



Published in final edited form as:

Brain Imaging Behav. 2020 October ; 14(5): 1792–1804. doi:10.1007/s11682-019-00115-6.

Detecting biological heterogeneity patterns in ADNI amnestic mild cognitive impairment based on volumetric MRI

Ali Ezzati^{1,2}, Andrea R. Zammit¹, Christian Habeck³, Charles B. Hall⁴, Richard B. Lipton^{1,2,4}, Alzheimer's Disease Neuroimaging Initiative

¹Department of Neurology, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, USA

²Department of Neurology, Montefiore Medical Center, Bronx, NY, USA

³Cognitive Neuroscience Division, Department of Neurology, Columbia University, New York, NY, USA

⁴Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, USA

Abstract

There is substantial biological heterogeneity among older adults with amnestic mild cognitive impairment (aMCI). We hypothesized that this heterogeneity can be detected solely based on volumetric MRI measures, which potentially have clinical implications and can improve our ability to predict clinical outcomes. We used latent class analysis (LCA) to identify subgroups among persons with aMCI ($n = 696$) enrolled in the Alzheimer's Disease Neuroimaging Initiative (ADNI), based on baseline volumetric MRI measures. We used volumetric measures of 10 different brain regions. The subgroups were validated with respect to demographics, cognitive performance, and other AD biomarkers. The subgroups were compared with each other and with normal and Alzheimer's disease (AD) groups with respect to baseline cognitive function and longitudinal rate of conversion. Four aMCI subgroups emerged with distinct MRI patterns: The first subgroup ($n = 404$), most similar to normal controls in volumetric characteristics and cognitive function, had the lowest incidence of AD. The second subgroup ($n = 230$) had the most similar MRI profile to early AD, along with poor performance in memory and executive function domains. The third subgroup ($n = 36$) had the highest global atrophy, very small hippocampus and worst overall cognitive performance. The fourth subgroup ($n = 26$) had the least amount of atrophy, however still had poor cognitive function specifically in the executive function domain.

Ali Ezzati, aezzati@einstein.yu.edu.

Conflict of interest None reported.

Compliance with ethical standards

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

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

Individuals with aMCI who were clinically categorized within one group other showed substantial heterogeneity based on MRI volumetric measures.

Keywords

Latent class analysis; Volumetric MRI; Cognitive function; Amnestic MCI; Alzheimer's disease

Introduction

Amnestic mild cognitive impairment (aMCI), an antecedent of Alzheimer's disease (AD), is sometimes conceptualized as a unitary entity; however, there is abundant evidence for heterogeneity (Bondi et al. 2014; Nettiksimmons et al. 2014). Individuals with aMCI differ in clinical course; some remit, others remain stable and many progress to AD, albeit at various rates (Nettiksimmons et al. 2014; Edmonds et al. 2015). Post-mortem studies demonstrate that aMCI has a variety of pathologic substrates (Dubois and Albert 2004). Identifying subgroups of aMCI may help clarify various pathways that lead to cognitive decline and ultimately identify groups that differ in prognosis and treatment response. For example, prior studies have indicated that aMCI individuals with smaller hippocampal volume have higher chance of conversion to AD (worse prognosis) (Devanand et al. 2007) and therefore might be a better target for interventional studies. Prior studies have used various approaches to characterizing subgroups of aMCI: Some studies have used cognitive tests which measure various domains to identify MCI subgroups (Edmonds et al. 2015; Eppig et al. 2017; Zammit et al. 2018a, b), while others have used a combination of neurocognitive tests and biological markers to detect such subgroups (Nettiksimmons et al. 2014).

The Alzheimer's Disease Neuroimaging Initiative (ADNI) recruits 3 distinct groups including normal controls, individuals with aMCI, and individuals with mild Alzheimer's disease (AD) (Mueller et al. 2005). This and other studies of similar design, attempt to identify and recruit clinically homogeneous groups at baseline. However, recent research has challenged the empirical validity of this conventional diagnostic approach by identifying considerable heterogeneity within each of these groups which might account for cognitive differences at baseline or during follow-up (Nettiksimmons et al. 2014; Edmonds et al. 2015; Eppig et al. 2017; Zammit et al. 2018c, d; Zammit et al. 2019). Although these approaches each have their own strengths, two major limitations might limit usage of the proposed criteria: (1) Data for all AD biomarkers, such as CSF proteins or PET, is not accessible in most clinical setting or even in many research studies; (2) Different research studies use diverse range of neurocognitive batteries, which makes harmonization and comparison of such data difficult. Beside that in real-world clinical settings due to limitations in time and resources, only a small subset of neurocognitive tests is obtained from each patient, which is typically inadequate to test all essential cognitive domains.

In the current study, we aimed to identify heterogeneous classes among the aMCI group in ADNI solely based on volumetric MRI measures. Structural MRIs are easily accessible for most patients and costs for the baseline MRI are covered by almost all national health

programs and other insurances. Furthermore, with the advances in automated analysis of MRIs, there is great potential in facilitating identification of these unique subgroups based on MRI measures. For the purpose of this study, we chose volumetric MRI data derived from FreeSurfer software, and used principal component analysis (PCA) to identify 10 regions of interests (ROIs). These 10 ROIs were entered into latent class analysis (LCA) models as predictors (see methods for details). We hypothesized that subgroups of individuals with MCI, identified by applying LCA to volumetric MRI data, would differ at cross-section in features not included in the model and longitudinally in rates of progression to dementia.

Methods

Participants

The data used for this analysis were downloaded from the ADNI database (www.adni.loni.usc.edu) in March 2018. The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The individuals included in the current study were recruited as part of ADNI-1, ADNI-GO, and ADNI-2 between September 2005, and December 2013. This study was approved by the Institutional Review Boards of all participating institutions. Informed written consent was obtained from all participants at each site. For up-to-date information, see www.adni.loni.usc.edu.

Eligible individuals completed baseline MRI and had automated volumetric analysis for all the ROIs, were included in the analysis. Participants whose scans failed to meet quality control or had unsuccessful automated image analysis were excluded from this study. A total of 1394 participants from ADNI-1, ADNI-GO, and ADNI-2 were eligible for this analysis. All enrolled ADNI MCI participants were diagnosed as having amnesic MCI; this diagnostic classification required Mini Mental State Examination (MMSE) scores between 24 and 30 (inclusive), a memory complaint, objective memory loss measured by education-adjusted scores on the Wechsler Memory Scale Logical Memory II, a Clinical Dementia Rating (CDR) of 0.5, absence of significant impairment in other cognitive domains, essentially preserved activities of daily living, and absence of dementia. The subjects with AD had to satisfy the National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria for probable AD, and have MMSE scores between 20 and 26 (inclusive), and CDR of 0.5 or 1. As part of study design, CN group were frequency matched to MCI and AD participants by age group. The cognitively normal (CN), MCI, and mild AD groups included in current study comprised of 449, 696, and 249 individuals, respectively.

MRI acquisition and measurements

A subset of MRIs for ADNI-1 was obtained using 1.5 T scanners and the rest of ADNI-1 as well as all participants of ADNI-GO and ADNI-2 completed MRIs using 3 T scanners across different sites of ADNI study (For more information please see www.adni.loni.usc.edu). MRI data were automatically processed using the FreeSurfer software package (version 4.3 and 5.1, available at <http://surfer.nmr.mgh.harvard.edu/>) by the Schuff and Tosun laboratory at the University of California-San Francisco as part of the

ADNI shared data-set. FreeSurfer methods for identifying and calculation of regional brain volume are previously described in detail (Fischl et al. 2002). For the purpose of this study, volume of all regions of interest (ROIs) were normalized for total brain volume and the ratio of ROIs' volume (ROI_v) to TBV [i.e. (ROI_v/TBV) x mean whole population ROI_v] was used in the analyses and reported throughout manuscript unless otherwise specified. White matter hyperintensity (WMH) volumes were calculated by 2 different automated methods for ADNI 1 and ADNI 2 by the Imaging of Dementia and Aging (IDeA) laboratory at the University of California–Davis. WMH volumes for ADNI 1 were derived from coregistered T1-, T2-, and proton density (PD)-weighted images (Schwarz et al. 2009) and for ADNI 2 they were derived from coregistered T1 and FLAIR images (DeCarli et al. 2013).

External validators

External validators are variables that have not been used in the class formation to determine if the classes are distinguishable on pre-existing characteristics and help in interpreting the results (Nylund et al. 2007a; Nagin and Tremblay 2005). In addition to demographic information, external validator variables included the following:

- APOE gene status: Apolipoprotein E (APOE) e4 allele frequency was accessible for 99.7% of MCI participants ($n = 694$) and included in the current study as a genetic marker of AD.
- CSF markers: CSF samples were batch processed by the ADNI Biomarker Core at the University of Pennsylvania School of Medicine (Shaw 2008). These data were available for 71.1% of the whole sample ($n = 991$) and 72.3% of MCI ($n = 503$). Individuals were classified according to CSF concentration thresholds (τ : >93 pg/mL; p- τ_{181p} : >23 pg/mL; $A\beta_{1-42} < 192$ pg/mL; p- $\tau_{181p}/A\beta_{1-42}$ ratio: >0.10) previously established to maximize sensitivity and specificity of autopsy confirmed AD (Shaw et al. 2009).
- Baseline FDG-PET was processed by the Jagust lab at the University of California-Berkeley and Lawrence Berkeley National Laboratory (Landau et al. 2011). An average of 18FDG-PET value for angular gyrus, temporal lobe, and posterior cingulate regions was selected as the variable of interest for this study. FDG-PET data were available for 77.2% of MCI group ($n = 537$), and 75.4% of the whole sample ($n = 1051$).
- Cognitive tests included Digit Span forward and backward, Digit Symbol Substitution, Trails B, Category Naming (sum of animal and vegetable scores), the Mini Mental State Examination (MMSE), Alzheimer's Disease Assessment Scale cognitive subscale (ADAS-cog), the Rey's Auditory Vocabulary List Test (RAVLT), and Logical Memory II (Folstein et al. 1975; Mirra et al. 1991; Partington and Leiter 1949; Rey 1958; Wechsler 2014; Mohs 1994). In addition, composite scores for the memory and executive function domains were used as separate external validators (Gibbons et al. 2012; Crane et al. 2012).
- Clinical outcomes: Longitudinal clinical outcomes were available on 94.0% of MCI participants ($n = 654$), with an average follow-up of 41.5 months. Variables used for the purpose of this study included the type of clinical conversion

(progression to dementia or reversion to normal) and the associated number of months to conversion.

Some of the external validator variables were not available for all participants for various sampling or quality related reasons, or because they were added at a later stage to the study.

Statistical analysis

Principal factor analysis of ROIs—Automatic parcellation of brain by FreeSurfer leads to more than 40 region-specific volume estimates. Considering the high correlation between values of some of these regions, inclusion of all of these variables in latent class analysis (LCA) might lead to a unidimensional pattern (i.e. identified classes would be based on the average of indicators and not the pattern of indicators). In addition, high number of indicators in LCA lead to complexity of interpretation. Therefore, as the first step we performed the PCA on all of the ROIs provided by FreeSurfer, which were implicated in aging, MCI, and AD based on prior studies (McDonald et al. 2009; De Jong et al. 2008). The factor analysis resulted in 10 distinct groups of ROIs (Table 1). From each group the ROI with highest loading was selected as the indicator for entry into the LCA model. The final 10 indicators included in the LCA models were hippocampus, middle Temporal, superior temporal, precuneus, anterior cingulate, medial orbitofrontal, frontal operculum, precentral, lingual, and caudate regions.

Latent class analysis of ROIs—LCA provides a unique opportunity to group individuals on the basis of continuously distributed indicators such as the volume of a region of interest, not based on (and therefore not biased by) artificial cutoffs or long-term trajectories. LCA is expected to identify subgroups comprised on individuals within each class that are as similar to each other as possible and classes which differ as much as possible. LCA uses a step-wise procedure and a variety of fit indices to determine whether the addition of classes improves the fit to the data (Beauchaine 2003). Thus, there is no gold standard regarding the number of participants and power for a proposed LCA. First, a one-class (unconditional) model is fit to the data. Next, two, three, ..., or N class solutions are applied until the model does not improve with the addition of extra classes (Nylund et al. 2007a). Several goodness of fit statistics are used to determine an optimal model, including Bayesian information criterion (BIC) (Raftery 1995); and entropy (Celeux and Soromenho 1996). Entropy identifies the solution with the best precision in distinguishing amongst classes. Monte Carlo simulation studies using a variety of sample sizes suggest that the BIC is the most robust fit index and thus were given the most weight in class selection (Nylund et al. 2007b). To obtain appropriate model convergence and a robust solution we applied 500 random starts in the initial stage and 40 optimizations in the final stage. Of note, demographics including age and sex, were not included in the LCA models as we did not want demographics to adjust, define, or contribute to, class membership.

Assessment of baseline and longitudinal differences—*Analysis of variance (ANOVA) and Chi-square tests for categorical variables* were used to compare demographics differences between LCA derived MCI subgroups. We used multiple analyses of covariance (ANCOVA, general linear model) with baseline and longitudinal characteristics as dependent variable, group membership as the independent variable, and age, gender,

education, and MRI field strength (categorical variable) as covariates. Age, gender and education are demographics known to affect cognitive performance and level of biomarker, therefore were added to the models as covariates. In addition, as mentioned above, the MRI protocol for ADNI1 (2004–2009) focused on consistent longitudinal structural imaging on 1.5 T scanners using T1-weighted sequences. While in ADNI-GO/ADNI2 (2010–2016), imaging was performed at 3 T with T1-weighted imaging parameters similar to ADNI1. Although FreeSurfer morphometric procedures have been demonstrated to show good test-retest reliability across scanner manufacturers and across field strengths (Dubois and Albert 2004; Devanand et al. 2007), we included MRI field strength as a covariate in ANCOVAs to account for potential effects of different field strength on outcomes. For the purpose of longitudinal analysis, and to account for the variation in duration of follow up in each MCI subgroup, we calculated incidence of AD for each MCI subgroup with the person-year approach (Armitage et al. 2008) by dividing the number of cases by the person-years at each subgroup. We multiplied these rates by 100 to report rates per 100 person-years. The number of person-years contributed by a participant who did not develop AD was the time between baseline examination to the last follow-up examination. The number of person-years contributed by a participant with AD was the time from baseline to the first visit that the participant was diagnosed with AD. Statistical significance was set at $\alpha = 0.01$ to control for Type-I errors.

For LCA modeling, we used MPlus version 8 (Muthén & Muthén, 1998–2017). All other statistical analyses were conducted using SPSS, version 25 (Chicago, IL: SPSS Inc.).

Results

Formation and selection of the optimal class solution

Model fit statistics for 2–6 class models are presented in Table 2. The fit improved markedly with every added class, up until the 4-class solution (BIC = 14,685), with an entropy of 82%. With 5 and 6-class solutions, BIC did not improve and entropy declined. Therefore, we judged that the 4-class solution was optimal.

Characteristics of the study population

Table 3 shows the baseline characteristics of the CN, MCI (separately for each class identified by LCA), and AD participants. The average age of the whole MCI group in this study was 72.5 years (SD = 7.4) and 58.2% were men. The subgroups differed on age, with MCI-1 and MCI-2 being markedly younger than the other two subgroups. MCI-2 (67.4%) and MCI-3 (77.8%) were predominantly male, while MCI-4 had the highest percentage of females (73.1%). MCI-1 also had slightly higher levels of education and lower CDR score in comparison with MCI-2. Subgroups were not different in terms of APOE status.

Baseline differences in demographics, volumetric MRI and cognitive function

Table 4 summarizes the difference between external validators (APOE status, neurocognitive battery, CSF biomarkers, and imaging biomarkers) for each of the 4 latent MCI classes as well as CN and AD groups.

MCI subgroups were created using baseline MRI measures and consequently showed significant differences in volumes of different ROIs at baseline. Table 5 shows volumes of ROIs included in LCA models for each subgroup. Each subgroup can be summarized as follows based on MRI measures: MCI-1 was the largest class (58% of whole MCI population). In comparison with other MCI subgroups, MCI-1 had the largest hippocampus (a medial temporal lobe structure). MCI-1 also had larger middle temporal and superior temporal lobules (lateral temporal lobe structures) in comparison with MCI-2 and MCI-3 subgroups. In addition, comparing to CN, the MCI-1 subgroup had larger volumetric measures in superior temporal, frontal operculum, medial orbitofrontal, anterior cingulate, precuneus, precentral, and lingual regions (Fig. 1). MCI-2, the second largest subgroup (33%) had the least healthy volumetric MRI profile, which was often near or beyond the averages in the AD group (Fig. 1). In comparison with MCI-1, MCI-2 had significantly smaller volumes in all ROIs except the caudate (Table 5). MCI-3, a small subgroup (5.2%), similar to MCI-2 had a poor volumetric profile, worse than averages in the AD group for some ROIs. MCI-3 had smaller volumes in all ROIs in comparison with MCI-1, except in caudate, which was significantly larger than MCI-1. Finally, MCI-4 was the smallest subgroup (3.7%). This subgroup had the smallest hippocampal volume among subgroups (significantly different only from MCI-1). However, MCI-4 had a healthy volumetric profile in most ROIs included in LCA models (Fig. 1 and Table 5).

Comparison of cognitive scores revealed that MCI-1 had the best overall cognitive scores, while MCI 2 and 3 had the most severe memory and executive function impairment, and MCI-4 had executive function impairment (Table 4 and Fig. 2). MCI-1 was significantly better than all other subgroups in executive function and had better memory in comparison with MCI-2 and MCI-3. MCI-3 had the worst scores on most cognitive scores among all subgroups, though in most cases they were not significantly different from MCI-2 and MCI-4, likely due to the small subgroup size and low statistical power.

Baseline differences in other AD biomarkers

MCI-1 had the least amount of brain atrophy among subgroups as evident by total brain volume (TBV) and ventricles volume (VV). MCI-3 had the smallest TBV and largest VV among all groups. MCI-2 had the least amount of white matter hyperintensities in comparison with other MCI subgroups and even in comparison with CN group. MCI-4, by far, had the worst WMHs portfolio among MCI subgroups (Table 3).

MCI-1 also had higher average FDG-PET of angular, temporal, and posterior cingulate in comparison with MCI-2 and MCI-3 subgroups. Comparison of CSF biomarkers showed that a significantly higher proportion of MCI-1 and MCI-4 subgroups had higher p-tau and P-tau/A β ratio than the other two subgroups (Table 4 & Fig. 3). Repeating the analysis for CSF biomarkers based on mean CSF biomarker concentrations, failed to show any significant difference between subgroups (results not shown).

Longitudinal conversion to AD and reversion to CN

Longitudinal clinical outcomes were available for 94% of the MCI participants included in this study, and all subgroups had at least one wave of follow up in more than 90% of cases.

However, duration of follow up in subgroups were not similar (MCI-1 = 2.4, MCI-2 = 2.1, MCI-3 = 1.5, and MCI-4 = 2.7). Of the total of 697 participants with MCI at baseline, 252 individuals developed AD during 2570 person-years of follow up resulting in an overall incidence rate of 9.8 per 100 person-years. MCI-1 had the lowest conversion rate to dementia, with an incidence rate of 7.7 per 100 person-years. Incidence rate of AD in MCI-2, MCI-3, and MCI-4 were 12.3, 16.4, and 12.4 per 100 person-years, respectively. Incidence rate of AD in MCI-2, MCI-3, and MCI-4 were significantly higher than group 1, but they were not significantly different from each other. In addition, only MCI-1 and MCI-2 subgroups had individuals who had reversion back to CN with an incidence rate of 1.9 (MCI-1) and 0.9 (MCI-2) per 100 person-years.

Discussion

We used LCA across 10 distinct cortical and subcortical MRI regional volumes to identify unique, empirically derived subgroups of amnesic MCI within the ADNI. The optimal solution revealed 4 subgroups, which differed on volumetric MRI measures. External measures of cognitive function and other biological markers also distinguished amongst the four subgroups. These results revealed some noteworthy findings. First, there is biological heterogeneity based on MRI at baseline among ADNI with aMCI participants based on volumetric MRI measures. Second, a large subgroup (58%) of participants (MCI-1) had better cognitive function at baseline and lower incidence rate of AD and was characterized by a biomarker profile much better in comparison with other subgroups implying that current MCI diagnoses might be overclassifying some individuals as cognitively impaired.

We found that the MCI-1 subgroup, in comparison with normal controls, had similar brain atrophy profile based on TBV and VV, and preserved volumes in most ROIs that were used as indicators in LCA models (except hippocampus). The MCI-1, despite having the significantly better memory in comparison with MCI-2 and MCI-3, and better executive function performance than the rest of MCI subgroups, on average had pronounced memory deficits and were distinguishable from normal controls. However, considering the higher rate of reversion to normal in comparison with all other subgroups, it is possible that a subset of subjects included in this subgroup had a false positive diagnosis of MCI as suggested by other studies (Bondi et al. 2014; Edmonds et al. 2015). Considering higher level of education (a marker for cognitive reserve) and the small extent of brain atrophy as evident by preserved TBV (an indicator of brain reserve and neurodegeneration), It is also possible that this group has higher cognitive and brain reserve compared to the rest of the classes (Stern 2002). Our data does not fully support the idea that this subgroup's memory impairment was only due to AD pathology, primarily because they had the lowest incidence rate of AD despite the higher CSF P-tau and P-tau/A β ratio in comparison with other groups. These findings, suggest that participants in MCI-1 subgroup might belong to a category of MCI participants previously described as "stable" amnesic MCI (Whitwell et al. 2008): the stable amnesic MCI participants had significantly smaller hippocampal volume than the normal controls, although there were no significant differences in gray matter loss. MCI-1 subgroup might have a mixed pathology, comprising a mix of hippocampal sclerosis (Zarow et al. 2008) and AD pathology.

MCI-2 and MCI-3 subgroups were very similar to each other in terms of MRI indicators, having evidence of atrophy in all ROIs studied with caudate and precentral area volumes being the only indicators differentiating the two subgroups in LCA models. However, MCI-3 had the highest rate of atrophy as evident by having smallest TBV and largest ventricle volumes. Another imaging finding differentiating these 2 subgroups was WMHs. While MCI-2 had very low levels of WMHs, even lower than normal controls, MCI-3 had significantly higher WMHs. MCI-2 and MCI-3 had the best CSF biomarker profile among subgroups with lower P-tau and P-tau/A β ratio, but a higher rate of incident AD. MCI-2 and MCI-3 had the worst performance on most memory and executive function tests in comparison with the other 2 subgroups; however, the difference between MCI-2, MCI-3 and MCI-4 was only a non-significant trend. This is likely due to our use of $\alpha = 0.01$ significance level to adjust for multiple comparisons and high variability of the tests in smaller subgroup (MCI-3), which produced a large standard error. These two subgroups might also differ on caudate dependent tasks (i.e. motor tasks) (Grahn et al. 2008), however due to lack of such tests in ADNI data-set, such comparison was not possible. Considering all these findings, it is plausible that both MCI-2 and MCI-3 represent a subgroup previously described as “progressive” amnesic MCI (Whitwell et al. 2008), with higher rates of Incident AD. The major difference between these subgroups is likely the pathophysiology, with one being predominantly AD (MCI-2) and the other possibly a mix of AD pathology and cerebrovascular disease (MCI-3).

MCI-4 represents the smallest subgroup in our LCA models. In comparison with other MCI subgroups and even CN group, MCI-4 had the largest volume in most cortical regions included in LCA. However, they had the very small hippocampus and very high cerebrovascular disease as evident by WMHs. CSF profile of this subgroup was also comparable with MCI-1 and slightly worse than the other two subgroups. MCI-4, similar to MCI-2 and MCI-3, had significant executive function impairment in comparison with MCI-1, with slightly better (but not significant) memory performance in comparison with MCI-2 and 3. Despite older age, substantial WMH, and 65.3% of its individuals having high tau/A β ratio, individuals in this group were still performing better than MCI-2 and MCI-3, indicating their ability to withstand more pathology potentially due to less global atrophy and likely higher brain reserve. It may be due to this reserve capacity that their cognitive function is preserved for a longer period, and why despite their MCI status, they are still performing relatively better than MCI 2 and 3 subgroups. Considering the small size of this subgroup, these results should be interpreted with some caution. Overall, our results indicate that MCI-4 is a distinct, yet very small subgroup, with clinical and biological characteristics, which are substantially different from the other subgroups, possibly with a mixed AD with predominant cerebrovascular pattern.

In general, among MCI subgroups, MCI-1 had the best cognitive scores and outcome, MCI 2 and 3 had the worst and very close to one another, and MCI-4 had impairment in executive function domain. This might suggest that all of the MCI classes are on the same path to AD, with some simply farther along the path. However, comparison of imaging and CSF biomarkers suggests that this is not the case. For example, MCI-1 and MCI-4 had higher p-tau and P-tau/A β ratio than the other two groups; MCI-4 had the highest WMHs and MCI-2 had the lowest WMHs among groups.

A few studies have previously shown heterogeneity in the ADNI MCI group using cluster analysis (Bondi et al. 2014; Nettiksimmons et al. 2014; Eppig et al. 2017). Although these studies used somewhat different analytical approaches and different indicators (i.e. neuropsychological data alone (Bondi et al. 2014; Eppig et al. 2017) or a mixture of biomarkers (Nettiksimmons et al. 2014), all of them found 4 subgroups in the MCI sample. Due to differences in methods, indicators and size of samples, a direct comparison of the present study with previous studies is not possible though there are some similarities. For example, all studies have a subgroup of individuals with MCI who are similar to normal controls and have a very low conversion rate to dementia in longitudinal follow up. There is another subgroup with a profile similar to AD; that subgroup has an elevated incidence of AD. Similar to our findings, Nettiksimmons et al. (Nettiksimmons et al. 2014) have also found four groups differentiated based on MRI, CSF and blood biomarkers, with one subgroup representing stable MCI (with smaller than normal hippocampal volume and insignificant differences in other gray matter volumes), one group with very high rate of conversion to AD, and at least one group with significant volume loss (TBV and hippocampal volume) and high WMH possibly representing the mixed pathology subgroup. Further studies with different numbers or types of indicators are needed to identify the optimal number and types of indicators that can help with improving the classification accuracy in population.

Perhaps the most important clinical implication of this work is the evidence that there is substantial biological heterogeneity among amnesic MCI subjects. This is particularly important because based on the ADNI study design, these subjects were intentionally selected to represent one clinical phenotype, presumed to be a precursor to AD (Morris and Cummings 2005). Similar study designs have been used in many AD clinical trials which have targeted subjects at various prodromal or early stages of AD (Schneider et al. 2014). Despite substantial efforts and use of various methods to develop interventions to promote cognitive health, results have been disappointing. (Willis et al. 2006; Ball et al. 2002; Godyn et al. 2016). A major reason for failure of these studies is potentially the heterogeneity in population of study, which leads to inferior subject selection criteria and low statistical power to detect meaningful effect of interventions. Considering the results of this study and other studies (Nettiksimmons et al. 2014; Eppig et al. 2017; Leoutsakos et al. 2015), which have used different approaches and indicators to detect subgroups in prodromal stages of AD, it can be concluded that there is a substantial heterogeneity in each pre-clinical stage of AD and therefore using a one dimensional classification (i.e. amnesic MCI class based on neuropsychological evaluation or subjective complaints) might not be the best approach to select subjects for an interventional study.

One strength of this study is the unsupervised latent class profiling method without using any cut-offs. All the indicators included in the model were derived only from a single MRI session, and only one MRI modality, which is potentially easily available for all patients being evaluated for concerns about their memory function. In addition, subgroup membership was not affected by cognitive function at baseline, cognitive outcomes, or other biomarkers. However, several limitations should be noted. First, ADNI is not a population-based study and there are strict inclusion and exclusion criteria for selection of subjects. Therefore, this study should be considered as an exploratory study and replication of this

work in a population-based cohort with large MRI data sets and uniform image acquisition techniques might provide better insight about the heterogeneity in the amnesic MCI group. In addition, some of the subgroups of this study were small, and therefore the true differences between these subgroups and others possibly were not detected in some cases. Using PCA to select indicators for the LCA is a new approach and might not be a perfect solution, however this approach can effectively reduce overlap between indicators and improve possibility of pattern recognition. Further studies are needed to investigate how to narrow down and select the best set of indicators for latent class modeling when there is a large pool of potentially informative indicators. Finally, it should be mentioned that the classes detected by LCA had a significant overlap and were not completely distinct. This might indicate that using clinical data (e.g. neuropsychological tests) and additional biomarkers (e.g. CSF biomarkers or PET measures, might improve classification. While it might seem easier to use all available data to identify heterogeneity, in practice it is important to know the separate and joint effect of each modality of data. For example, neuropsychological tests are relatively obtained easier -and are usually inexpensive-however there is no consensus on the “best neuropsychological tests” and each clinician, cohort or trial obtains a different subset of tests. Therefore, models developed in one sample are rarely applicable to other samples without a rigorous harmonization process. On the other hand, MRIs or PET scans provide objective measures that are more reliable and relatively stable across sites and studies, however they are more expensive and not always feasible. Recently our group has used latent class modeling and neuropsychological data to parse heterogeneity in 2 large cohorts of older adults (Zammit et al. 2018c, d; Zammit et al. 2019), and in this paper we showed the effectiveness of using MRI volumetrics regardless of cognitive scores. In future studies we plan to assess if using cognitive score and MRI measures (including but not limited to volumetrics) simultaneously in latent class models can help with identifying more homogenous subgroups.

In conclusion, our findings revealed biological heterogeneity based on volumetric MRI in the amnesic MCI group in ADNI. Characterization of these heterogeneous subgroups indicated that there are further substantial differences in cognitive function, cognitive outcomes, and AD biomarkers amongst subgroups.

Acknowledgements

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 https://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

Funding Data collection and sharing for ADNI project was funded by the Alzheimer’s Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer’s Association; Alzheimer’s Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd. and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health

Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Authors of this study were supported in part by National Institutes of Health grants NIA 2 P01 AG03949, NIA 1R01AG039409-01, NIA R03 AG045474, NIH K01AG054700, the Leonard and Sylvia Marx Foundation, and the Czap Foundation.

References

- Armitage P, Berry G, Matthews JNS (2008). *Statistical methods in medical research*. John Wiley & Sons.
- Ball K, Berch DB, Helmers KF, Jobe JB, Leveck MD, Marsiske M, Morris JN, Rebok GW, Smith DM, Tennstedt SL, Unverzagt FW, Willis SL, & ACTIVE Study Group. (2002). Effects of cognitive training interventions with older adults: A randomized controlled trial. *JAMA*, 288, 2271–2281. [PubMed: 12425704]
- Beauchaine TP (2003). Taxometrics and developmental psychopathology. *Development and Psychopathology*, 15, 501–527. [PubMed: 14582930]
- Bondi MW, Edmonds EC, Jak AJ, Clark LR, Delano-Wood L, McDonald CR, Nation DA, Libon DJ, Au R, Galasko D, & Salmon DP (2014). Neuropsychological criteria for mild cognitive impairment improves diagnostic precision, biomarker associations, and progression rates. *Journal of Alzheimer's Disease*, 42, 275–289.
- Celeux G, & Soromenho G (1996). An entropy criterion for assessing the number of clusters in a mixture model. *Journal of Classification*, 13, 195–212.
- Crane PK, Carle A, Gibbons LE, Insel P, Mackin RS, Gross A, et al. (2012). Development and assessment of a composite score for memory in the Alzheimer's Disease Neuroimaging Initiative (ADNI). *Brain Imaging and Behavior*, 6, 502–516. [PubMed: 22782295]
- De Jong L, Van der Hiele K, Veer I, Houwing J, Westendorp R, Bollen E, et al. (2008). Strongly reduced volumes of putamen and thalamus in Alzheimer's disease: An MRI study. *Brain.*, 131, 3277–3285. [PubMed: 19022861]
- DeCarli C, Maillard P, Fletcher E (2013). Four tissue segmentation in ADNI II Alzheimer's Disease Neuroimaging Initiative. Department of Neurology and Center for Neuroscience, University of California at Davis.
- Devanand D, Pradhaban G, Liu X, Khandji A, De Santi S, Segal S, et al. (2007). Hippocampal and entorhinal atrophy in mild cognitive impairment prediction of Alzheimer disease. *Neurology.*, 68, 828–836. [PubMed: 17353470]
- Dubois B, & Albert ML (2004). Amnesic MCI or prodromal Alzheimer's disease? *The Lancet Neurology.*, 3, 246–248. [PubMed: 15039037]
- Edmonds EC, Delano-Wood L, Clark LR, Jak AJ, Nation DA, McDonald CR, Libon DJ, Au R, Galasko D, Salmon DP, Bondi MW, & Alzheimer's Disease Neuroimaging Initiative. (2015). Susceptibility of the conventional criteria for mild cognitive impairment to false-positive diagnostic errors. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, 11, 415–424.
- Eppig JS, Edmonds EC, Campbell L, Sanderson-Cimino M, Delano-Wood L, Bondi MW, & for the Alzheimer's Disease Neuroimaging Initiative. (2017). Statistically derived subtypes and associations with cerebrospinal fluid and genetic biomarkers in mild cognitive impairment: A latent profile analysis. *Journal of the International Neuropsychological Society*, 23, 564–576. [PubMed: 28578726]
- Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, van der Kouwe A, Killiany R, Kennedy D, Klaveness S, Montillo A, Makris N, Rosen B, & Dale AM (2002). Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33, 341–355. [PubMed: 11832223]

- Folstein MF, Folstein SE, & McHugh PR (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189–198. [PubMed: 1202204]
- Gibbons LE, Carle AC, Mackin RS, 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 Imaging and Behavior*, 6, 517–527. [PubMed: 22644789]
- Godyn J, Jonczyk J, Panek D, & Malawska B (2016). Therapeutic strategies for Alzheimer's disease in clinical trials. *Pharmacological Reports*, 68, 127–138. [PubMed: 26721364]
- Grahn JA, Parkinson JA, & Owen AM (2008). The cognitive functions of the caudate nucleus. *Progress in Neurobiology*, 86, 141–155. [PubMed: 18824075]
- Landau SM, Harvey D, Madison CM, Koeppe RA, Reiman EM, Foster NL, Weiner MW, Jagust WJ, & Alzheimer's Disease Neuroimaging Initiative. (2011). Associations between cognitive, functional, and FDG-PET measures of decline in AD and MCI. *Neurobiology of Aging*, 32, 1207–1218. [PubMed: 19660834]
- Leoutsakos J-MS, Forrester SN, Lyketsos C, & Smith GS (2015). Latent classes of neuropsychiatric symptoms in NACC controls and conversion to mild cognitive impairment or dementia. *Journal of Alzheimer's Disease*, 48, 483–493.
- McDonald C, McEvoy L, Gharapetian L, Fennema-Notestine C, Hagler D, Holland D, et al. (2009). Regional rates of neocortical atrophy from normal aging to early Alzheimer disease. *Neurology*, 73, 457–465. [PubMed: 19667321]
- Mirra SS, Heyman A, McKeel D, Sumi S, Crain BJ, Brownlee L, et al. (1991). The consortium to establish a registry for Alzheimer's disease (CERAD) part II. Standardization of the neuropathologic assessment of Alzheimer's disease. *Neurology*, 41, 479–486. [PubMed: 2011243]
- Mohs R (1994). Administration and scoring manual for the Alzheimer's disease assessment scale, 1994 revised edition. New York: The Mount Sinai School of Medicine.
- Morris JC, & Cummings J (2005). Mild cognitive impairment (MCI) represents early-stage Alzheimer's disease. *Journal of Alzheimer's Disease*, 7, 235–239.
- Mueller SG, Weiner MW, Thal LJ, Petersen RC, Jack CR, Jagust W, Trojanowski JQ, Toga AW, & Beckett L (2005). Ways toward an early diagnosis in Alzheimer's disease: The Alzheimer's Disease Neuroimaging Initiative (ADNI). *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, 1, 55–66.
- Nagin DS, & Tremblay RE (2005). What has been learned from group-based trajectory modeling? Examples from physical aggression and other problem behaviors. *The Annals of the American Academy of Political and Social Science*, 602, 82–117.
- Nettiksimmons J, DeCarli C, Landau S, & Beckett L (2014). Biological heterogeneity in ADNI amnesic mild cognitive impairment. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, 10, 511–21.e1.
- Nylund KL, Asparouhov T, & Muthén BO (2007a). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, 14, 535–569.
- Nylund K, Bellmore A, Nishina A, & Graham S (2007b). Subtypes, severity, and structural stability of peer victimization: What does latent class analysis say? *Child Development*, 78, 1706–1722. [PubMed: 17988316]
- Partington JE, Leiter RG (1949). Partington's pathways test. *Psychological Service Center Journal*.
- Raftery AE (1995). Bayesian model selection in social research. *Sociological Methodology*, 25, 111–163.
- Rey A (1958). L'examen clinique en psychologie.
- Schneider LS, Mangialasche F, Andreasen N, Feldman H, Giacobini E, Jones R, Mantua V, Mecocci P, Pani L, Winblad B, & Kivipelto M (2014). Clinical trials and late-stage drug development for Alzheimer's disease: An appraisal from 1984 to 2014. *Journal of Internal Medicine*, 275, 251–283. [PubMed: 24605808]

- Schwarz C, Fletcher E, DeCarli C, Carmichael O (2009). Fully-automated white matter hyperintensity detection with anatomical prior knowledge and without FLAIR. *International conference on information processing in medical imaging*: Springer; p. 239–51.
- Shaw LM (2008). PENN biomarker core of the Alzheimer’s disease neuroimaging initiative. *Neurosignals.*, 16, 19–23. [PubMed: 18097156]
- Shaw LM, Vanderstichele H, Knapik-Czajka M, Clark CM, Aisen PS, Petersen RC, Blennow K, Soares H, Simon A, Lewczuk P, Dean R, Siemers E, Potter W, Lee VMY, Trojanowski JQ, & Alzheimer’s Disease Neuroimaging Initiative. (2009). Cerebrospinal fluid biomarker signature in Alzheimer’s disease neuroimaging initiative subjects. *Annals of Neurology*, 65, 403–413. [PubMed: 19296504]
- Stern Y (2002). What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society*, 8, 448–460. [PubMed: 11939702]
- Wechsler D (2014). Wechsler adult intelligence scale–fourth edition (WAIS–IV). San Antonio: Psychological Corporation.
- Whitwell JL, Shiung MM, Przybelski S, Weigand SD, Knopman DS, Boeve BF, et al. (2008). MRI patterns of atrophy associated with progression to AD in amnesic mild cognitive impairment. *Neurology.*, 70, 512–520. [PubMed: 17898323]
- Willis SL, Tennstedt SL, Marsiske M, Ball K, Elias J, Koepke KM, Morris JN, Rebok GW, Unverzagt FW, Stoddard AM, Wright E, & ACTIVE Study Group. (2006). Long-term effects of cognitive training on everyday functional outcomes in older adults. *JAMA*, 296, 2805–2814. [PubMed: 17179457]
- Zammit AR, Muniz-Terrera G, Katz MJ, Hall CB, Ezzati A, Bennett DA, et al. (2018a). Subtypes based on neuropsychological performance predict incident dementia: Findings from the rush memory and aging project. *Journal of Alzheimer’s Disease*, 1–11.
- Zammit AR, Hall CB, Katz MJ, Muniz-Terrera G, Ezzati A, Bennett DA, et al. (2018b). Class-specific incidence of all-cause dementia and Alzheimer’s disease: A latent class approach. *Journal of Alzheimer’s Disease*, 1–11.
- Zammit AR, Hall CB, Katz MJ, Terrera GM, & Lipton RB (2018c). Process of identifying A clinical model of latent neuropsychological profiles in community-dwelling older adults. *Alzheimer’s & Dementia: The Journal of the Alzheimer’s Association.*, 14, P1298–P12P9.
- Zammit AR, Hall CB, Lipton RB, Katz MJ, & Muniz-Terrera G (2018d). Identification of heterogeneous cognitive subgroups in community-dwelling older adults: A latent class analysis of the Einstein aging study. *Journal of the International Neuropsychological Society*, 24, 511–523. [PubMed: 29317003]
- Zarow C, Sitzer TE, & Chui HC (2008). Understanding hippocampal sclerosis in the elderly: Epidemiology, characterization, and diagnostic issues. *Current Neurology and Neuroscience Reports*, 8, 363–370. [PubMed: 18713571]

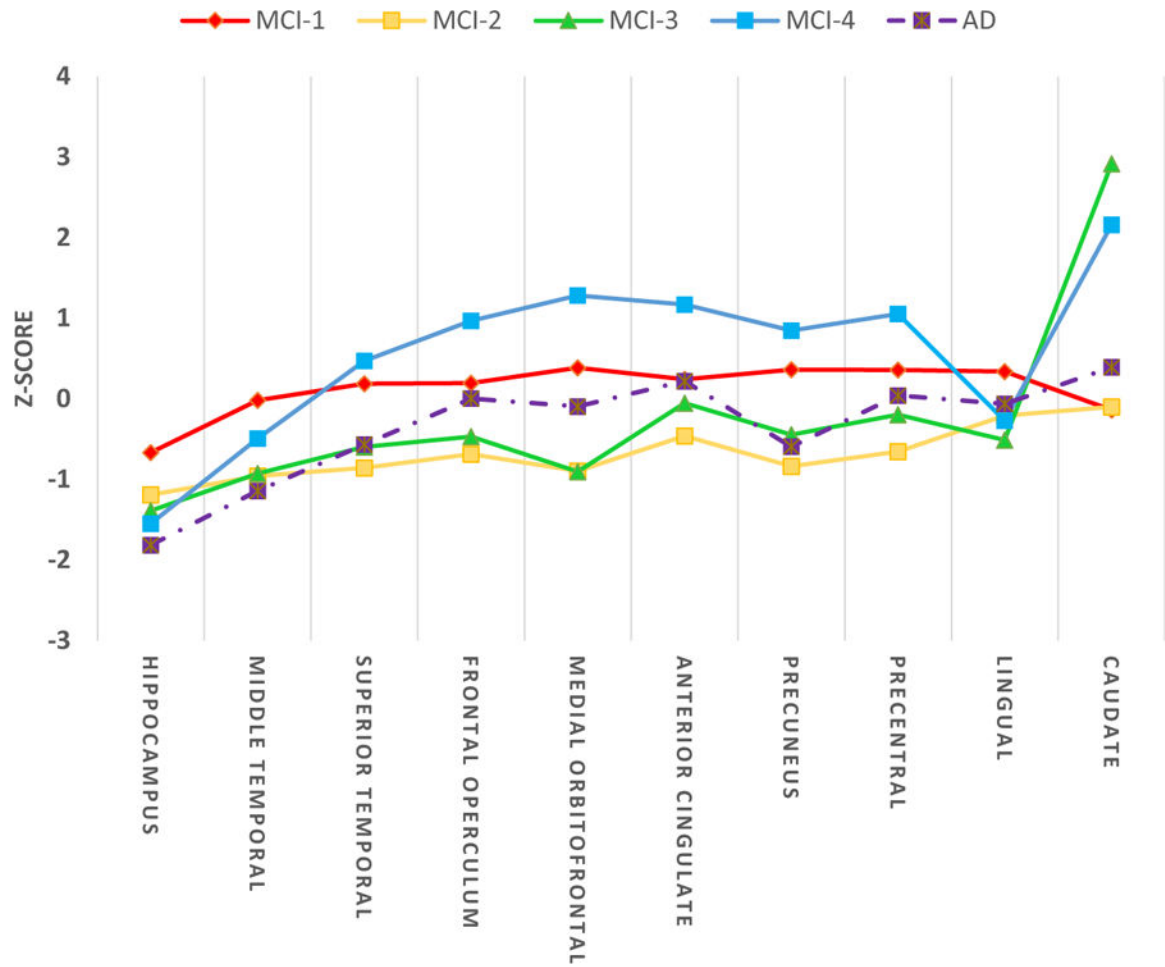


Fig. 1. Comparison of volumetric measures of ROIs included in LCA. The z scores shown were created by subtracting the ADNI normal control (CN) mean and dividing by the ADNI CN standard deviation for each biomarker shown so that zero represents the mean of the CN group

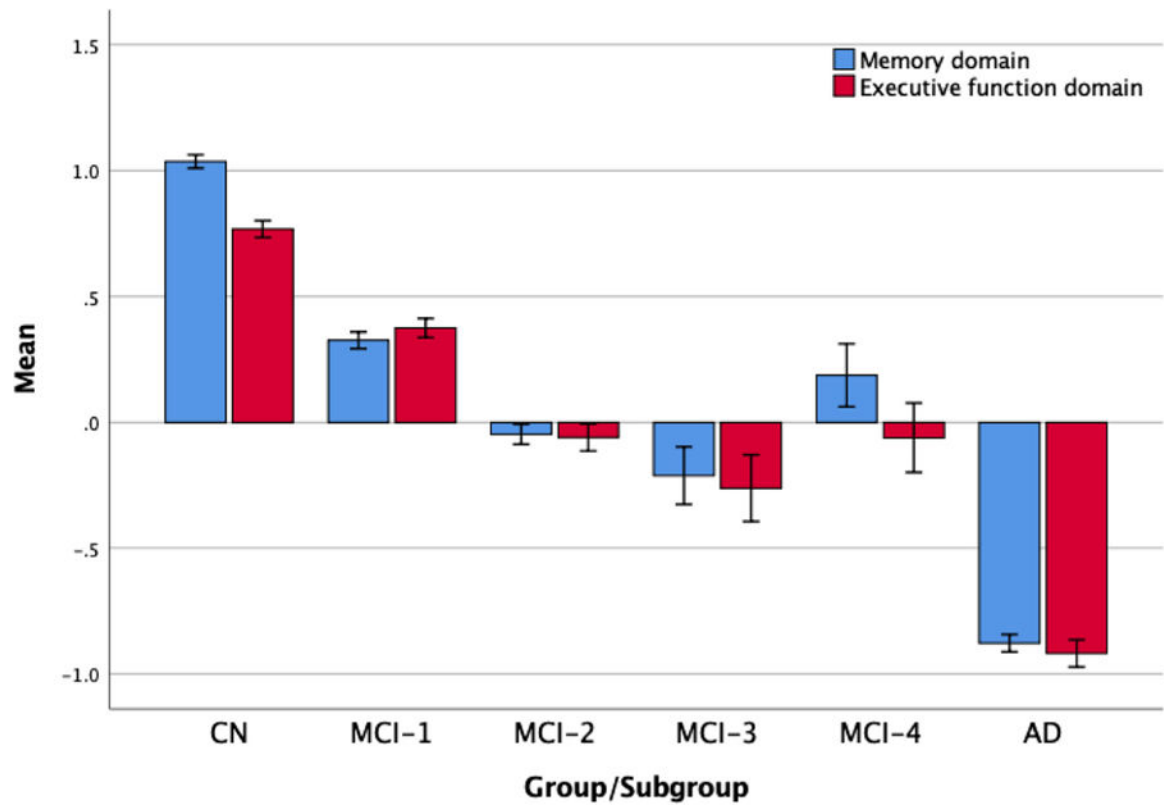


Fig. 2. Neuropsychological performance for the CN, AD, and MCI subgroups. Error bars denote 1 standard error of mean

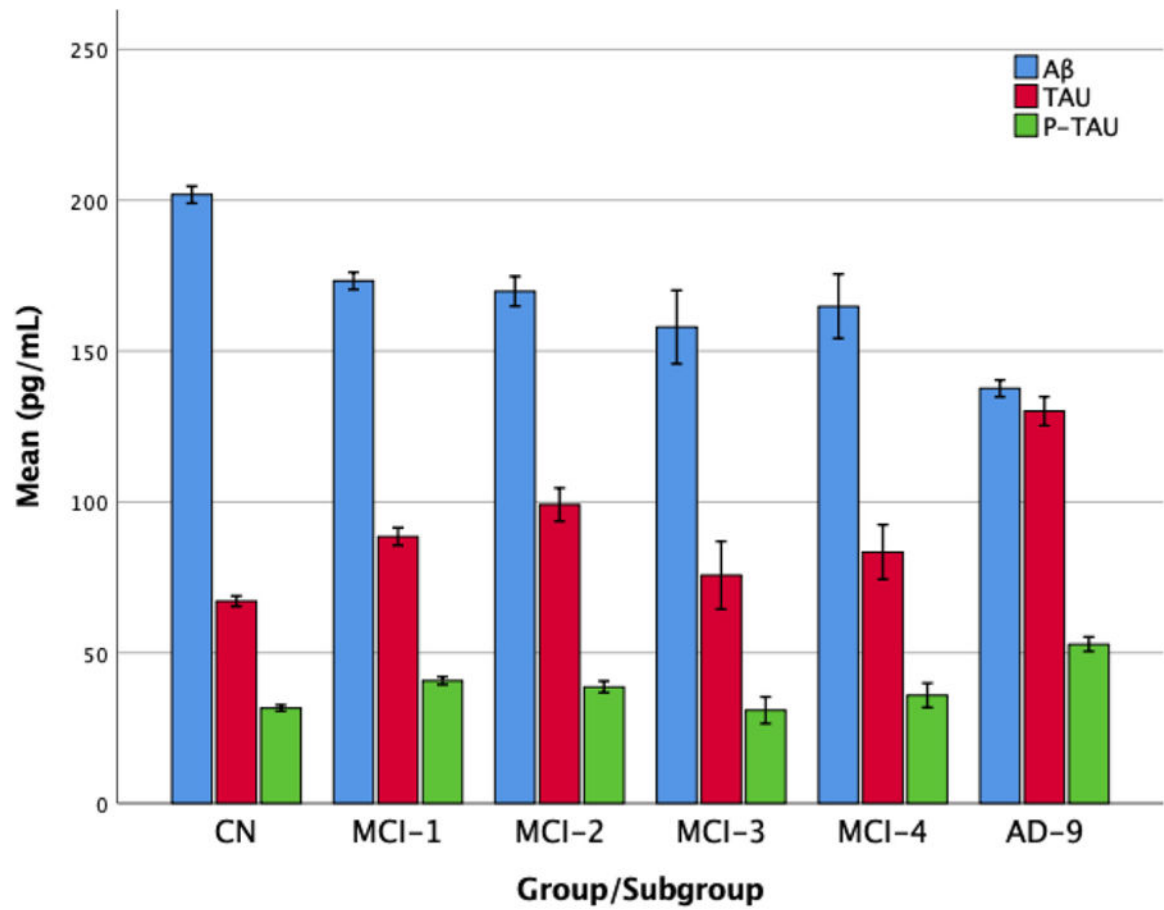


Fig. 3. CSF biomarkers for the CN, AD, and MCI subgroups. Error bars denote 1 standard error of mean

Table 1

Exploratory factor analysis on FreeSurfer derived ROI volumes

Factors	1	2	3	4	5	6	7	8	9	10
% of explained variance	8.3	8.2	7.9	7.8	7.3	5.3	5.2	5.0	4.0	3.8
Brain Regions										
Medial Orbitofrontal	.77									
Hippocampus		.82								
Precentral			.70							
Middle Temporal				.76						
Frontal Operculum					.87					
Superior Temporal						.73				
Lingual							.73			
Caudate								.90		
Anterior Cingulate									.72	
Precuneus										.56
Frontal Pole	.73									
Rostral Middle Frontal	.46									.45
Lateral Orbitofrontal	.50									.42
Isthmus Cingulate	.55									
Insula	.51									
Temporal Pole	.41	.52								
Entorhinal		.75								
Amygdala		.78								
Para-hippocampal		.66								
Caudal Middle Frontal			.66							
Superior Frontal			.63							
Paracentral			.66							
Postcentral			.56							
Inferior Temporal				.51						
Supramarginal				.45						.49
Banks of the superior temporal sulcus					.73					

Factors	1	2	3	4	5	6	7	8	9	10
Inferior Parietal				.68						
Pericalcarine					.80					
Thalamus					.55					
Transverse Temporal						.62				
Fusiform						.41				
Cuneus							.68			
Lateral Occipital							.71			
Putamen								.72		
Ventricles								.52		
Posterior Cingulate									.57	
Superior Parietal										.54

Extraction method: Principal Component Analysis, Rotation Method: Varimax. Loadings less than 0.4 are excluded from the table for the purpose of simplicity

Table 2

Information criteria of the class solutions on the MCI sample

Classes	Bayesian information criterion (BIC)	Entropy
2	14,757	0.719
3	14,703	0.800
4	14,685	0.828
5	14,685	0.741
6	14,700	0.765

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

Participants demographics for different subgroups of study

	CN	MCI 1	MCI 2	MCI 3	MCI 4	p	AD
N (% of diagnostic group)	449	404 (58)	230(33)	36 (5.2)	26(3.7)		249
Age, mean (SD)	73.9(5.6)	71.4(7.1) ^{3,4}	72.8(7.7) ^{3,4}	77.2(5.4) ^{1,2}	78.9(5.9) ^{1,2}	<.001	74.3(7.9)
Sex-male, %	51.4%	53.2%	67.4%	77.8%	26.9%	<.001	45.4%
Education, mean (SD)	16.3(2.7)	16.2(2.7) ²	15.4(3.0) ¹	15.9(3.2)	15.2(2.9)	.003	15.2(2.9)
Apo E (% e4 carrier)	27.7%	50.1%	52.6%	55.6%	50%	.920	70.4%

P-values are representative of outcomes of analyses of variance (ANOVA) for continuous variables, and Chi-square tests for categorical variables with MCI subgroup membership as the independent variable. MCI subgroup membership as the independent variable. Numbered superscripts denote significant post-hoc differences at *p* < 0.01 between each aMCI subgroup and the subgroup number indicated

Table 4

Clinical, genetic, CSF biomarker, and longitudinal information separately for CN, AD, and MCI subgroups based on latent class analysis

	CN	MCI1	MCI2	MCI3	MCI4	P	AD
Clinical							
CDR	.03(.13)	1.42(.86) ²	1.64(.88) ²	1.76(.74)	1.53(.81)	.010	4.36(1.6)
FAQ	.20(.7)	2.6(3.8) ²	3.9(4.4)	4.1(4.4)	3.4(4.8)	.010	12.8(6.9)
GDS	.86(1.1)	1.7(1.4)	1.7(1.4)	1.6(1.5)	1.6(1.2)	.931	1.64(1.4)
Attention							
Digit Span forward	8.73(2.0)	8.4(2.0)	8.0(1.9)	8.4(2.2)	7.2(2.6)	.271	7.5(1.9)
Digit Span backward	7.14(2.1)	6.2(2.0)	6.0(1.9)	5.3(2.2)	7.0(3.3)	.244	4.8(2.1)
Executive function							
Digit Symbol	46.03(10.1)	38.3(10.3)	36.6(11.3)	32.6(10.4)	41.7(22.8)	.123	26.0(13.1)
Trails B	85.09(42.8)	100.7(51.7) ^{2,3,4}	135.1(78.6) ¹	146.8(73.8) ¹	149.3(79.1) ¹	<.001	201.4(86.2)
Category naming	34.68(8.1)	27.4(7.2)	26.9(7.3)	26.8(8.0)	28.5(1.8)	.677	20.1(6.9)
Memory							
MMSE	29.0(1.1)	28.0(1.6) ^{2,3}	26.9(1.8) ^{1,4}	26.8(1.7) ^{1,4}	27.8(1.8) ^{2,3}	<.001	23.12(2.0)
ADAS-cog	5.8(2.8)	9.3(4.3) ^{2,3}	11.2(4.2) ¹	11.8(11.9) ¹	10.9(5.5)	<.001	19.9(6.9)
RAVLT (sum of five trials)	44.9(9.7)	36.5(10.7) ^{2,3}	31.0(9.4) ¹	28.8(10.7) ¹	35.1(9.1)	<.001	22.8(7.4)
RAVLT 30 min	3.73(2.7)	4.6(2.5)	4.5(2.3)	4.6(4.5)	4.8(5.8)	.905	4.4(1.8)
Logical Memory II	13.1(3.3)	6.5(3.4) ^{2,3}	4.4(2.9) ^{1,4}	4.5(3.3) ^{1,4}	5.9(3.4) ^{2,3}	<.001	1.3(1.9)
Summary scores							
Memory domain	1.03(.56)	0.32(.67) ^{2,3}	-0.05(.60) ¹	-0.21(.69) ¹	0.18(.64)	<.001	-87(.54)
Executive function domain	.76(.71)	0.37(.76) ^{2,3,4}	-0.06(.82) ¹	-0.26(.79) ¹	-0.06(.69) ¹	<.001	-92(.85)
Imaging Biomarkers							
Total brain Volume (cm ³)*	1050(70)	1053(73) ^{2,3}	1005(72) ^{1,3}	940(76) ^{1,2,4}	1006(63) ³	<.001	975(68)
Ventricle volume (cm ³)*	33.4(18.0)	34.7(19.4) ^{3,4}	40.3(19.7) ^{3,4}	75.9(31.8) ^{1,3,4}	56.4(22.5) ^{1,2,3}	<.001	50.5(22)
WMH (Log cm ³)	-.07(.8)	0.05(.82) ^{2,4}	-0.41(.76) ^{1,3,4}	0.22(.76) ^{2,4}	0.89(.72) ^{1,2}	<.001	.047(0.04)
Average FDG-PET of angular, temporal, and posterior cingulate	1.32(.1)	1.27(.13) ^{2,3}	1.21(.12) ¹	1.15(.12) ¹	1.21(.14)	<.001	1.06(.14)
CSF Biomarkers							
Low Aβ ₁₋₄₂ (%)	28.7(4.5)	52.2 ²	39.5 ¹	38.8	57.6	<.009	64.2(4.8)

	CN	MCI 1	MCI 2	MCI 3	MCI 4	P	AD
High Tau (%)	12.4(3.3)	27.9	22.1	13.8	23.0	.153	45.7(5.0)
High P-tau _{181p} (%)	43.6(4.9)	62.1 ^{2,3}	41.7 ¹	25.0 ^{1,4}	57.6 ³	<.001	64.2(4.8)
High P-tau/A β ratio (%)	48.7(5.0)	64.1 ^{2,3}	45.2 ¹	38.8 ^{1,4}	65.3 ³	<.001	66.2(4.7)
Longitudinal outcome							
Incidence rate of conversion to AD ^{**}		7.6 (6.4–9.2) ^{2,3}	12.3 (10.2–14.9) ¹	16.5 (10.3–26.1) ¹	12.3 (6.8–22.3) ¹	<.001	
Rate of Reversion to CN ^{**}		1.9 (1.3–2.8) ^{2,3}	0.9 (0.4–1.8) ^{1,3,4}	0 ¹	0 ¹	<.001	

Data summarized as mean and (SE) unless otherwise noted

P-values are representative of outcomes of ACNOVA with age, gender, education and MRI field strength as covariates and MCI subgroup membership as independent variable. Numbered superscripts denote significant post-hoc differences at $p < 0.01$ between each aMCI subgroup and the subgroup number indicated

* Measures adjusted by total intracranial volume

** rates are reported “per 100 person-years”, with 95% confidence interval

Table 5

Volumetric measures of ROIs used in latent class models by class

	CN Mean(SD)	MCI-1 Mean (SD)	MCI-2 Mean (SD)	MCI-3 Mean(SD)	MCI-4 Mean(SD)	p	AD Mean(SD)
Medial temporal*							
Hippocampus	3.67(.35)	3.43(.46) ^{2,3,4}	3.25(.43) ¹	3.18(.40)	3.12(.39) ¹	<.001	3.02(.45)
Lateral temporal							
Middle Temporal	10.02(.88)	10.00(.87) ^{2,3}	9.18(.94) ¹	9.20(1.13) ¹	9.58(.90)	<.001	9.01(1.04)
Superior Temporal	10.04(.77)	10.21(.77) ^{2,3}	9.29(.82) ^{1,4}	9.52(.74) ^{1,4}	10.46(.77) ^{2,3}	<.001	9.54(.93)
Parietal							
Precuneus	8.24(.63)	8.52(.63) ^{2,3,4}	7.59(.73) ^{1,4}	7.90(.47) ^{1,4}	8.90(.83) ^{1,2,3}	<.001	7.78(.94)
Cingulate							
Anterior Cingulate	5.99(.50)	4.04(.50) ^{2,4}	3.68(.44) ^{1,4}	3.89(.47) ^{1,4}	4.51(.50) ^{1,2,3}	<.001	6.94(.50)
Frontal							
Medial Orbitofrontal	4.24(.43)	4.45(.43) ^{2,3,4}	3.76(.42) ^{1,4}	3.76(.37) ^{1,4}	4.94(.45) ^{1,2}	<.001	4.19(.60)
Frontal Operculum	9.01(.77)	9.18(.77) ^{2,3,4}	8.43(.76) ^{1,4}	8.61(.92) ^{1,4}	9.84(.68) ^{1,2,3}	<.001	9.01(.92)
Primary Sensory/motor							
Precentral	11.39(.91)	11.80(.91) ^{2,3,4}	10.67(.99) ^{1,3,4}	11.18(1.18) ^{1,2,4}	12.57(.94) ^{1,2,3}	<.001	11.44(1.24)
Occipital							
Lingual	5.75(.71)	5.98(.70) ^{2,3,4}	5.62(.65) ¹	5.41(.49) ¹	5.58(.86) ¹	<.001	5.71(1.78)
Subcortical structures							
Caudate	3.43(.36)	3.36(.36) ^{3,4}	3.38(.38) ^{3,4}	4.87(.53) ^{1,2,4}	4.50(.49) ^{1,2,3}	<.001	3.62(.55)

P-values are representative of outcomes of ACNOVA with age, gender, education and MRI field strength as independent variable. Numbered superscripts denote significant post-hoc differences at $p < 0.01$ between each aMCI subgroup and the subgroup number indicated. All structural volumes are reported in cubic centimeters