

1) *Clusters reflect domains most severely affected:* By repeating the cluster analyses for each time point, we sought to identify if the same profiles of impairment could be observed. In addition, we investigated the longitudinal progression of cluster-defining impairment when projecting follow-up data using baseline criteria.
2) *Clusters reflect disease severity:* If true, clusters should differ in well-established markers of global disease severity, such as the Mini Mental State Examination (MMSE) [14]. We also tested if clustering is predictive of disease progression.

METHODS

Participants

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (http://adni.loni.usc.edu). The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations, as a $60 million, 5-year public-private partnership. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cogni-

AUTHOR COPY

tive impairment (MCI) and early AD. Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials. The Principal Investigator of this initiative is Michael W. Weiner, MD, VA Medical Center and University of California–San Francisco. ADNI is the result of efforts of many co-investigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the U.S. and Canada. The initial goal of ADNI was to recruit 800 subjects but ADNI has been followed by ADNI-GO and ADNI-2. The ADNI protocol was approved by the human studies committees at all institutions in the United States and Canada. Written and verbal informed consents were obtained from participants at screening and enrollment. To date these three protocols have recruited over 1500 adults, ages 55 to 90, to participate in the research, consisting of cognitively normal older individuals, people with early or late MCI, and people with early AD. The follow up duration of each group is specified in the protocols for ADNI-1, ADNI-2, and ADNI-GO. Subjects originally recruited for ADNI-1 and ADNI-GO had the option to be followed in ADNI-2. For up-to-date information, see http://www.adni-info.org.

A total of 127 individuals (aged 76.5 ± 7.2; 51% male) with probable AD (according to NINCDS/ADRDA criteria) were included in our study (see Supplementary Table 1). Data of 0, 6, 12, and 24 months follow-up were used for statistical analysis. AD patients were not assessed at 18 months. Patients with complete data from at least three time-points were included in the analysis.

Neuropsychological battery

All available ADNI neuropsychological scores were included in the analysis, except those scores which were used for AD diagnosis (i.e., logical memory, MMSE) to avoid circularity. Trail Making Test A and B were excluded because of almost 20% missing values. Thus, the following neuropsychological sub-scores were included in the analysis: Clock Drawing Test [15]: drawing (total score) and copying (total score); Rey Auditory Verbal Learning Test [16]: immediate recall (recalled words within trial 1 to 5), intrusions (recalled words that were not part of the list), immediate recall after interference (trial 6), delayed recall after 30 minutes (trial 7), recognition (hits and false alarms); Digit Span from the Wechsler Memory Scale [17]: forward, backwards (number); Category Fluency [18]: animals, vegetables (count per minute), a reduced version (30 items) of the Boston Naming Test [19]: spontaneously named items; semantic cues given and phonemic cues given (maximum 30). As the Digit Span and the Category Fluency Vegetables were not part of the ADNI 2 study (see http://www.adni-info.org/scientists/ADNIStudyProcedures.aspx) only data of ADNI 1 were used in our study. Of note, scores of the MMSE were used as markers of disease severity but did not enter the factor analyses.

Data analyses

Time point specific analyses

The analysis of the neuropsychological data was conducted in a two-stage sequence following Stopford and colleagues [12]. First, a PCA was carried out to represent items that are related to one another by a more general term, following a cluster analysis with the extracted factors using SPSS software package (version 21.0; IBM Inc.; USA). In the following, each step is explained in greater detail.

Principal component analysis: A PCA with orthogonal rotation (varimax) was done using non-normalized raw data of all subjects. To verify sample adequacy Kaiser-Meyer-Olkin (KMO) criterion and Bartlett's test of sphericity were computed. Factor solutions were extracted using both the Kaiser-Guttmann criterion [20] and the scree-test [21].

Cluster analysis: The extracted factor scores were entered into a hierarchical cluster analysis using Ward's method [22] with squared Euclidian distance as a measure of proximity. After determining the optimum number of clusters using an agglomeration schedule (i.e., the change in agglomeration coefficients as the number of clusters increase), cluster analysis was rerun to assign each patient to the corresponding cluster.

Cluster membership and disease severity

We reasoned that high between-cluster differences (i.e., high effect sizes) unexplained by markers of global disease severity (i.e., MMSE), would be a strong indicator that clusters reflect cognitive domains most severely affected. Significance tests were done using a critical $p < 0.05$ after applying Bonferroni's correction for multiple comparisons.

Within time-point analyses: Factor scores entered a Mann-Whitney-U Test to compare one cluster against a combination of the remaining clusters (e.g., cluster 'intrusions' versus all other clusters together).

AUTHOR COPY

Significant results were qualified by effect sizes computed as $r = z - score/\sqrt{N}$, with N being the total sample size [23]. Effect sizes were evaluated according to Cohen [24] with $r \geq 0.1$; $r \geq 0.3$; and $r \geq 0.5$ as being small, medium, and large, respectively.

At each time point, MMSE scores of clusters were compared using Kruskal-Wallis Tests followed by *post-hoc* Mann-Whitney U-Tests.

Between time-point analyses: Using factor scores and cluster assignment defined at baseline, we examined the longitudinal development of the cognitive profile. Therefore, baseline factor loadings were multiplied with z-transformed neuropsychological follow-up scores (i.e., z-transformation of all subjects of 6-months, 12-months, or 24-months scores) to identify the longitudinal progression of patients in each cluster for any given factor by projecting follow-up data on baseline data. Z-transformation, done implicitly through PCA and explicitly for the follow-up data, allowed us to correct for global effects of disease progression. A higher z-score of change in cognitive score at follow-up thus indicates that the corresponding factor is affected more than average.

Longitudinal progression of MMSE scores were computed using difference scores (baseline minus 12-months follow-up or 24-months follow-up, respectively) and Kruskal-Wallis Tests followed by Mann-Whitney U-tests comparing MMSE difference scores of one cluster against another.

Stability of cluster assignment

Using the identical statistical pipeline as at baseline: We examined what percentage of subjects would fall into the same cluster at follow-up using the identical statistical pipeline as at baseline (i.e., PCA followed by hierarchical clustering). Cohen's kappa [25] was computed to test the consistency of cluster assignment given different group sizes. Cohen's kappa was evaluated with $\kappa \leq 0.2$; $\kappa \leq 0.4$; $\kappa \leq 0.6$; $\kappa \leq 0.8$; and $\kappa \geq 0.8$ as being a slight, fair, moderate, substantial, or almost perfect agreement, respectively. If clusters separate disease stages, we would expect that increasingly many subjects would fall into the cluster for the more advanced disease stages.

Using a projection of follow-up on baseline data: Since minor fluctuations could affect clustering, the longitudinal stability of patients' global cognitive profiles was evaluated using projected factor scores of each patient. Therefore, baseline factor loadings were multiplied with z-transformed neuropsychological scores of a given time-point to compute factor scores.

RESULTS

Time-point specific PCA and clustering

Principal component analysis

A PCA with varimax rotation was conducted on a data matrix of 127 subjects and 24 neuropsychological scores.

At all time-points, the KMO measure verified sampling adequacy (rated as 'middling' to 'meritorious' according to Kaiser [26]). Likewise, Bartlett's test of sphericity indicated that correlations between variables were sufficiently high. At each time point, six relevant components were identified (see Supplementary Table 2).

At the first three time points, the item loadings were interpreted as follows: component 1: verbal episodic memory; component 2: intrusions; component 3: semantic knowledge; component 4: working memory; component 5: visuo-constructive abilities, and component 6: recognition. At 24-months follow-up, representations were similar with the exception that visuo-construction and working memory loaded on to the same factor while immediate recall and delayed recall formed two separate ones (see Supplementary Tables 3–6 for factor loadings).

Cluster analysis

As with factor loading, clustering was very similar for the first three time points with agglomeration schedules indicating four distinct clusters ('typical AD', 'focal semantic impairment', 'preserved memory with focal visuo-constructive impairment', 'focal intrusions'), while five distinct clusters were identified at 24-months follow-up (additional 'preserved delayed recall').

Mann-Whitney U-Tests were applied using Bonferroni correction reporting significant results at $p < 0.002$. A comparison of the 4 distinct clusters and their corresponding neuropsychological characteristic is presented in Fig. 1.

Cluster membership and disease severity

Within time point analyses

No significant differences could be observed between cluster specific MMSE-scores at either baseline or 6-months follow-up.

At 12-months follow-up MMSE scores differed significantly between clusters ($H(3) = 225.548$, $p < 0.001$). *Post-hoc* Mann-Whitney U-Tests (critical $p < 0.003$) revealed that cluster 'semantic knowledge'

AUTHOR COPY

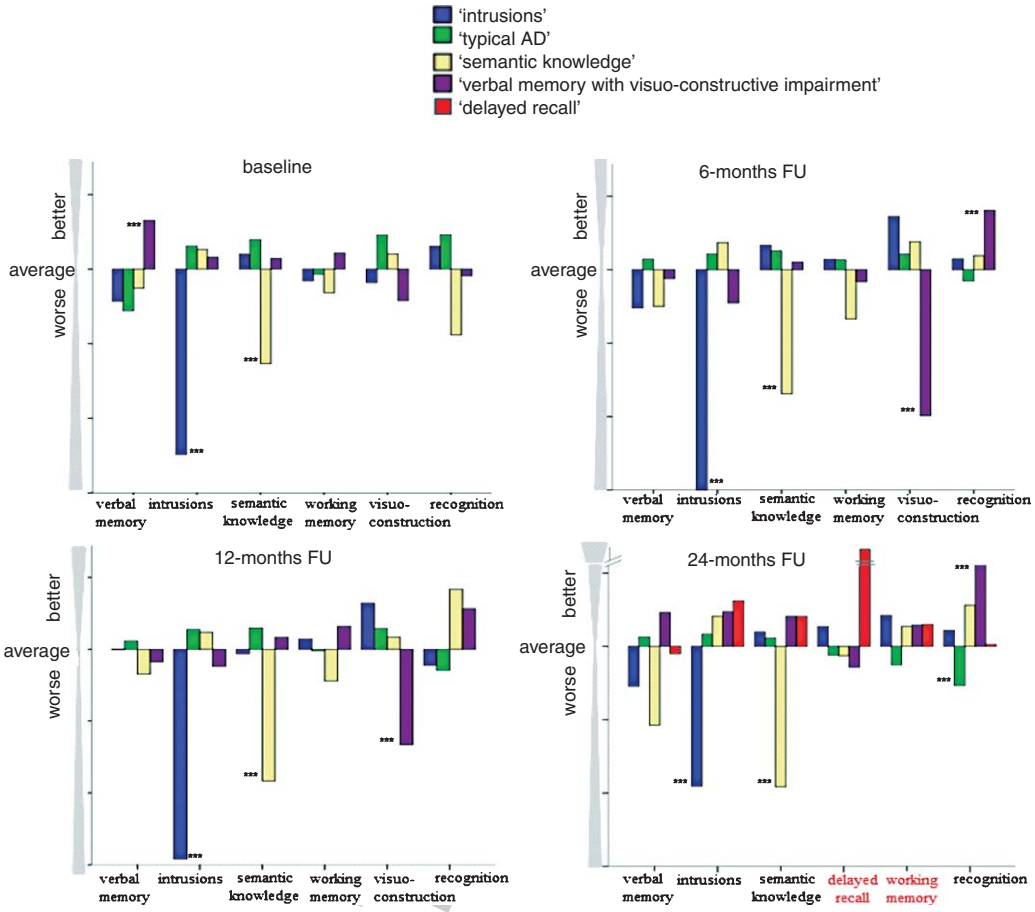

Fig. 1. Progression of cognitive profiles over time (within time-point analyses). At 24-months follow-up two slightly different factors (shown in red) were revealed by principal component analysis. Stars indicate significant different performance in the respective domain (at $p < 0.001$) between the indicated cluster and the combination of all other clusters.

Table 1
MMSE progression for all clusters (within time-point analyses). Mean values with one standard deviation are reported

	Baseline			6-months FU			12-months FU			24-months FU		
cluster	n/f	Age	MMSE	n/f	Age	MMSE	n/f	Age	MMSE	n/f	Age	MMSE
intrusions	11/6	72.1 ± 9.3	23.9 ± 1.3	6/4	64.5 ± 8.5	24.3 ± 1.8	7/5	76.1 ± 6.5	24.1 ± 2.7	12/5	74.8 ± 6.8	21.5 ± 4.1
typ. AD	43/17	76.3 ± 6.3	23.4 ± 2.1	89/42	77.9 ± 6.3	23.4 ± 2.7	82/40	77.0 ± 7.1	22.6 ± 3.4	57/31	77.3 ± 7.5	20.0 ± 4.6
sem.know	21/12	77.9 ± 6.9	23.0 ± 2.2	15/9	75.4 ± 7.1	22.1 ± 2.1	15/7	75.1 ± 8.9	19.2 ± 4.5	9/5	73.4 ± 6.2	13.7 ± 5.6
verb.mem	52/27	77.0 ± 7.5	24.1 ± 1.9	14/7	73.5 ± 6.9	21.8 ± 3.6	23/10	75.7 ± 6.8	20.8 ± 3.9	18/7	77.5 ± 5.9	22.2 ± 3.3
del.recall										2/0	72.5 ± 9.8	24.0 ± 1.4

FU, follow up; MMSE, Mini-Mental Status Examination; n, number of subjects; f, number of female subjects; SD, standard deviation; AD, Alzheimer's disease; typ.AD, typical AD; sem.know, semantic knowledge; verb.mem, preserved verbal memory with visuo-constructive impairment; del. recall, preserved delayed recall.

showed significantly lower MMSE scores compared to cluster 'typical AD' ($p < 0.002$). At 24-months follow-up MMSE scores differed significantly between clusters (H (3) = 11.482, $p < 0.001$). *Post-hoc* tests again revealed that cluster 'semantic knowledge' showed significantly lower MMSE scores compared to cluster 'typical AD' ($p = 0.002$) and cluster 'preserved memory with visuo-constructive impairment' ($p < 0.001$) (see Table 1 and Fig. 2).

Between time point analyses

We observed a significant main effect (critical $p < 0.025$) of cluster specific MMSE progression after 12 months (H (3) = 18.1, $p < 0.001$) and 24 months

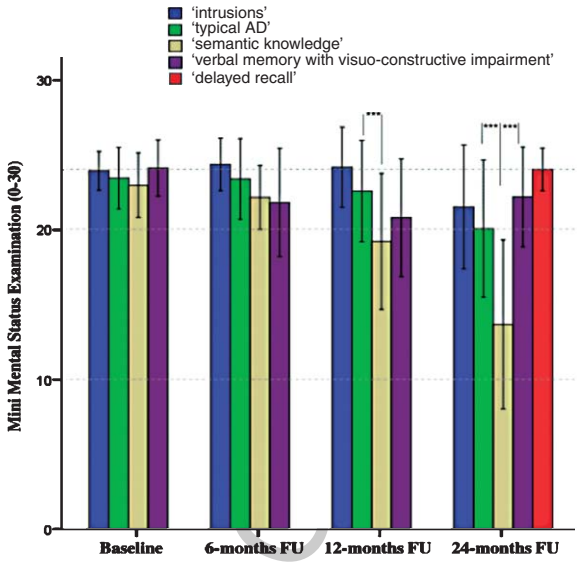

Fig. 2. Scores of Mini-Mental State Examination (MMSE) for clusters within and across time points. Stars indicate significant differences at $p < 0.001$. Dotted lines indicate MMSE cut-off scores for mild, moderate and severe stage of AD. Error bars indicate 1 standard deviation from the mean.

Table 2
MMSE progression of baseline cluster members (between time-point analyses). Mean values with one standard deviation are reported

	Baseline ($n = 127$, 51% male)		6-months FU ($n = 124$, 50% male)	12-months FU ($n = 127$, 51% male)	24-months FU ($n = 98$, 51% male)
	Age	MMSE	MMSE	MMSE	MMSE
all subjects	76.5 ± 7.2	23.7 ± 2.0	22.9 ± 2.9	21.9 ± 3.8	20.1 ± 4.8
'intrusions'	72.1 ± 9.3	23.9 ± 1.3	23.7 ± 3.6	19.8 ± 5.7	18.9 ± 8.2
'typical AD'	76.3 ± 6.3	23.4 ± 2.1	22.8 ± 2.9	22.8 ± 3.4	20.7 ± 4.6
'semantic knowledge'	77.9 ± 6.9	23.0 ± 2.2	21.7 ± 2.4	18.6 ± 3.6	16.6 ± 4.7
'verbal memory'	77.0 ± 7.5	24.1 ± 1.9	23.5 ± 2.8	22.9 ± 2.6	21.2 ± 3.7

FU, follow up; MMSE, Mini-Mental Status Examination; n, number of subjects; SD, standard deviation; AD, Alzheimer's disease.

(H (3) = 10.1, $p < 0.01$). *Post-hoc* Mann-Whitney U-tests (critical $p < 0.004$) revealed that significance was due to cluster 'semantic knowledge' showing a significantly steeper progression compared to cluster 'typical AD' and cluster 'verbal memory with visuo-constructive impairment' after 12 months ($p < 0.001$) and again after 24 months ($p = 0.003$) (see MMSE scores in Table 2).

Stability of cluster assignment

Performing PCA and clustering separately for each time point

Figure 3 provides an overview of stability of cluster assignment over time where we observed the following typical pathways: No typical pathways for patients in the cluster 'intrusions'. In contrast, the majority of cluster 'typical AD' (77%; 33/43) were assigned to the same cluster after both six months and 12 months (65%; 28/43). 17 patients (40%; 17/43) were assigned to the cluster at all 4 time-points. In clusters 'semantic knowledge' (57% two times, 38% three times, 24% four times) and 'verbal memory with visuo-constructive impairment' (12% two times, 6% three times, 4% four times), high proportions of subjects either remained in the initial cluster of changed to 'typical AD' (semantic knowledge: 14%; 'verbal memory with visuo-constructive impairment': 40%).

Comparing cluster assignment for each time point to the next was significantly different from change. Cohen's kappa indicated only 'slight' to 'fair' agreement except when comparing 6-months and 12-months follow up ($p < 0.001$; $\kappa = 0.46$, 'moderate' agreement).

Using a projection on baseline data

Longitudinal stability of patients' global cognitive profiles was evaluated using their individual factor scores with baseline factor loadings. Progression of

AUTHOR COPY

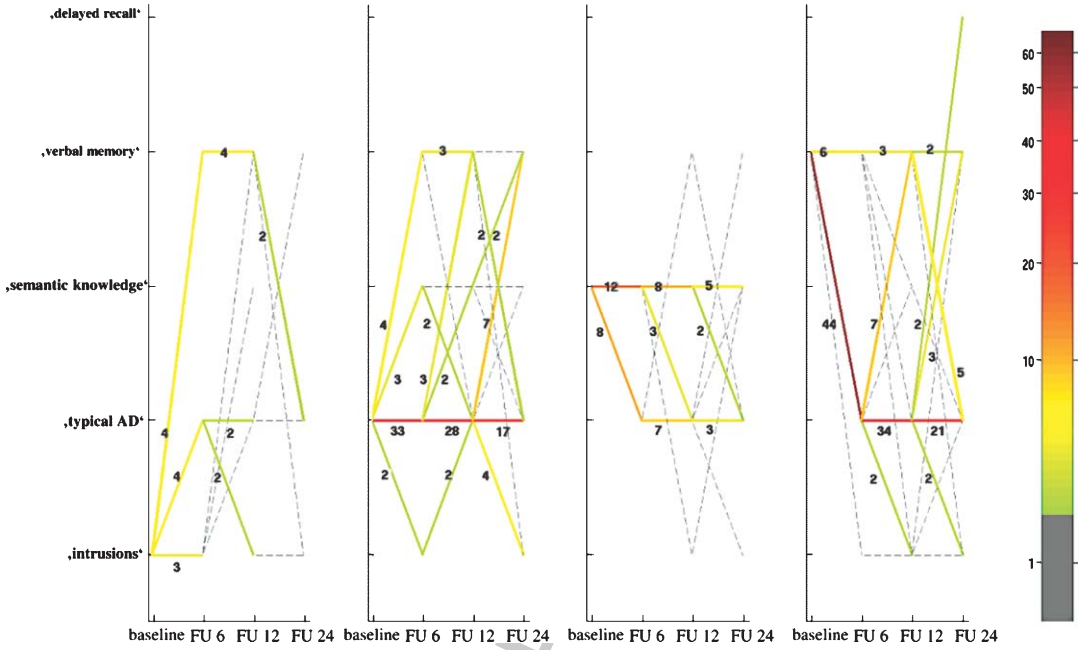

Fig. 3. Cluster change over time based on clustering at every time point for each cluster (within time-point analyses). Dashed lines indicate cluster change of one patient. Warmer colors indicate more patients with identical cluster change.

cognitive profiles is shown in Fig. 4. Subjects in most clusters kept a similar profile but the domain most affected at baseline often deteriorated faster than the others. Subjects in the cluster 'semantic knowledge' continued to be particularly impaired in that domain but several other domains declined rapidly making this the overall fastest declining cluster.

DISCUSSION

The aim of our study was to identify subgroups of AD and to investigate their stability over time.

Baseline factor analyses

In line with previous single time-point studies, PCA revealed memory, visuospatial abilities, and language [27–29] as separate cognitive components in AD. In contrast to previous studies, memory recall and recognition did not load on to the same factor, suggesting two separate aspects of memory. This notion is supported by studies hypothesizing different stages of the verbal learning test to map onto dissociable brain regions (e.g., memory recall is related to hippocampus whereas recognition is associated with perirhinal/entorhinal cortex) [30]. Number of intrusions (recalled items which were not part of the list) loaded on to a distinct factor. Interestingly, previous studies didn't include intrusions in the analysis of verbal episodic memory [e.g., 5, 12] although they reflect distortions of existing memories possibly due to source-monitoring deficits in AD [31]. To our knowledge, this is the first study revealing a distinct factor of 'intrusions' with PCA in AD.

Cluster analyses

Cluster analyses revealed 4 distinct clusters at baseline, 6-month, and 12-months follow-up, while 5 distinct clusters were found at 24-months follow-up. Of note, 29 out of 127 cases dropped out of the study, possibly explaining slight differences in cluster definition at the last time-point.

Compared to the 'typical AD' cluster ($n = 43$), 66% of all patients ($n = 84$) were distinctively impaired in at least one cognitive domain (i.e., one specific factor) at baseline (6-months follow-up: 28%; 12-months follow-up: 35%; 24-months follow-up: 42%). These findings are consistent with former studies showing AD patients with relative isolated word-finding and other language deficits [3, 30], or impairments in visuospatial abilities [32], which might reflect pronounced neuronal loss in the left or right hemispheres respectively.

AUTHOR

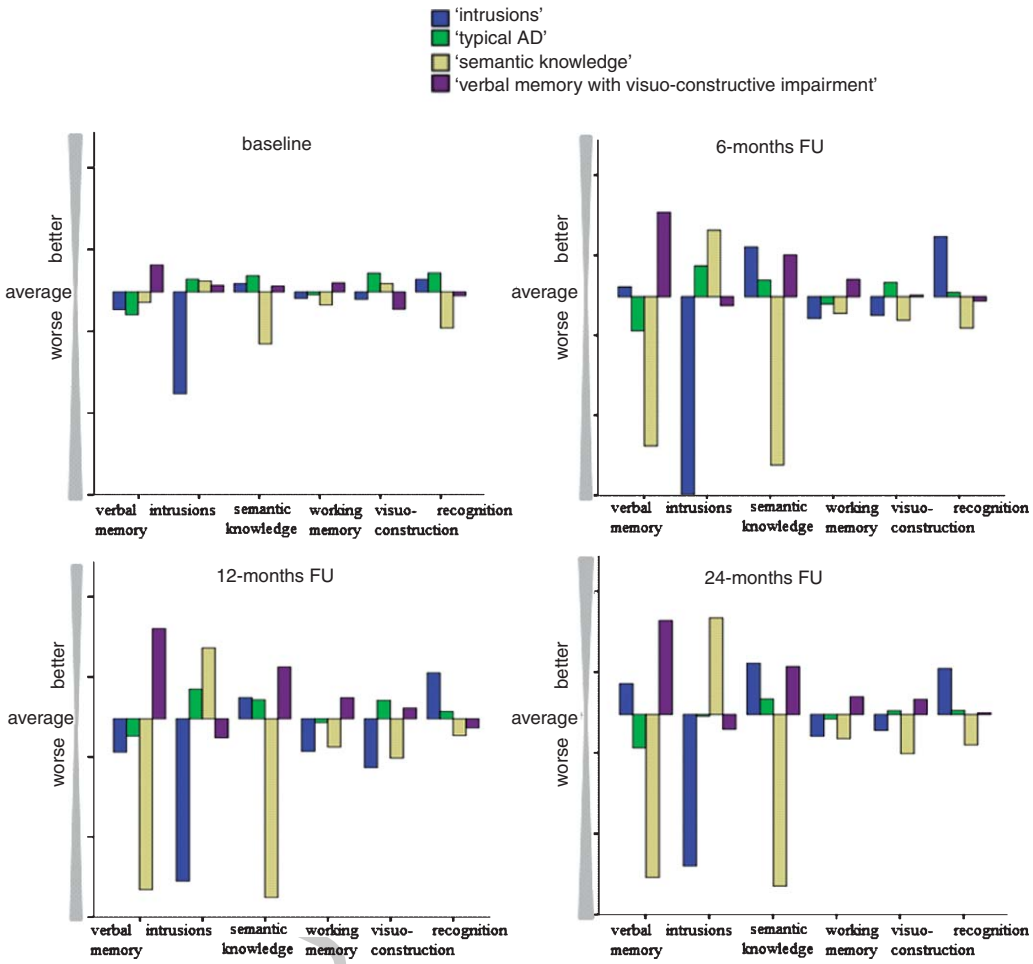

Fig. 4. Progression of cognitive profiles over time using projections of baseline factor loadings (between time-point analyses).

We repeatedly observed a cluster with pronounced difficulties in semantic knowledge as did Martin and colleagues [29] and Fisher and colleagues [33]. Semantic memory deficits in AD may be caused by an impaired access to vocabulary [38] or reflect a genuine deterioration in the structure and organization of semantic knowledge [39]. Although Stopford et al. [12] found a group of patients impaired in a factor termed 'language', their definition was substantially broader and is thus not directly comparable to our cluster 'semantic knowledge'. They suggested that naming deficits are attributable to impaired retrieval rather than semantic knowledge since naming and category fluency both loaded on to different factors although both have significant semantic demands. Our study indicates a different view, since category fluency and naming loaded on to one factor, namely 'semantic knowledge'. This has been corroborated by studies investigating semantic deficits in AD [34] and by distinctive FDG-PET and MRI findings [35]. Some authors argued that the episodic memory deficit in AD is portrayed by poor performance on both recall and recognition measures indicative of impairment in both encoding and storing information [36, 37]. This is not supported by our study as we found different clusters with unequal impairment on both factors (i.e., patients of cluster 'semantic knowledge' were impaired in verbal memory but well preserved in recognition).

Another cluster which was found at all time-points represented preserved verbal memory functions with visuo-constructive impairment (characterized by the clock drawing test). A similar cluster was also found by Fisher et al. [33]. Of note, we observed a subgroup characterized by frequent intrusions at all time-points. Intrusions may be related to memory distortion and seem to indicate a breakdown in the ability to monitor and constrain recall responses [38]. Former studies never included intrusions in their analysis (neither PCA

nor cluster analysis) although intrusions are typically assessed in every verbal memory tests.

Comparing the pattern of deficits in the four distinct clusters to clinically defined variants of AD indicates that patients in the cluster 'verbal memory with visuo-constructive abilities' might be comparable to the variant 'posterior cortical atrophy', showing preserved verbal memory and recognition with deficits in visuo-construction. However, this interpretation is made less likely as this cluster contains the most patients at baseline and patients from that cluster tended to fall into the 'average AD' cluster at follow-up (see Fig. 3). Patients in the cluster 'semantic knowledge' showing deficits in semantic knowledge and verbal memory might represent an aphasia subtype of AD but additional aphasia tests would be required to put such an interpretation on firmer grounds.

Of note, frequency of APOE allele (distribution of genotypes APOE 2/3; APOE 2/4; APOE 3/3; APOE 3/4; and APOE 4/4) did not differ significantly between clusters at any time-point as measured using χ^2-test on cross-tables (baseline: $\chi^2 = 12.3$, $p = 0.42$; 6m-FU: $\chi^2 = 12.9$, $p = 0.37$; 12m-FU $\chi^2 = 13.2$, $p = 0.32$; 24m-FU $\chi^2 = 17.7$, $p = 0.34$). Nevertheless, in the 'focal semantic impairment' cluster (the fastest progressing cluster), the frequency (in percent) of APOE 4/4 genotype was the highest at all time-points, although not significantly so.

Longitudinal stability of cluster assignment

Cluster definitions remained relatively constant over time with time point specific PCA and cluster analyses. A significant amount of patients were assigned to the same cluster at follow-up. 53 patients (42%) were assigned to the same cluster at three time points subsequently and 21 patients (17%) were assigned to the same cluster at every time-point. In general, stability of cluster assignment was significant but to a slight extent. The highest stability was observed in the cluster 'typical AD' followed by the cluster 'semantic knowledge'. Figure 3 indicates that subjects in clusters 'semantic knowledge' and 'verbal memory' frequently changed to the 'typical AD' cluster. Figure 4 on the other hand, indicates more pronounced impairments in the domains that were already predominantly affected at baseline, even when correcting for the overall disease progression by using z-scores. This apparent contradiction could result if those patients remaining in the cluster would continue to develop a more pronounced deficit and would dominate the cluster average.

Effects of disease severity on clustering

Using MMSE as a marker of disease severity, within time-point differences in MMSE were insignificant for baseline and 6 months follow-up. This observations contradicts the argument, that diseases severity is the primary distinction between clusters [5]. Of note, Ralph et al. [5] included MMSE in the factor analysis, which could lead to circularity when MMSE is used to rate disease severity. While we retrospectively compared MMSE between clusters, one could also regress out disease severity from the raw data before PCA and clustering. This would allow observing a clustering independently of disease severity. Ideally, non-cognitive markers (e.g., hippocampus atrophy) should be used for that purpose, as the MMSE is not independent from cognitive scores. Regressing out disease severity using MMSE scores could therefore distort the relation between cognitive scores and alter the results substantially.

Comparing cognitive decline over time averaged across all participants of this study (i.e., about 2 points per year on the MMSE), patients assigned to the cluster 'semantic knowledge' at baseline showed a significantly steeper decline of about 4 points per year (see Table 2). Lower MMSE scores in subjects with 'focal language' deficits have also been reported by Stopford et al. [12] who explained this by higher demands of the MMSE test on language compared to, for example, visuo-construction. This is not supported by former studies, identifying two factors using PCA with the 'verbal abilities' factor explaining only 20% of total variance and not being related to cognitive deterioration [39]. Of note, when repeating these analyses using Clinical Dementia Rating-Sum of Boxes [40], a measure of cognitive and functional performance without reference to psychometric presentation as an index of disease severity, no significant differences could be observed for any time point but descriptively patients of cluster 'semantic knowledge' increased fastest.

Limitation

The neuropsychological tests available in ADNI are partly unbalanced with some cognitive domains represented with multiple scores and others with one or even not at all. For instance, genuine tests of executive functions are underrepresented. Future studies should therefore include a wider test set.

In summary, the observed similar cluster definitions at each time point support the stability of the clusters.

AUTHOR COPY

Cluster specific longitudinal increases of focal impairments and significant between cluster differences in disease progression make this approach useful for stratified inclusions into clinical trials.

ACKNOWLEDGMENTS

AA and SK were funded by a grant from the Deutsche Forschungsgemeinschaft (KL2415/2-1) to SK.

Authors' disclosures available online (http://www.j-alz.com/disclosures/view.php?id=2279).

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of California, Los Angeles. This research was also supported by NIH grants P30 AG010129 and K01 AG030514.

SUPPLEMENTARY MATERIAL

Supplementary tables are available in the electronic version of this article: http://dx.doi.org/10.3233/JAD-140261.

AUTHOR COPY

REFERENCES

[1] Alladi S, Xuereb J, Bak T, Nestor P, Knibb J, Patterson K, Hodges JR (2007) Focal cortical presentations of Alzheimer's disease. *Brain* **130**, 2636-2645.

[2] Petersen RC (1998) Clinical subtypes of Alzheimer's disease. *Dement Geriatr Cogn Disord* **9**(Suppl 3), 16-24.

[3] Galton CJ, Patterson K, Xuereb JH, Hodges JR (2000) Atypical and typical presentations of Alzheimer's disease: A clinical, neuropsychological, neuroimaging and pathological study of 13 cases. *Brain* **123 Pt 3**, 484-498.

[4] McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM (1984) Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* **34**, 939-944.

[5] Ralph MAL, Patterson K, Graham N, Dawson K, Hodges JR (2003) Homogeneity and heterogeneity in mild cognitive impairment and Alzheimer's disease: A cross-sectional and longitudinal study of 55 cases. *Brain* **126**, 2350-2362.

[6] Schmidtke K, Hüll M, Talazko J (2005) Posterior cortical atrophy: Variant of Alzheimer's disease? A case series with PET findings. *J Neurol* **252**, 27-35.

[7] Duker AP, Espay AJ, Wszolek ZK, Rademakers R, Dickson DW, Kelley BJ (2012) Atypical motor and behavioral presentations of Alzheimer disease: A case-based approach. *Neurologist* **18**, 266-272.

[8] Schott JMRB (2006) Apolipoprotein E genotype modifies the phenotype of alzheimer disease. *Arch Neurol* **63**, 155-156.

[9] Gorno-Tempini ML, Brambati SM, Ginex V, Ogar J, Dronkers NF, Marcone A, Perani D, Garibotto V, Cappa SF, Miller BL (2008) The logopenic/phonological variant of primary progressive aphasia. *Neurology* **71**, 1227-1234.

[10] Janocko NJ, Brodersen KA, Soto-Ortolaza AI, Ross OA, Liesinger AM, Duara R, Graff-Radford NR, Dickson DW, Murray ME (2012) Neuropathologically defined subtypes of Alzheimer's disease differ significantly from neurofibrillary tangle-predominant dementia. *Acta Neuropathol* **124**, 681-692.

[11] Murray ME, Dickson D (2008) O1-01-02: Alzheimer's disease with relative hippocampal sparing: A distinct clinicopathologic variant. *Alzheimers Dement* **4**, T106-T106.

[12] Stopford CL, Snowden JS, Thompson JC, Neary D (2008) Variability in cognitive presentation of Alzheimer's disease. *Cortex* **44**, 185-195.

[13] Davidson JE, Irizarry MC, Bray BC, Wetten S, Galwey N, Gibson R, Borrie M, Delisle R, Feldman HH, Hsiung G-Y, Fornazzari L, Gauthier S, Guzman D, Loy-English I, Keren R, Kertesz A, George-Hyslop PS, Wherrett J, Monsch AU (2010) An exploration of cognitive subgroups in Alzheimer's disease. *J Int Neuropsychol Soc* **16**, 233-243.

[14] Folstein MF, Folstein SE, McHugh PR (1975) "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* **12**, 189-198.

[15] Shulman KI, Shedletsky R, Silver IL (1986) The challenge of time: Clock-drawing and cognitive function in the elderly. *Int J Geriatr Psychiatry* **1**, 135-140.

[16] Rey A (1964) *L'examen clinique en psychologie*. Presses universitaires de France, Paris.

[17] Wechsler D (1945) A standardized memory scale for clinical use. *J Psychol* **19**, 87-95.

[18] Benton AL (1968) Differential behavioral effects in frontal lobe disease. *Neuropsychologia* **6**, 53-60.

[19] Kaplan E, Goodglass H, Weintraub S (1983) *Boston Naming Test*. Lea & Febiger, Philadelphia.
[20] Guttman L (1954) Some necessary conditions for common-factor analysis. *Psychometrika* **19**, 149-161.
[21] Cattell RB (1966) The Scree test for the number of factors. *Multivariate Behav Res* **1**, 245-276.
[22] Ward JH (1963) Hierarchical grouping to optimize an objective function. *J Am Stat Assoc* **58**, 236-244.
[23] Rosnow RL, Rosenthal R (1996) Computing contrasts, effect sizes, and counternulls on other people's published data: General procedures for research consumers. *Psychol Methods* **1**, 331-340.
[24] Cohen J (1988) *Statistical power analysis for the behavioral sciences.* L. Erlbaum Associates, Hillsdale, NJ.
[25] Cohen J (1960) A coefficient of agreement for nominal scales. *Educ Psychol Meas* **20**, 37-46.
[26] Kaiser HF (1970) A second generation little jiffy. *Psychometrika* **35**, 401-415.
[27] Kramer JH, Jurik J, Sha SJ, Rankin KP, Rosen HJ, Johnson JK, Miller BL (2003) Distinctive neuropsychological patterns in frontotemporal dementia, semantic dementia, and Alzheimer disease. *Cogn Behav Neurol* **16**, 211-218.
[28] Price BH, Gurvit H, Weintraub S, Geula C, Leimkuhler E, Mesulam M (1993) Neuropsychological patterns and language deficits in 20 consecutive cases of autopsy-confirmed Alzheimer's disease. *Arch Neurol* **50**, 931-937.
[29] Martin A, Brouwers P, Lalonde F, Cox C, Teleska P, Fedio P, Foster NL, Chase TN (1986) Towards a behavioral typology of Alzheimer's patients. *J Clin Exp Neuropsychol* **8**, 594-610.
[30] Wolk DA, Dickerson BC (2011) Fractionating verbal episodic memory in Alzheimer's disease. *NeuroImage* **54**, 1530-1539.
[31] Gallo DA, Sullivan AL, Daffner KR, Schacter DL, Budson AE (2004) Associative recognition in Alzheimer's disease: Evidence for impaired recall-to-reject. *Neuropsychology* **18**, 556-563.
[32] Strite D, Massman PJ, Cooke N, Doody RS (1997) Neuropsychological asymmetry in Alzheimer's disease: Verbal versus visuoconstructional deficits across stages of dementia. *J Int Neuropsychol Soc* **3**, 420-427.
[33] Fisher NJ, Rourke BP, Bieliauskas LA (1999) Neuropsychological subgroups of patients with Alzheimer's disease: An examination of the first 10 years of CERAD data. *J Clin Exp Neuropsychol* **21**, 488-518.
[34] Zahn R, Garrard P, Talazko J, Gondan M, Bubrowski P, Juengling F, Slawik H, Dykierek P, Koester B, Hull M (2006) Patterns of regional brain hypometabolism associated with knowledge of semantic features and categories in Alzheimer's disease. *J Cogn Neurosci* **18**, 2138-2151.
[35] Frings L, Klöppel S, Teipel S, Peters O, Frölich L, Pantel J, Schröder J, Gertz H-J, Arlt S, Heuser I, Kornhuber J, Wiltfang J, Maier W, Jessen F, Hampel H, Hüll M (2011) Left anterior temporal lobe sustains naming in Alzheimer's dementia and mild cognitive impairment. *Curr Alzheimer Res* **8**, 893-901.
[36] Greenaway MC, Lacritz LH, Binegar D, Weiner MF, Lipton A, Munro Cullum C (2006) Patterns of verbal memory performance in mild cognitive impairment, Alzheimer disease, and normal aging. *Cogn Behav Neurol* **19**, 79-84.
[37] Tierney MC, Black SE, Szalai JP, Snow WG, Fisher RH, Nadon G, Chui HC (2001) Recognition memory and verbal fluency differentiate probable Alzheimer disease from subcortical ischemic vascular dementia. *Arch Neurol* **58**, 1654-1659.
[38] MacDuffie KE, Atkins AS, Flegal KE, Clark CM, Reuter-Lorenz PA (2012) Memory distortion in Alzheimer's disease: Deficient monitoring of short- and long-term memory. *Neuropsychology* **26**, 509-516.
[39] Brugnolo A, Nobili F, Barbieri MP, Dessi B, Ferro A, Girtler N, Palummeri E, Partinico D, Raiteri U, Regesta G, Servetto G, Tanganelli P, Uva V, Mazzei D, Donadio S, De Carli F, Colazzo G, Serrati C, Rodriguez G (2009) The factorial structure of the mini mental state examination (MMSE) in Alzheimer's disease. *Arch Gerontol Geriatr* **49**, 180-185.
[40] Lynch CA, Walsh C, Blanco A, Moran M, Coen RF, Walsh JB, Lawlor BA (2006) The clinical dementia rating sum of box score in mild dementia. *Dement Geriatr Cogn Disord* **21**, 40-43.

AUTHOR COPY