NeuroImage: Clinical 20 (2018) 1222–1232

Contents lists available at ScienceDirect

NeuroImage: Clinical

journal homepage: www.elsevier.com/locate/ynicl

Profiling heterogeneity of Alzheimer's disease using white-matter impairment factors

Xiuchao Sui, Jagath C. Rajapakse*, The Alzheimer's Disease Neuroimaging Initiative[1]

School of Computer Science and Engineering, Nanyang Technological University, 639798, Singapore

ARTICLE INFO

Keywords:
Heterogeneity
Alzheimer's disease
Mental disorder subtypes
White matter impairment
Latent dirichlet allocation

ABSTRACT

The clinical presentation of Alzheimer's disease (AD) is not unitary as heterogeneity exists in the disease's clinical and anatomical characteristics. MRI studies have revealed that heterogeneous gray matter atrophy patterns are associated with specific traits of cognitive decline. Although white matter (WM) impairment also contributes to AD pathology, its heterogeneity remains unclear. The Latent Dirichlet Allocation (LDA) method is a suitable framework to study heterogeneity and allows to identify latent impairment factors of AD instead of simply mapping an overall disease effect. By exploring whole brain WM skeleton images by using LDA, three latent factors were revealed in AD: a temporal-frontal impairment factor (temporal and frontal lobes, especially hippocampus and para-hippocampus), a parietal factor (parietal lobe, especially precuneus), and a long fibre bundle factor (corpus callosum and superior longitudinal fasciculus). As revealed by longitudinal analysis, the latent factors have distinct impact on cognitive decline: for executive function (EF), the temporal-frontal factor was more strongly associated with baseline EF compared with the parietal factor, while the long-fibre bundle factor was most associated with decline rate of EF; for memory, the three factors showed almost equal effect on the baseline memory and decline rate. For each participant, LDA estimates his/her composition profile of latent impairment factors, which indicates disease subtype. We also found that the APOE genotype affects the AD subtype. Specifically, APOE ε4 was more associated with the long fibre bundle factor and APOE ε2 was more associated with temporal-frontal factor. By investigating heterogeneity and subtypes of AD through white matter impairment factors, our study could facilitate precision medicine.

1. Introduction

Alzheimer's disease (AD), the most common form of dementia, affects 11% of individuals over age 65 worldwide with no known cure. Although memory loss is the primary symptom, AD patients manifest heterogeneous cognitive profiles, i.e., each individual could express different extents of decline in memory, executive skills, language, and perceptuospatial abilities (Scheltens et al., 2015; Stopford et al., 2008; Snowden et al., 2007), which complicate the diagnosis. Hence, there is a pressing need to identify AD subtypes for improving diagnosis accuracy and monitor disease progression.

AD is a complex neurodegenerative disease in that pathological and structural heterogeneity is constantly observed in AD (Kramer and Miller, 2000; Dickerson et al., 2010; Johnson et al., 1999; Butters et al., 1996; Tang-Wai et al., 2004; Gefen et al., 2012). For example, neuropathological studies found that 25–30% of the AD cases had untypical distributions of amyloid plaques and neurofibrillary tangles (NFTs), and exhibited different rates of disease progression compared with typical AD cases (Murray et al., 2011; Ossenkoppele et al., 2016; Whitwell et al., 2012). MRI studies reported that 10–30% of the AD patients exhibited untypical hippocampal atrophy (Frisoni et al., 2010; Scheltens et al., 2002; Lowe et al., 2013). Moreover, a recent MRI study suggested that the varying patterns of the whole brain gray matter (GM) atrophy were associated with distinct trajectories of multidomain cognitive decline in AD (Zhang et al., 2016; Noh et al., 2014; Park et al., 2017).

White matter (WM) impairment also contributes to AD pathology (Sachdev et al., 2013; Agosta et al., 2011) and can be detected even before the development of cortical atrophy in at-risk and heathy participants (Agosta et al., 2011; Maier-Hein et al., 2015). Most WM studies on AD assumed a single disease effect that charts only the overall associations between WM changes and cognition decline in AD (Acosta-Cabronero et al., 2009; Bozzali et al., 2002). One study reported that

* Corresponding author at: N4-2a06, 50 Nanyang Avenue, Nanyang Technological University, 639798, Singapore.
E-mail address: asjagath@ntu.edu.sg (J.C. Rajapakse).

[1] Data used in the preparation of this article were obtained from the ADNI database (adni.loni.usc.edu). A complete list of the ADNI investigators can be found at http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

https://doi.org/10.1016/j.nicl.2018.10.026
Received 9 July 2018; Received in revised form 28 September 2018; Accepted 23 October 2018
Available online 29 October 2018

Table 1
Study demographics, cognitive scores, and imaging metrics of baseline scans.

	NC	EMCI	LMCI	AD
Participants at baseline	50	90	44	48
3 months	68%	61%	77%	75%
6 months	74%	46%	66%	65%
1 year	76%	56%	66%	67%
2 years	62%	56%	45%	13%
Sex (female/male)	27/23	36/54	17/27	20/28
Age (years)	72.5(6.2)	72.9(7.9)	73.9(7.1)	75.0(8.7)
Education (years)	16(3)[c]	16(3)	16(3)	15(3)
Memory (z-score)[#]	0.99(0.60)[a,b,c]	0.49(0.64)[d,e]	−0.31(0.65)[f]	−0.92(0.50)
Executive function (z-score)[$]	0.81(0.73)[a,b,c]	0.19(0.74)[e]	0.14(0.77)[f]	−0.77(0.81)
Intracranial vault (ICV) (cm^3)	1380(113)[b]	1406(137)	1443(132)	1395(158)
Total gray matter (%ICV)	42.4(1.2)[b,c]	42.2(1.5)[d,e]	41.4(1.5)[f]	40.6(1.5)
APOE ε3/4 or 4/4	0.31[a,b,c]	0.57[d,e]	0.88	0.82
APOE ε2/3 or 2/2	0.11	0.11	0.07	0.07
CSF $A\beta_{1-42}$ (pg/ml)	203.8(50.3)[a,b,c]	174.3(49.6)[e]	155.1(52.1)	137.3(36.3)

The list of participants ($N = 232$, data-points = 893) is available at github.com/xiuchao/ldaFA.
[#] ADNI-EF is available for 679 time-points, among which 504 have amyloid information available;
[$] ADNI-Mem is available for 691 time-points, among which 508 have amyloid information available;
[a] NC v. EMCI ($p < 0.05$)
[b] NC v. LMCI ($p < 0.05$).
[c] NC v. AD ($p < 0.05$).
[d] EMCI v. LMCI ($p < 0.05$)
[e] EMCI v.AD ($p < 0.05$).
[f] LMCI v. AD ($p < 0.05$).

WM impairment in AD was associated with impairment in different cognitive domains (Gouw et al., 2008). A recent study assessed WM impairment in typical and atypical AD and identified syndrome-specific patterns of WM microstructure breakdown (Caso et al., 2015). Inspired by these findings, we hypothesize that heterogeneity exists in WM impairment in AD, which could explain the variations of multidomain cognitive decline among participants across the disease spectrum.

In this work, we explored the heterogeneity of WM impairment in AD by using Latent Dirichlet Allocation (LDA) to identify multiple latent disease factors from WM skeletons constructed using diffusion tensor imaging (DTI) data (Zhang et al., 2016; Blei et al., 2003). LDA provides a probabilistic description of multiple impairment factors. For example, concerning patient A, the WM impairment pattern might be 70% owing to factor 1 and 30% to factor 2; while concerning patient B, the impairment might be 20% owing to factor 1 and 80% to factor 2. It remains an open question whether subtypes, which manifest heterogeneity in AD cases, are exclusive or mixing up in an individual. LDA explores the mixed effects of multiple latent factors in a probabilistic framework and is therefore well suited for studying heterogeneity in AD.

AD develops slowly from a preclinical phase and the pathophysiological changes start about a decade before the final diagnosis (Villemagne et al., 2013). In this work, we found that heterogeneity of WM impairment already appears before clinical syndromes are fully expressed. We revealed the association between the WM impairment factors and the cognitive decline in participants throughout the disease spectrum. In addition, we verified that the genotype was associated with the latent factors, i.e., APOE ε4 was most strongly associated with the long fibre bundle factor and APOE ε2 was more associated with temporal-frontal factor. Our work provides insights into disease heterogeneity in brain white matter throughout the prolonged course of AD, sheds light on the association between the genotype and the disease subtype, and facilitates precision medicine.

2. Material and methods

2.1. Overview

We performed the following analyses on fractional anisotropy (FA) maps derived from DTI scans, which measure WM integrity. First, the LDA model was used to estimate the degeneration probabilities $p(Voxel \mid Factor)$ of each voxel in each factor, from the AD patients. Afterwards, the estimated latent factors were used to infer the factor compositions $p(Factor \mid Patient)$ of both AD patients and non-demented participants, including participants with early mild cognitive impairment (EMCI), late mild cognitive impairment (LMCI) and normal controls (NC). Second, we investigated whether the factor compositions are stable over 3/6/12/24 months as the disease progresses. Third, using longitudinal analyses, we explored whether different WM impairment factors are associated with memory and executive function decline differently across the prolonged course of AD. Last, we investigated whether the *APOE* genotype is associated with the disease subtype (i.e., latent factor compositions) by using a general linear model (GLM).

2.2. Data

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership and led by Principle Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to examine whether structural MRI, functional MRI, DTI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early AD (up-to-date information is at www.adni-info.org).

Our work was based on ADNI-2 and ADNI-GO (extracted on April 08, 2018), and enrolled all the available participants having both corrected FA images and structural MRI images of good quality. As listed in Table 1, this study included 232 participants at baseline, comprised of 48 AD patients, 90 EMCI, 44 LMCI and 50 NC. These participants were longitudinally followed at 3, 6, 12, 24, 36, and 48 months later (most participants have scans at more than two time points), totalling 893 time points of 232 participants.

The 48 AD patients at baseline were used as input of the LDA model to identify latent WM impairment factors in AD. Based on the derived latent factors, individual factor composition was inferred for every participant. With the factor composition profiles (at 893 time-points) quantified by the LDA framework, we delved into further statistical

analysis to reveal the associations between latent impairment factors and cognition and *APOE* genotype.

For neuropsychological data, memory (ADNI-Mem) (Crane et al., 2012) and executive function (ADNI-EF) (Gibbons et al., 2012) scores were explored. For genotype data, *APOE* ε4 and *APOE* ε2 were considered. 170 participants at baseline have genotype information available, which were used to explore genotype effects on disease subtypes (i.e. latent factor compositions).

2.3. Tract-Based Spatial Statistics (TBSS)

DTI scans help us understand the regional basis of white matter tissue degeneration in clinical conditions. Changes in diffusion properties demonstrate distinct zones of alterations, which potentially stem from the differences of the underlying pathology. FA was considered as the primary metric of interest in this study, which is quantified by the orientational coherence of water diffusion within a voxel. FA reflects the extent of astrogliosis, myelin and axonal loss.

The FA images were analysed using TBSS (fsl.fmrib.ox.ac.uk/fsl/fslwiki/TBSS) that comes as a part of FSL (Smith et al., 2004). First, 232 FA images at baseline scans (Table 1) were aligned into a common space by using nonlinear registration. Second, the mean FA image was generated and thinned in order to create a group's mean FA skeleton, representing the centre of the tracts. Every participant's aligned FA data was then projected onto this skeleton. Third, the $\log_{10}$ transformation was applied to the resulting skeletonized FA data. Fourth, effects of age, sex, and education on log transformed FA values were regressed out with a general linear model (GLM) estimated from 50 NC participants.

For the follow-up scans (647 time-points), the FA images went through similar TBSS procedures without generating a new WM skeleton. Instead we used the WM skeleton derived from the baseline scans as the standard WM template to derive skeletonized FA maps.

2.4. Latent Dirichlet Allocation (LDA)

In this study, our premise is that an individual AD patient expresses one or multiple WM impairment factors that are associated with distinct WM impairment patterns (Fig.1). The LDA method (Blei et al., 2003) is well suited for modelling DTI data under such a premise, which.

As LDA works on discrete data, the continuous FA maps need to be discretized by mapping greater WM impairment to a larger positive integer. Specifically, FA values were first log-transformed to make data better conform to normality, and then z-transformed (with respect to NC baseline scans) for each AD patient at baseline. A z score below zero at a given voxel of a particular participant is an indication of degeneration relative to the NC participants. Z scores above zero suggest unimpaired voxels and were set to zero. The z scores were discretized by multiplying by -10 and rounding to the nearest integer. Larger positive values imply more severe impairment.

For a specified number K of WM impairment factors, LDA analysis was performed on the transformed FA images of AD patients to identify latent factors and infer factor compositions. Specifically, the probabilistic distributions $P(Voxel \mid Factor)$of WM impairment, estimated from AD patients were used to infer factor compositions $P(Factor|Patient)$ of individual participants. The factor composition of a participant can be visualized as a point inside the factor space (e.g., $K = 3$ is represented by a triangle) as shown in Fig.4. The factor composition represents a subtype of AD.

Moreover, the robustness of LDA were also evaluated on two aspects (*SI Appendix)*. First, the stableness of identified latent factors across different initializations was investigated. For each K, the LDA algorithm was run for 20 times with random initializations. We observed that different random initializations led to almost identical solutions, suggesting that the identified latent factors are robust across different runs. The run with the highest data likelihood (indicating the model best fits the data) was selected as the optimal factor decomposition of the current K. Second, the spatial consistency of two sets of latent factors identified by LDA with different K's was investigated. The spatial consistency of latent factors is defined as the situation that latent factors reside in approximately the same brain regions regardless of a particular K value. For K in candidate sets of $\{2, 3, \ldots, 6\}$ (!!! Invalid Citation !!!, n.d.), we observed that the identified latent factors were robust across different K's (*SI Appendix, Fig.S1*).

2.5. Choosing the number K of factors

The number K of latent factors is an important parameter of LDA. To determine K, we proposed to maximize the spatial separation criterion, i.e., the latent factors that best represent heterogeneity are those whose spatial distributions are maximally separated or minimally overlapped. Intuitively, this criterion expects that any two latent impairment factors should affect different brain regions, so that the factors are meaningful, interpretable, and heterogeneous. To measure the spatial separation of latent factors within the K factor model, the correlation coefficients

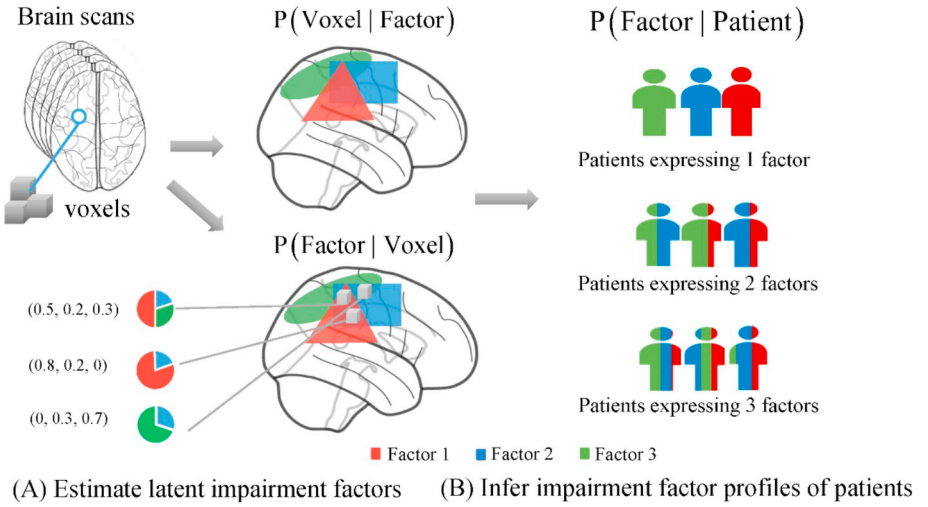

Fig. 1. Schematic illustration of impairment factor decomposition using the Latent Dirichlet Allocation (LDA) framework on brain images. (A) Given a cohort of brain images, LDA estimates group-wise latent impairment factors and each factor is associated with distinct patterns of structural impairment, represented as $p(Voxel \mid Factor)$. In this schematic figure, the red triangle reveals the brain region affected by latent impairment factor 1, the blue rectangle reveals impairment factor 2 and the green ellipse reveals impairment factor 3. Each voxel is associated with one or more latent impairment factors, denoted by $p(Factor \mid Voxel)$, as illustrated as the pie charts. (B) LDA is used to further infer $p(Factor \mid Patient)$, the probability that a patient expresses a particular factor. $p(Factor \mid Patient)$ serves as the profile of the disease subtype. As illustrated, a patient expresses one or multiple latent impairment factors was originally proposed to discover latent topics in a set of text documents. In the brain imaging scenario, each participant's WM image is a collection of voxels, and each voxel is associated with a subset within total K impairment factors; this is analogous to text analysis in that a document is collection of words, and each word is associated with a subset within total K latent topics. Each latent impairment factor is represented by a probability distribution over all WM voxels $p(Voxel \mid Factor)$, which can be visualized as a probabilistic impairment map overlaid on the FSL MNI152 FA skeleton template (Fig.2). Each AD patient expresses multiple WM impairment factors to different extents and $p(Factor \mid Patient)$ denotes the composition of impairment factors expressed by a patient.

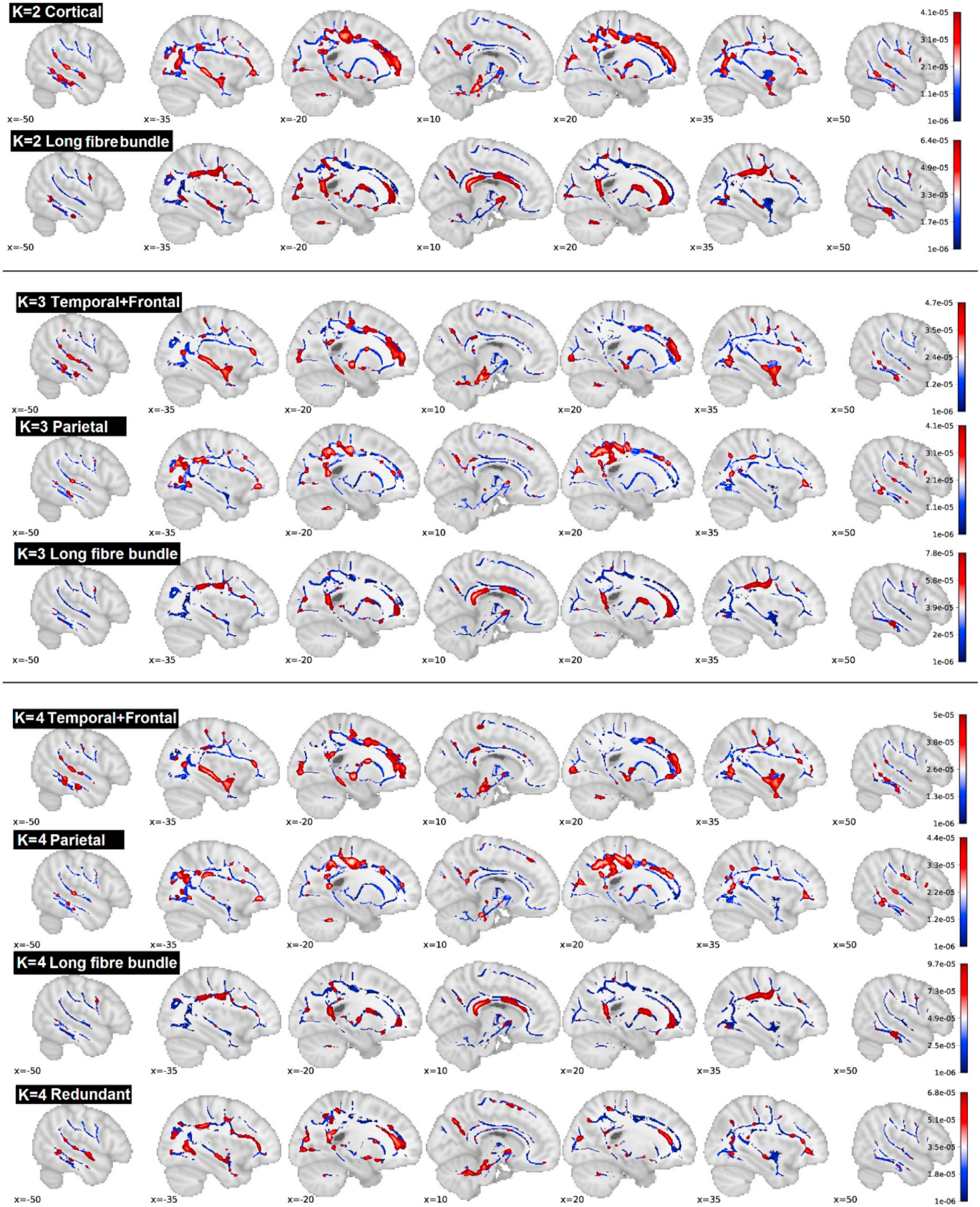

Fig. 2. Latent white matter (WM) impairment factors in AD identified by the LDA model with $K = 3$. The probabilistic impairment distribution $p(Voxel \mid Factor)$ of each factor was quantified and displayed on the WM skeleton. Red colour indicates brain regions that are more susceptible disease targets while blue colour indicates less susceptible regions. Three impairment factors were the temporal-frontal factor, the parietal factor, and the long fibre bundle factor.

Table 2
Comparison of baseline scores and decline rates of (A) executive function and (B) memory across factors.

	(A) ADNI - executive function			(B) ADNI –memory		
	$p - tf$	$l - tf$	$l - p$	$p - tf$	$l - tf$	$l - p$
Baseline score	−0.84 (0.37) **p = 0.02**	- 0.09 (0.37) $p = 0.81$	+0.75 (0.37) $p = 0.052$	- 0.07 (0.36) $p = 0.84$	−0.13 (0.36) $p = 0.73$	−0.05 (0.36) $p = 0.89$
Decline rate	+0.15 (0.11) $p = 0.15$	+ 0.39 (0.15) **p = 0.009**	+0.23 (0.14) $p = 0.08^{\$}$	- 0.05 (0.08) $p = 0.56$	+0.14 (0.12) $p = 0.24$	+0.19 (0.10) $p = 0.07^{\$}$

tf denotes the temporal-frontal factor, *p* the parietal factor, and *l* the long fibre bundle factor. For example, the second row suggests that participants expressing more temporal-frontal factor has higher baseline executive function than those expressing more long-fibre bundle factor. Significant comparisons are highlighted in bold.

between all the $C_K{}^2$ pairs of factors were calculated, and the maximum correlation was used to obtain the degree of spatial separation defined as:

$$1 - \max_{\forall i,j \in \{1,\ldots,K\}, i \neq j} (\mathrm{corr}(P(Voxel \mid Factor_i), P(Voxel \mid Factor_j))).$$

We first observed qualitatively how the estimated factors vary with K from 2 to 4 (Fig.2) and then adopted grid search to quantitatively choose the best K that maximize the separation of latent factors for K from 2 to 6. After determining K, LDA analysis was performed on baseline AD scans to identify latent WM impairment factors. Then, the LDA analysis inferred factor compositions for all 893 time-points of 232 participants (Table 1), which were further employed in the following statistical analyses.

The following statistical analyses were based on the LDA model with $K = 3$, which identified three latent WM impairment factors including the temporal-frontal factor *tf*, the parietal factor *p* and the long fibre bundle factor *l*.

2.6. Evaluating the longitudinal stability of latent factors

One question naturally arises: does the latent WM impairment factors change as the disease progresses? In order to quantify the stability of latent factors, we examined participants who had both their baseline scans and follow-up scans available (3/6/12/24 months respectively). If the latent factors are stable as AD progresses, then the factor compositions of individuals should remain the same over time. To investigate the stability of the factor compositions, we drew scatter plots where each dot represents an individual with the *x* coordinate representing the composition proportion of a particular factor at baseline and the *y* coordinate representing the corresponding factor composition proportion after 12 months (Fig.5, *SI Appendix, Table S1 and Fig.S6*). A linear relationship with a slope close to 1 in the scatter plot would indicate stable the WM impairment factor over the disease progression.

2.7. Modelling the longitudinal cognitive decline using latent WM impairment factors

Here we attempt to decouple the associations between WM impairment factors and cognitive decline in AD. Relating to memory and executive function, we asked two questions: (i) whether the WM impairment factors correspond to the same baseline cognitive scores and the decline rates throughout the disease course; and (ii) when modelling the longitudinal decline, whether the latent factors lead to higher sensitivity compared to the traditional modelling approach.

Accordingly, linear mixed effect (LME) models (Bernal-Rusiel et al., 2013) were used to answer the two questions. For the first question, the response variable y of the LME model consisted of the longitudinal ADNI-Mem (508 time-points) and ADNI-EF scores respectively (504 time-points, as described in Table 1). The explanatory fixed-effects variables included: (i) mean FA values on impaired regions *FA* (*SI Appendix)* indicating the severity of WM impairment; (ii) factor compositions indicating the subtype, i.e., the compositions of parietal factor *p* and the long fibre bundle factor *l* (the composition of temporal-frontal factor *tf* is implicit in $p + l + tf = 1$); (iii) elapsed time t from the baseline; (iv) interactions between factor compositions and time from baseline, i.e. $p \cdot t$ and $l \cdot t$; and (v) nuisance variables, including amyloid x_1, sex x_2, education x_3 and gray matter volume x_4 (age and white matter volume were not significant associated with cognition as evaluated by ANOVA). The resulting LME model, denoted as M, is given by

$$y = (\beta_0 + \beta_{FA} \cdot FA + \beta_p \cdot p + \beta_l \cdot l) + (\beta_{t0} + \beta_{tp} \cdot p + \beta_{tl} \cdot l) \cdot t + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \varepsilon,$$

where y denotes the ADNI-Mem/ADNI-EF score, ε is the residual, and β_* indicate the regression coefficients. As to interpret the LME model, $\beta_0 + \beta_p$reflects the parietal factor's baseline cognitive score,β_{t0}reflects the temporal-frontal factor's decline rate, and $\beta_{t0} + \beta_{tp}$ reflects the parietal factor's decline rate. We aimed to investigate whether the three WM impairment factors have the same baseline cognitive scores and decline rates through the disease course. Accordingly, as for hypothesis tests, we could ask whether the temporal-frontal factor have significantly stronger effect on ADNI-Mem than the long fibre bundle factor, and whether the parietal factor and the temporal-frontal factor equally affect the decline rate of ADNI-EF throughout the disease stages. The results of the statistical tests are shown in Table 2.

To compare our model with the traditional model and demonstrate the predictive power of the latent factor compositions, we compared our model M with the traditional model that simply deploys the traditional measures without factoring the compositions (or into subtypes). A traditional model, denoted as $M_{classic}$, would be

$$y = \beta_0 + \beta_{FA} \cdot FA + \beta_{t0} \cdot t + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \varepsilon,$$

The ANOVA analyses was performed to evaluate M and $M_{classic}$, respectively, in predicting memory and executive function.

2.8. Association of APOE genotype with WM impairment factors

We investigated whether *APOE* genotype, as a well-known genetic risk factor, was associated with the disease subtype. More specifically, participants with different genotypes might have different associations with the latent impairment factors. To test this hypothesis, we performed an analysis with a GLM model, including the following fixed-effects variables: (i) binary group indicators, i.e., the EMCI group *em*, the LMCI group *lm* and the AD group *ad* (the NC group is implicit); and (ii) factor compositions, i.e., the parietal factor *p* and the long fibre bundle factor *l* (the temporal-frontal factor *tf* is implicit in $tf + p + l = 1$). The resulting GLM model is given by

$$y = \beta_0 + \beta_{em} \cdot em + \beta_{lm} \cdot lm + \beta_{ad} \cdot ad + \beta_p p + \beta_l l + \varepsilon,$$

where y denotes the *APOE* $\varepsilon 4/\varepsilon 2$ carrier status (0, 1 carrier or 2 carriers), β_* indicate the regression coefficients, and ε is the residual. Note that β_p captures the difference between the responses of the parietal factor *p* and the temporal-frontal factor *tf*, and β_l captures the difference between the responses of the long fibre bundle factor *l* and temporal factor *tf*.

In order to compare the association between the genetic status and different factors, we further tested the null hypothesis $H_0 : C\beta = 0$, where $\beta = [\beta_0, \beta_{em}, \beta_{lm}, \beta_{ad}, \beta_p, \beta_l]^T$and C denotes the linear contrast corresponding to a particular scientific question. For example, to test the differences

Table 3
Comparison of effects of APOE ε4 and APOE ε2 on different latent factors.

(A) APOE ε4			(B) APOE ε2		
$p - tf$	$l - tf$	$l - p$	$p - tf$	$l - tf$	$l - p$
+0.19 (0.23) $p = 0.42$	+ 1.1 (0.28) $\mathbf{p = 0.0001}$	+ 0.91 (0.25) $\mathbf{p = 0.0006}$	- 0.26 (0.12) $\mathbf{p = 0.02}$	−0.15 (0.14) $p = 0.28$	−0.11 (0.13) $p = 0.39$

tf denotes the temporal-frontal factor, *p* the parietal factor, and *l* the long fibre bundle factor. Significant comparisons are highlighted in bold.

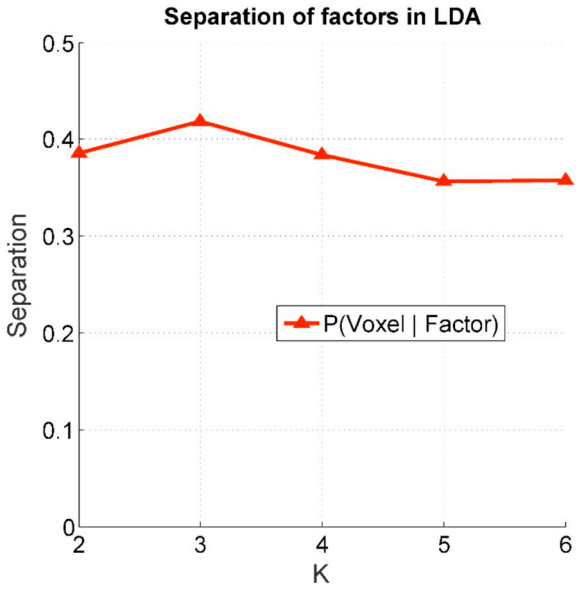

Fig. 3. The degree of spatial separation among latent factors changes with respect to *K*, the number of factors. $K = 3$ achieves the highest separation among latent factors.

between all factors, we chose $C = [0,0,0,0,1,0;0,0,0,0,0,1]$; for differences between the parietal factor *p* and the long fibre bundle factor *l*, we choose $C = [0,0,0,0,1,-1]$. The results of the statistical tests are shown in Table 3.

2.9. Implementation details

For LDA analysis, we used the VEM implementation (www.cs.princeton.edu/~blei/lda-c/) and for the LME analysis, the LME toolbox was used (github.com/NeuroStats/lme). The code for identifying the coherence of the WM impairment factors was adapted from (github.com/ThomasYeoLab/CBIG). We used nilearn (nilearn.github.io) to visualize WM impairment factors, and ggplot2 (www.ggplot2.org) and ggtern (www.ggtern.com) for visualizing factor compositions. The code and intermediate results will be released on GitHub (github.com/xiuchao/ldaFA).

3. Results

3.1. Overall approach

Our study consisted of the following parts: (i) identifying multiple WM impairment factors in AD patients and inferring factor compositions of individual participants (both AD and non-AD) by using LDA; (ii) exploring the stability of latent factors as the disease progresses; (iii) examining the associations between WM impairment factors and cognition (i.e. ADNI-Mem and ADNI-EF); and (iv) examining the association between the genotype (*APOE ε4/ε2*) and the disease subtype (i.e. factor compositions). The details of the results are described as follows.

3.2. WM impairment factors in AD

We hypothesized that AD patients express one or more WM impairment factors and this hypothesis can be captured by the LDA model (Fig.1). For the 48 AD patients at baseline, given the pre-processed FA skeleton maps and the number of *K* factors, the LDA method estimated the probability distribution of each impairment factor $p(Voxel \mid Factor)$ and the factor composition $p(Factor \mid Patient)$ of individual participant.

The number of impairment factors *K* is a free parameter in the LDA model. We first qualitatively inspected how the estimated factors $p(Voxel \mid Factor)$ changed as *K* was varied from 2 to 4. The impairment factors were spatially consistent (*Section 2.3*) for all $K = 2$ to 4 (Fig.2). Specifically, the $K = 2$ model revealed an impairment factor throughout cortical regions, which included temporal regions (parahippocampal WM and temporal WM), frontal and parietal WM, and another long fibre bundle factor that spreads over corpus callosum and superior longitudinal fasciculus (SLF). The $K = 3$ model revealed a similar long fibre bundle factor, whereas the cortical factor split into a parietal factor with WM impairment mainly in parietal lobe and a temporal-frontal factor associated with extensive WM impairment in para-hippocampus gyrus, temporal lobe and frontal lobe. The $K = 4$ model revealed three latent factors similar with the $K = 3$ model, as well as a redundant factor that spanned over the brain regions identified by the other three factors. A closer inspection revealed that the redundant factor was a mixture of the other three factors, and thus did not provide any new insight into susceptible WM tissues. In addition, when *K* further increased, more redundant factors were observed. To summarize, the latent impairment factors were spatially consistent as *K* was varied (Fig.2, *Fig.S1*); whereas redundant factors appeared when *K* increased from 3 to 4 or more.

We further quantitatively evaluated the spatial separation among individual latent factors within each LDA model of a particular *K*. As visually seen in Fig.3, $K = 3$ achieved the largest spatial separation between latent factors. Therefore, $K = 3$ was selected as the best number of latent factors on this data. Notably, our findings are consistent with previous AD subtype studies in terms of the number of subtypes. For example, Murray's study revealed three AD subtypes, which is based on neuropathologcial evidence (Murray et al., 2011). In addition, other previous studies based on structural MRI also reported three distinct patterns of GM atrophy (Zhang et al., 2016; Noh et al., 2014) and NFT changes (Whitwell et al., 2012).

In addition, among the 48 AD patients, 38 had their cerebrospinal fluid (CSF) amyloid data available and 35 of 38 patients were Aβ + (CSF amyloid concentration < 192 pg/mL). We performed LDA on the 35 Aβ + AD dementia patients alone and compared the WM impairment factors with those derived using all 48 AD patients (*SI Appendix Fig.S2*). The WM impairment factors obtained under the two settings (48 AD vs. 35 Aβ + AD) were almost identical, with an average correlation $r = 0.95$. Considering such a high similarity, we used the factors derived from the larger sample for further analysis.

3.3. The stability of latent factors as AD progresses

The probabilistic WM impairment distributions $p(Voxel \mid Factor)$ estimated from the AD patients were used to infer factor compositions $p(Factor \mid Patient)$ of the non-demented participants, including EMCI, LMCI and NC, by using the variational expectation-maximization

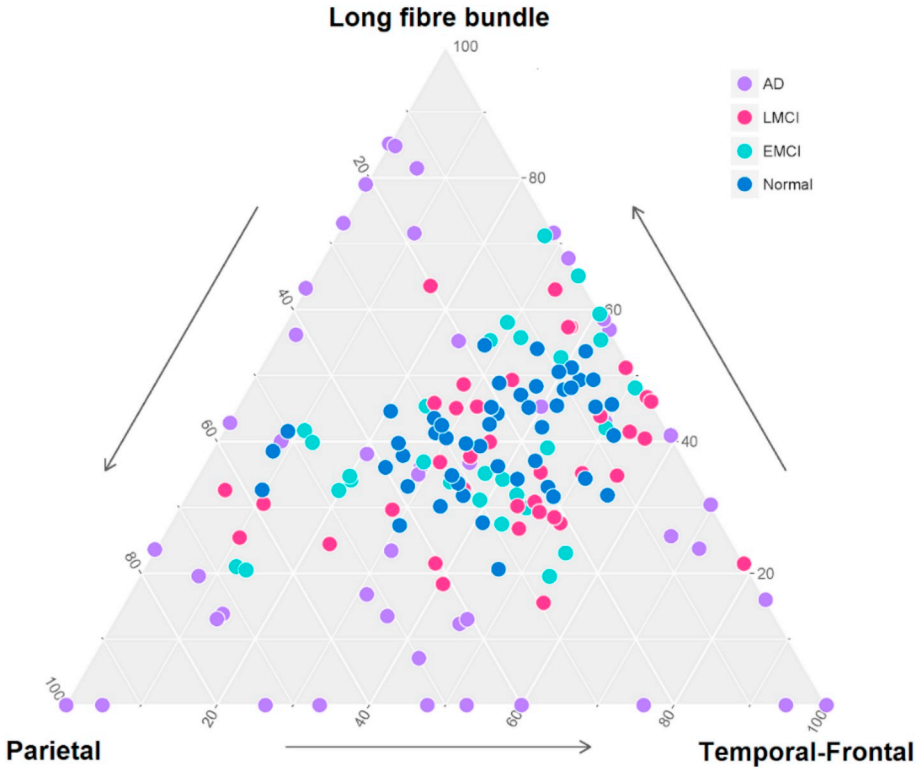

Fig. 4. Factor compositions of participants at baseline. Each dot corresponds to a participant and its location represents his/her factor composition. Corners of this triangle represent pure factors; a closer distance to a corner indicates a higher probability to the respective factor. As can be seen, most participants express multiple factors, and NC participants are clustered more closely than AD patients, indicating the heterogeneity of AD.

(VEM) LDA algorithm (Blei et al., 2003). As shown in Fig.4, most of the participants expressed multiple WM impairment factors instead of a single predominant factor. It is worth noting that NC participants were more homogeneous (i.e., clustered close to each other), while AD patients were more heterogeneous.

To investigate whether the expression of impairment factors changed over the progression of AD, we compared the factor compositions of participants who had both the baseline scan and the 12-month follow-up scan available. As shown in Fig.5, the factor probabilities were highly consistent ($r > 0.96$ across all the three factors), indicating that the factor compositions remained stable over the disease progression (for time-points at 3/6/24 months, see *SI Appendix Table S1*). This result is in accordance with earlier study on GM, which reported that the atrophy factor compositions were stable after two years (Zhang et al., 2016). This finding further supports that the impairment factors represents different impairment subtypes instead of different disease stages.

3.4. Association between WM impairment factors and cognition

For memory (ADNI-Mem) and executive function (ADNI-EF), we investigated whether distinct WM impairment factors were associated with the same baseline scores and decline rate as the disease progressed.

First, diagnostic groups differed in cognition as expected (Table 1). ADNI-Mem was significantly lower in each diagnostic group (NC > EMCI > LMCI > AD, $p < 0.05$) while ADNI-EF followed similar pattern (NC > EMCI = LMCI > AD, $p < 0.05$).

Second, an LME model was used to relate longitudinal cognitive scores (ADNI-Mem/ADNI-EF) with the observed predictors, including

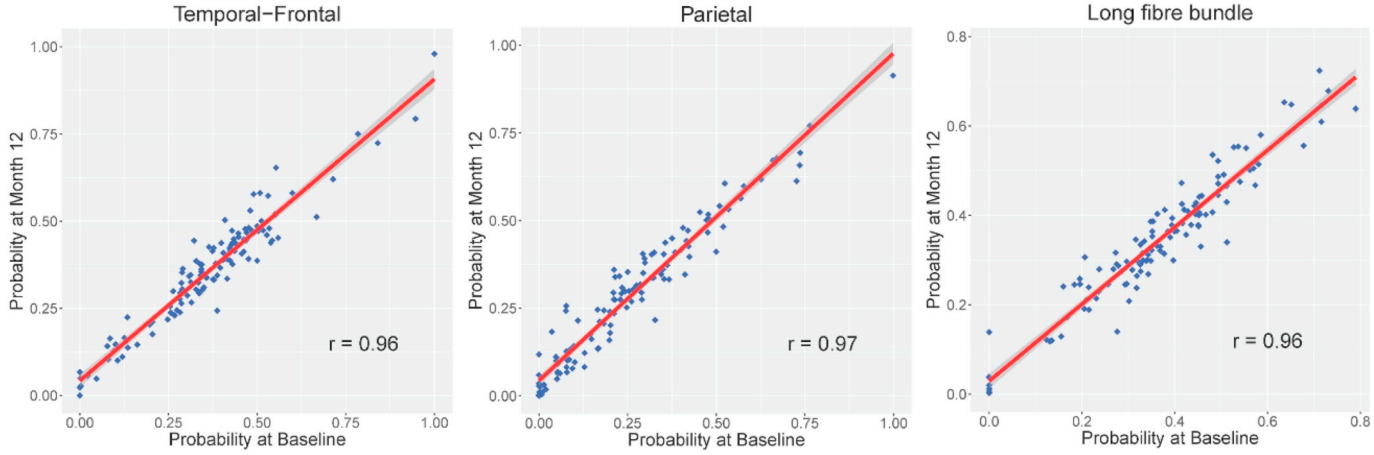

Fig. 5. Stability of factor compositions of the three factors over 12 months. Each dot represents a participant, of which the x and y axes represent the probabilities of factors at the baseline and at 12 months later, respectively. Linear fits to the scatters are close to the $y = x$ line with correlations $r > 0.96$ for all the three impairment factors, suggesting that the factor compositions are stable during disease progression.

the mean FA of impaired brain regions (*SI Appendix*), the factor compositions, the progression time, and interactions between time and factor compositions (*Method 2.6*). We studied whether each factor had the same baseline cognition and decline rate through the disease course. For EF (Table 2A), the temporal-frontal factor was more associated with baseline EF compared with the parietal factor ($p = 0.02$); the long fibre bundle factor was significantly more associated with the EF decline rate compared with the temporal-frontal factor ($p = 0.009$), and tended to be more associated with the EF decline rate than the parietal factor ($p = 0.08$). For memory (Table 2B), the three factors were almost equally associated with baseline memory and decline rate across the disease spectrum. The long fibre bundle factor tended to be more associated with memory decline rate than the parietal factor in AD patients ($p = 0.07$). Our findings suggest that latent factors have distinct effects on different cognitive domains.

Third, we evaluated whether incorporating factor compositions could lead to better modelling of longitudinal effects on cognition. ANOVA analysis was performed to compare the latent factor based model M and the traditional model $M_{classic}$ in distinct cognitive domains. We found that the latent factor compositions significantly contributed to better prediction of EF ($p = 0.01$) while had no significant effect on memory prediction ($p = 0.48$). These results comply with the aforementioned findings (in Table 2) and show that by enrolling subtype information (through latent factor composition), longitudinal changes can be predicted with higher sensitivity in specific cognition domains.

3.5. Association between APOE genotype and latent factors

The *APOE* genotype has been identified as genetic susceptibility factor in AD (Filippini et al., 2009; Drzezga et al., 2009). We investigated whether the *APOE* genotype has different effects on different white matter impairment factors by using a GLM model *(Method 2.7)*. As summarized in Table 3A, the long fibre bundle factor was more associated with *APOE* ε4 than the temporal-frontal ($p = 0.0001$) and the parietal factor ($p = 0.0006$). As summarized in Table 3B, the temporal-frontal factor was more associated with APOE ε2 than the parietal factor ($p = 0.02$). Our findings suggest that the genotype affects the disease subtype (i.e. factor compositions) differently.

4. Discussion

In this work, instead of assuming an overall disease effect on WM impairment, we identified distinct latent impairment factors in AD patients by using LDA modelling of DTI data (Fig.1). Three latent impairment factors were revealed in AD (Figs.2–3). Each participant was allowed to exhibit their own unique factor compositions (Fig.4) under the LDA framework, which were found to be stable across stages (Fig.5). These impairment factors deteriorated longitudinal memory and executive function differently along the full disease course of AD (Table 2). Moreover, *APOE* genotype has distinct effects on different WM impairment factors (Table 3). To summarize, our work suggests that the heterogeneity of WM impairment in AD patients are associated with the variance in longitudinal cognitive decline and *APOE* genotype is associated with latent factors. In summary, this work revealed the links between WM impairment factors, *APOE* genotype, and cognition in distinct domains.

4.1. WM impairment factors in AD

$K = 3$ model achieved the largest dissociation between factors within each LDA model and the three impairment factors were investigated in this study. Earlier studies have consistently reported three subtypes in AD (Murray et al., 2011; Zhang et al., 2016; Noh et al., 2014; Park et al., 2017). Our three-factor model includes a temporal-frontal factor, a parietal factor, and a long fibre bundle factor (Fig.2). Specifically, we observed that the temporal-frontal factor was linked with WM impairment in hippocampus, parahippocampus, medial temporal lobe and frontal lobe. Earlier work demonstrated that changes in parahippocampal region may result in isolation of the hippocampus, which contribute to impairment in memory (Salat et al., 2010); The frontal lobe regions have been suggested to associate with executive function (Smith et al., 2011; Brinkman et al., 2012). The parietal factor was associated with the impairment in precuneus and parietal lobe. The precuneus has been associated with retrieval of episodic memory and self-processing operations (Shallice et al., 1994; Malouin et al., 2003; Kircher et al., 2002). The long fibre bundle factor was associated with the WM impairment in splenium and the body of corpus callosum (CC) and SLF. The splenium of CC predominantly connects the temporal lobe and the occipital lobe between the two hemispheres. The SLF is an association fibre tract that contains three separate components (Makris et al., 2004). SLF II (Table S2), the major component of SLF, connects the caudal-inferior parietal cortex and dorsolateral prefrontal cortex, and serves to transmit working memory in the prefrontal cortex to provide the parietal cortex with information to regulate spatial attention and retrieval of spatial information (De Schotten et al., 2011).

Accumulating evidence suggests that structural and histological changes in AD are heterogeneous. For example, heterogeneity in cortical atrophy has been evinced by using structural MRI (Zhang et al., 2016; Noh et al., 2014; Park et al., 2017) and heterogeneity in NFT distributions has been demonstrated from thioflavinS fluorescent microscopy (Murray et al., 2011). While a few studies have investigated the heterogeneous white matter alterations in AD, using DTI data (Gouw et al., 2008; Bartzokis et al., 2004), to our best knowledge, this is the first study that employ latent factor decomposition on white matter impairment to profile subtypes of AD. By referring to the field knowledge, we can interpret a particular WM impairment factor and how it may be associated with various cognitive domains. However, it is hard to make direct comparison with earlier subtype studies which were mostly based on GM atrophy. We could see rough spatial correspondence between WM impairment factors and GM atrophy factors identified in earlier studies. For example, a recent GM atrophy study identified three subtypes: the parietal predominant subtype that roughly correspond to our parietal WM impairment factor, the medial temporal predominant subtype that roughly correspond to our temporal-frontal factor, and the diffuse atrophy subtype that roughly corresponds to the long fibre bundle factor (Park et al., 2017). Yet such correspondence is not strict due to the tissue difference and should be interpreted with caution. Secondary degeneration of WM, as a result of cortical atrophy, may contribute to the observed rough spatial correspondence. Furthermore, we identified GM atrophy factors based on the structural MRI data of the same dataset (SI Appendix, *Fig.S5*), which also showed no exact one-to-one correspondence with the WM impairment factors. This further illustrates that despite partly as a secondary effect of cortical atrophy, WM impairment still provides unique information on AD. Previous study on ADNI dataset also testified that DTI provides unique information about cognition instead of merely recapitulations of established association between other biomarkers of AD (e.G. *hippocampus* volume) (Scott et al., 2017). Accordingly, WM impairment factors uniquely contribute to portraying subtypes of AD. As future work, one can probe AD subtypes by simultaneously considering GM atrophy and WM impairment.

4.2. Factor compositions were stable as the disease progresses

The LDA method in the Bayesian framework treats AD cases as entities with mixed membership of one or multiple factors, and allows each participant to exhibit his/her own unique factor composition. For a hypothetical patient, his/her factor composition profile could be portrayed as expressing 60% the temporal-frontal factor, 20% the parietal factor and 10% the long fibre bundle factor. Such probabilistic factor composition extends previous approaches that classified patients into distinct subtypes (Whitwell et al., 2012; Byun et al., 2015), and

makes it more feasible to study how the heterogeneity of WM impairment was associated with variance in cognitive decline in distinct cognition domain. However, a natural question arises: was the observed heterogeneity attributed to disease stages, or disease expressions, i.e., subtypes (Ritchie and Touchon, 1992; Young et al., 2014)?

To address this question, we investigated whether that factor compositions were stable after 3/6/12 months (Fig.5, *SI Appendix, Table S1 and Fig.S6*). If the disease stages hypothesis stands, a hypothetical individual would express more cortical factor after 12 months, which is not supported by our results. As the factor compositions were stable as the disease progresses, we can conclude that the heterogeneity in WM degeneration is primarily due to the disease subtypes. Our result complies with an earlier AD subtype study that reported stable GM atrophy factor across disease stages (Zhang et al., 2016).

4.3. Factors reflect distinct association with longitudinal cognitive decline

Our work revealed that distinct WM impairment factors were associated with different cognition domains differently as the disease progressed (Table 2).

Executive function includes higher-order cognitive processes such as shifting attention, cognitive flexibility, cognitive fluency, working memory, planning and organization. The temporal-frontal factor showed more association with baseline EF compared with the parietal factor. This finding was in line with earlier work that reported that frontal lobe brain regions were associated with executive function (Smith et al., 2011; Brinkman et al., 2012), and the frontal-parietal network has been implicated in facilitating the integration and control of executive processes (Barbey et al., 2012). The long fibre bundle factor was more associated with EF decline rate compared with the temporal-frontal factor ($p = 0.009$), and was inclined to be more associated with EF decline rate than the parietal factor ($p = 0.08$). By looking into the long fibre bundle factor, the largest commissural tract CC predominantly connects the temporal lobe and the occipital lobe between the two hemispheres; the association fibre tract SLF connects the caudal-inferior parietal cortex and dorsolateral prefrontal cortex. The prior knowledge that CC and SLF serve to transmit working memory in the prefrontal cortex and to provide the parietal cortex with information for regulating spatial attention and retrieval of spatial information (De Schotten et al., 2011), partially supports our discovery that the long fibre bundle factor was most associated with to EF decline rate in AD.

For memory, the three factors contribute equally to the memory decline in each group. Also, as illustrated in *Section 4.1*, the temporal-frontal factor involve WM in hippocampus and temporal lobe, the most well-known brain regions related with memory (Salat et al., 2010); the parietal factor includes the precuneus and parietal lobe WM that is related with episodic memory retrieval (Shallice et al., 1994; Malouin et al., 2003; Kircher et al., 2002); the long fibre bundle factor includes CC and SLF, which are related with visual memory (De Schotten et al., 2011; Stanislav et al., 2013). We can assume that the three factors are all associated with memory decline, yet further studies are necessary to examine whether a dominant factor exists by using a larger data set.

Concerning previous DTI studies on the same ADNI dataset, though cross-sectional studies (Grieve et al., 2007; Bennett and Madden, 2014) showed the association between WM impairment measured by DTI metrics and executive function decline, yet longitudinal study (Scott et al., 2017) suggested the insensitivity of DTI metrics to reveal the changing rate of cognitive decline. Our study, based on the same dataset, revealed that the long fibre bundle factor was more associated with EF decline rate than the other two latent WM impairment factors. Our findings suggested that by differentiating the subtypes of AD, we could achieve a finer view in interpreting the data. Our work corresponds with earlier cross-sectional studies (Grieve et al., 2007; Bennett and Madden, 2014) and expands the longitudinal study (Scott et al., 2017) by illustrating the effect of latent factors on longitudinal cognitive decline of multi-domain.

4.4. APOE genotype was associated with latent factors

Since disease subtypes may have genetic origins, we investigated whether risk genes are associated with latent impairment factors. *APOE* has been the best-established genetic risk factor, which associated with structural and functional changes in the elderly and AD patients (Filippini et al., 2009; Drzezga et al., 2009). The major role of apoE protein in the brain is to transport lipid components that facilitate the shaping of the myelin sheath (Mahley, 1988; Han, 2007), thus affecting WM integrity (Heise et al., 2011). In this study, we found *APOE* ε4 was most associated with the long-fibre bundle factor, while *APOE* ε2 was more associated with the temporal-frontal factor compared with the parietal factor (Table 3). Our work lends support to the hypothesis that the *APOE* genotype modulates the phenotype of AD through influences on specific large-scale brain networks (Wolk et al., 2010).

Revisiting the discoveries in LME analysis, the baseline executive function was more associated with temporal-frontal factor compared with the parietal factor, while *APOE* ε2 was also more associated temporal-frontal factor compared with the parietal factor (Tables 2–3). In addition, the EF decline rate was most associated with the long fibre bundle factor, while *APOE* ε4 was also most associated with the long fibre bundle factor. Such consistency suggested that our work moved one more step towards dissociating the links between gene, brain structure and cognition. More inspiringly, assume a hypothetical patient with *APOE* ε4/ε4 (two ε4 alleles), we could expect a higher chance of expressing more long fibre bundle factor and faster EF decline. Similarly, a hypothetical patient with APOE ε2/ε2 (two ε2 alleles) could expect a higher chance of expressing more temporal-frontal factor compared with the parietal factor, as well as higher baseline EF scores.

To summarize, variation in genotypes of patients were associated with expressions of WM impairment factors (i.e., subtypes), which would manifest in features of cognitive decline. By looking such association, we could take one-step forward to precision medicine.

4.5. Limitations

AD dementia is a heterogeneous disease in that the origins and spread of the disease may differ substantially between different subtypes. Earlier work explored the subtypes, using neurological scores (Stopford et al., 2008), the distribution of neurofibrillary tangles (Murray et al., 2011), and cortical atrophy (Zhang et al., 2016; Noh et al., 2014; Park et al., 2017). Along this line of research, our work demonstrated that the patterns of WM impairment could also be used to portray AD subtypes. The probabilistic profiles of subtypes, portrayed in the LDA framework, were associated with distinct trajectories of cognitive decline and may originate from different genotypes. However, there are limitations of our work.

First, the pathological basis of these WM impairment factors is still obscure. A recent study suggested that myelin and oligodxendrocytes are especially vulnerable to amyloid pathology (Dean et al., 2017). Further studies are needed to address the characterisation of AD subtypes using other biomarkers, e.g. white matter hyper-intensity, amyloid and tau immunohistochemistry, which may lead to a better understanding of AD heterogeneity and further facilitate early diagnosis and the monitoring of AD progression (Frisoni et al., 2010; Cash et al., 2014).

Second, pathological findings from autopsied cases provide the golden standard for revealing structural heterogeneity as such studies represent advanced stages of AD. However, for degeneration at the early disease stages, we have been unable to find a well-recognized reference to evaluate the structural heterogeneity findings. In addition, it is hard to make direct comparisons between neuropathologically

defined subtypes, GM atrophy defined subtypes and WM impairment defined subtypes, especially as these observations were based on different participants and different analysis protocols. For this study, we presented our subtype observations on DTI data under the LDA framework, which could serve as a reference for clinicians and pathologist to unveil the broader picture of subtypes in AD.

Third, to simplify the analysis pipeline and increase reproducibility of this study, we used preprocessed images by ADNI where the corrected FA images provided the largest sample set. Therefore, we used FA measure for this study. As an initial trial to probe subtypes using DTI, this work has ignored other DTI metrics that characterize the same underlying microstructures from different aspects. For example, mean diffusivity (MD) was used as widely as FA to map white matter changes in AD; radial diffusivity, and axonal diffusivity gained more popularity as DTI markers of axonal and myeline damage, respectively (Acosta-Cabronero et al., 2009; Winklewski et al., 2018). Moreover, though FA has been widely used, it was suggested that FA does not always correlate well with the actual individual fibre anisotropy and maybe suboptimal in detecting disease processes that affect myelination (Leow et al., 2009). Future subtyping studies based on DTI may benefit from looking into other diffusion measures that provide more information in the diffusion profile. Moreover, as a previous study (Konukoglu et al., 2016) suggested, joint analysis of the diffusion measures could be more powerful in characterizing histopathology and provide information about disease progresses not available through examination of any measure in isolation. Future DTI studies may benefit from such multivariate method for joint analysis of diffusion measures.

Finally, our study utilized the DTI images of ADNI-GO and ADNI-2 dataset, in which the enrolled patients were followed up for no > 5 years. As ADNI is still collecting more longitudinal data, future studies may benefit from enriched data resources and yield more statistically reliable analytical results. Also, as a study based on DTI data, this study has limitations inherent in DTI. More advanced imaging and modelling techniques, such as high-angular resolution diffusion imaging (HARDI), can resolve more complex diffusion geometries and further facilitate finer portrait of regional tissue impairment that would lead to more accurate subtype profiles.

5. Conclusions

By identifying latent WM impairment factors with LDA, our work lends support to the growing belief that AD is heterogeneous rather than a single disease entity. We identified three impairment factors: the temporal-frontal factor, the parietal factor and the long fibre bundle factor. These factors were associated with distinct decline trajectories of memory and executive function across the full course of AD. Moreover, individual participant could express multiple WM impairment factors to different degrees. The heterogeneous factor compositions among participants were associated with different genotypes and can facilitate predicting individual-level cognition and contribute to precision medicine.

Conflict of interest

The authors declare they have no conflicts of interest.

Acknowledgements

This work was supported by AcRF Tier-1 grant RG 149/17 grant by the Ministry of Education, Singapore. We thank Dr. Shaohua Li for his valuable comments and suggestions. Data collection and sharing for this project was funded by the ADNI, United States (NIH Grant U01 AG024904) and the Department of Defense (DOD), United States, ADNI (DOD Grant W81XWH-12-2-0012). The ADNI is funded by the National Institute on Aging, United States and the National Institute of Biomedical Imaging and Bioengineering, United States 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). 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. The ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. Data used in preparation of this article were obtained from the ADNI database (adni.loni.usc.edu). 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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.nicl.2018.10.026.

References

!!!Invalid Citation!!!.

Acosta-Cabronero, J., Williams, G.B., Pengas, G., Nestor, P.J., 2009. Absolute diffusivities define the landscape of white matter degeneration in Alzheimer's disease. Brain 133, 529–539.

Agosta, F., Pievani, M., Sala, S., Geroldi, C., Galluzzi, S., Frisoni, G.B., et al., 2011. White matter damage in Alzheimer disease and its relationship to gray matter atrophy. Radiology 258, 853–863.

Barbey, A.K., Colom, R., Solomon, J., Krueger, F., Forbes, C., Grafman, J., 2012. An integrative architecture for general intelligence and executive function revealed by lesion mapping. Brain 135, 1154–1164.

Bartzokis, G., Sultzer, D., Lu, P.H., Nuechterlein, K.H., Mintz, J., Cummings, J.L., 2004. Heterogeneous age-related breakdown of white matter structural integrity: implications for cortical "disconnection" in aging and Alzheimer's disease. Neurobiol. Aging 25, 843–851.

Bennett, I.J., Madden, D.J., 2014. Disconnected aging: cerebral white matter integrity and age-related differences in cognition. Neuroscience 276, 187–205.

Bernal-Rusiel, J.L., Greve, D.N., Reuter, M., Fischl, B., Sabuncu, M.R., A. S. D. N. Initiative, 2013. Statistical analysis of longitudinal neuroimage data with linear mixed effects models. NeuroImage 66, 249–260.

Blei, D.M., Ng, A.Y., Jordan, M.I., 2003. Latent dirichlet allocation. J. Mach. Learn. Res. 3, 993–1022.

Bozzali, M., Falini, A., Franceschi, M., Cercignani, M., Zuffi, M., Scotti, G., et al., 2002. White matter damage in Alzheimer's disease assessed in vivo using diffusion tensor magnetic resonance imaging. *J. Neurol.* Neurosurg. Psychiatry 72, 742–746.

Brinkman, T.M., Reddick, W.E., Luxton, J., Glass, J.O., Sabin, N.D., Srivastava, D.K., et al., 2012. Cerebral white matter integrity and executive function in adult survivors of childhood medulloblastoma. Neuro-Oncology 14, iv25–iv36.

Butters, M.A., Lopez, O.L., Becker, J.T., 1996. Focal temporal lobe dysfunction in probable Alzheimer's disease predicts a slow rate of cognitive decline. Neurology 46, 687–692.

Byun, M.S., Kim, S.E., Park, J., Yi, D., Choe, Y.M., Sohn, B.K., et al., 2015. Heterogeneity of regional brain atrophy patterns associated with distinct progression rates in Alzheimer's disease. PLoS ONE 10, e0142756.

Cash, D.M., Rohrer, J.D., Ryan, N.S., Ourselin, S., Fox, N.C., 2014. Imaging endpoints for clinical trials in Alzheimer's disease. Alzheimers Res. Ther. 6, 87.

Caso, F., Agosta, F., Mattavelli, D., Migliaccio, R., Canu, E., Magnani, G., et al., 2015. White matter degeneration in atypical Alzheimer disease. Radiology 277, 162–172.

Crane, P.K., Carle, A., Gibbons, L.E., Insel, P., Mackin, R.S., Gross, A., et al., 2012. Development and assessment of a composite score for memory in the Alzheimer's Disease Neuroimaging Initiative (ADNI). Brain Imag. Behav. 6, 502–516.

De Schotten, M.T., Dell'Acqua, F., Forkel, S.J., Simmons, A., Vergani, F., Murphy, D.G., et al., 2011. A lateralized brain network for visuospatial attention. Nat. Neurosci. 14, 1245–1246.

Dean, D.C., Hurley, S.A., Kecskemeti, S.R., O'Grady, J.P., Canda, C., Davenport-Sis, N.J.,

et al., 2017. Association of amyloid pathology with myelin alteration in preclinical Alzheimer Disease. JAMA Neurol. 74, 41–49.
Dickerson, B.C., David, A., 2011. Wolk, and Alzheimer's Disease Neuroimaging Initiative. "Dysexecutive versus amnesic phenotypes of very mild Alzheimer's disease are associated with distinct clinical, genetic and cortical thinning characteristics. J. Neurol. Neurosurg. Psychiatry 82 (1), 45–51.
Drzezga, A., Grimmer, T., Henriksen, G., Mühlau, M., Perneczky, R., Miederer, I., et al., 2009. Effect of APOE genotype on amyloid plaque load and gray matter volume in Alzheimer disease. Neurology 72, 1487–1494.
Filippini, N., Rao, A., Wetten, S., Gibson, R.A., Borrie, M., Guzman, D., et al., 2009. Anatomically-distinct genetic associations of APOE ε4 allele load with regional cortical atrophy in Alzheimer's disease. NeuroImage 44, 724–728.
Frisoni, G.B., Fox, N.C., Jack, C.R., Scheltens, P., Thompson, P.M., 2010. The clinical use of structural MRI in Alzheimer disease. Nat. Rev. Neurol. 6, 67–77.
Gefen, T., Gasho, K., Rademaker, A., Lalehzari, M., Weintraub, S., Rogalski, E., et al., 2012. Clinically concordant variations of Alzheimer pathology in aphasic versus amnestic dementia. Brain 135, 1554–1565.
Gibbons, L.E., Carle, A.C., Mackin, R.S., Harvey, D., Mukherjee, S., Insel, P., et al., 2012. A composite score for executive functioning, validated in Alzheimer's Disease Neuroimaging Initiative (ADNI) participants with baseline mild cognitive impairment. Brain Imag. Behav. 6, 517–527.
Gouw, A., Seewann, A., Vrenken, H., Van Der Flier, W., Rozemuller, J., Barkhof, F., et al., 2008. Heterogeneity of white matter hyperintensities in Alzheimer's disease: post-mortem quantitative MRI and neuropathology. Brain 131, 3286–3298.
Grieve, S., Williams, L., Paul, R., Clark, C., Gordon, E., 2007. Cognitive aging, executive function, and fractional anisotropy: a diffusion tensor MR imaging study. Am. J. Neuroradiol. 28, 226–235.
Han, X., 2007. Potential mechanisms contributing to sulfatide depletion at the earliest clinically recognizable stage of Alzheimer's disease: a tale of shotgun lipidomics. J. Neurochem. 103, 171–179.
Heise, V., Filippini, N., Ebmeier, K., MacKay, C., 2011. The APOE [varepsilon] 4 allele modulates brain white matter integrity in healthy adults. Mol. Psychiatry 16, 908.
Johnson, J.K., Head, E., Kim, R., Starr, A., Cotman, C.W., 1999. Clinical and pathological evidence for a frontal variant of Alzheimer disease. Arch. Neurol. 56, 1233–1239.
Kircher, T.T., Brammer, M., Bullmore, E., Simmons, A., Bartels, M., David, A.S., 2002. The neural correlates of intentional and incidental self processing. Neuropsychologia 40, 683–692.
Konukoglu, E., Coutu, J.-P., Salat, D.H., Fischl, B., A. S. D. N. Initiative, 2016. Multivariate statistical analysis of diffusion imaging parameters using partial least squares: Application to white matter variations in Alzheimer's disease. NeuroImage 134, 573–586.
Kramer, J.H., Miller, B.L., 2000. Alzheimer's disease and its focal variants. In: Seminars in Neurology, pp. 447–454.
Leow, A.D., Zhan, L., Zhu, S., Hageman, N., Chiang, M.-C., Barysheva, M., et al., 2009. White matter integrity measured by fractional anisotropy correlates poorly with actual individual fiber anisotropy. In: Biomedical Imaging: From Nano to Macro, 2009. ISBI'09, pp. 622–625 (IEEE International Symposium on 2009).
Lowe, V.J., Peller, P.J., Weigand, S.D., Quintero, C.M., Tosakulwong, N., Vemuri, P., et al., 2013. Application of the National institute on aging–Alzheimer's association AD criteria to ADNI. Neurology 80, 2130–2137.
Mahley, R.W., 1988. Apolipoprotein E: cholesterol transport protein with expanding role in cell biology. Science 240, 622–630.
Maier-Hein, K.H., Westin, C.-F., Shenton, M.E., Weiner, M.W., Raj, A., Thomann, P., et al., 2015. Widespread white matter degeneration preceding the onset of dementia. Alzheimer's Dementia 11, 485–493 e2.
Makris, N., Kennedy, D.N., McInerney, S., Sorensen, A.G., Wang, R., Caviness Jr., V.S., et al., 2004. Segmentation of subcomponents within the superior longitudinal fascicle in humans: a quantitative, in vivo, DT-MRI study. Cereb. Cortex 15, 854–869.
Malouin, F., Richards, C.L., Jackson, P.L., Dumas, F., Doyon, J., 2003. Brain activations during motor imagery of locomotor-related tasks: a PET study. Hum. Brain Mapp. 19, 47–62.
Murray, M.E., Graff-Radford, N.R., Ross, O.A., Petersen, R.C., Duara, R., Dickson, D.W., 2011. Neuropathologically defined subtypes of Alzheimer's disease with distinct clinical characteristics: a retrospective study. Lancet Neurol. 10, 785–796.
Noh, Y., Jeon, S., Lee, J.M., Seo, S.W., Kim, G.H., Cho, H., et al., 2014. Anatomical heterogeneity of Alzheimer disease based on cortical thickness on MRIs. Neurology 83, 1936–1944.
Ossenkoppele, R., Schonhaut, D.R., Schöll, M., Lockhart, S.N., Ayakta, N., Baker, S.L., et al., 2016. Tau PET patterns mirror clinical and neuroanatomical variability in Alzheimer's disease. Brain 139 (5), 1551–1567. https://doi.org/10.1093/brain/aww027.
Park, J.-Y., Na, H.K., Kim, S., Kim, H., Kim, H.J., Seo, S.W., et al., 2017. Robust Identification of Alzheimer's Disease subtypes based on cortical atrophy patterns. Sci. Rep. 7.
Ritchie, K., Touchon, J., 1992. Heterogeneity in senile dementia of the Alzheimer type: individual differences, progressive deterioration or clinical sub-types? J. Clin. Epidemiol. 45, 1391–1398.
Sachdev, P.S., Zhuang, L., Braidy, N., Wen, W., 2013. Is Alzheimer's a disease of the white matter? Curr. Opin. Psychiatry 26, 244–251.
Salat, D., Tuch, D., Van der Kouwe, A., Greve, D., Pappu, V., Lee, S., et al., 2010. White matter pathology isolates the hippocampal formation in Alzheimer's disease. Neurobiol. Aging 31, 244–256.
Scheltens, P., Fox, N., Barkhof, F., De Carli, C., 2002. Structural magnetic resonance imaging in the practical assessment of dementia: beyond exclusion. Lancet Neurol. 1, 13–21.
Scheltens, N.M., Galindo-Garre, F., Pijnenburg, Y.A., van der Vlies, A.E., Smits, L.L., Koene, T., et al., 2016. The identification of cognitive subtypes in Alzheimer's disease dementia using latent class analysis. J. Neurol. Neurosurg. Psychiatry 87 (3), 235–243. https://doi.org/10.1136/jnnp-2014-309582.
Scott, J.A., Tosun, D., Braskie, M.N., Maillard, P., Thompson, P.M., Weiner, M., et al., 2017. Independent value added by diffusion MRI for prediction of cognitive function in older adults. NeuroImage 14, 166–173.
Shallice, T., Fletcher, P., Frith, C., Grasby, P., Frackowiak, R., Dolan, R., 1994. Brain regions associated with acquisition and retrieval of verbal episodic memory. Nature 368, 633.
Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E., Johansen-Berg, H., et al., 2004. Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage 23, S208–S219.
Smith, E., Salat, D., Jeng, J., McCreary, C., Fischl, B., Schmahmann, J., et al., 2011. Correlations between MRI white matter lesion location and executive function and episodic memory. Neurology 76, 1492–1499.
Snowden, J.S., Stopford, C.L., Julien, C.L., Thompson, J.C., Davidson, Y., Gibbons, L., et al., 2007. Cognitive phenotypes in Alzheimer's disease and genetic risk. Cortex 43, 835–845.
Stanislav, K., Alexander, V., Maria, P., Evgenia, N., Boris, V., 2013. Anatomical characteristics of cingulate cortex and neuropsychological memory tests performance. Procedia Soc. Behav. Sci. 86, 128–133.
Stopford, C.L., Snowden, J.S., Thompson, J.C., Neary, D., 2008. Variability in cognitive presentation of Alzheimer's disease. Cortex 44, 185–195.
Tang-Wai, D.F., Graff-Radford, N., Boeve, B.F., Dickson, D.W., Parisi, J.E., Crook, R., et al., 2004. Clinical, genetic, and neuropathologic characteristics of posterior cortical atrophy. Neurology 63, 1168–1174.
Villemagne, V.L., Burnham, S., Bourgeat, P., Brown, B., Ellis, K.A., Salvado, O., et al., 2013. Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. Lancet Neurol. 12, 357–367.
Whitwell, J.L., Dickson, D.W., Murray, M.E., Weigand, S.D., Tosakulwong, N., Senjem, M.L., et al., 2012. Neuroimaging correlates of pathologically defined subtypes of Alzheimer's disease: a case-control study. Lancet Neurol. 11, 868–877.
Winklewski, P.J., Sabisz, A., Naumczyk, P., Jodzio, K., Szurowska, E., Szarmach, A., 2018. Understanding the physiopathology behind axial and radial diffusivity changes—-what do we know? Front. Neurol. 9, 92.
Wolk, D.A., Dickerson, B.C., Weiner, M., Aiello, M., Aisen, P., Albert, M.S., et al., 2010. Apolipoprotein E (APOE) genotype has dissociable effects on memory and attentional–executive network function in Alzheimer's disease. Proc. Natl. Acad. Sci. 107, 10256–10261.
Young, A.L., Oxtoby, N.P., Daga, P., Cash, D.M., Fox, N.C., Ourselin, S., et al., 2014. A data-driven model of biomarker changes in sporadic Alzheimer's disease. Brain 137, 2564–2577.
Zhang, X., Mormino, E.C., Sun, N., Sperling, R.A., Sabuncu, M.R., Yeo, B.T., et al., 2016. Bayesian model reveals latent atrophy factors with dissociable cognitive trajectories in Alzheimer's disease. Proc. Natl. Acad. Sci. 113, E6535–E6544.