Published in final edited form as:
*Arch Neurol*. 2012 May ; 69(5): 614–622. doi:10.1001/archneurol.2011.3029.

# Comparison of imaging biomarkers in ADNI versus the Mayo Clinic Study of Aging

**Jennifer L. Whitwell, PhD**[1], **Heather J. Wiste, BA**[2], **Stephen D. Weigand, MS**[2], **Walter A. Rocca, MD, MPH**[3,4], **David S. Knopman, MD**[4], **Rosebud O. Roberts, MB, Ch.B**[3], **Bradley F. Boeve, MD**[4], **Ronald C. Petersen, MD, PhD**[4], **Clifford R. Jack Jr, MD**, and **The Alzheimer's Disease Neuroimaging Initiative***

[1]Department of Radiology, Mayo Clinic, Rochester MN

[2]Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester MN

[3]Division of Epidemiology, Department of Health Sciences Research, Mayo Clinic, Rochester MN

[4]Department of Neurology, College of Medicine, Mayo Clinic, Rochester MN

## Abstract

**Objective—**To determine whether MRI measurements observed in the Alzheimer's Disease Neuroimaging Initiative (ADNI; convenience-sample) differ from those observed in the Mayo Clinic Study of Aging (MCSA; population-based sample).

**Design—**Comparison of two samples.

**Setting—**59 recruiting sites for the ADNI in US/Canada, and the MCSA, a population-based cohort in Olmsted County, MN.

**Patients—**Cognitively normal (CN) subjects and amnestic mild cognitive impairment (aMCI) subjects were selected from the ADNI convenience cohort and MCSA population-based cohort. Two samples were selected; the first was a simple random sample of subjects from both cohorts in the same age range, and the second applied matching for age, sex, education, apolipoprotein E genotype, and Mini-Mental State Examination.

**Main outcome measures—**Baseline hippocampal volumes and annual percent decline in hippocampal volume.

**Results—**In the population-based sample, MCSA subjects were older, less educated, performed worse on MMSE, and less often had family history of AD than ADNI subjects. Baseline hippocampal volumes were larger in ADNI compared to MCSA CN subjects in the random sample, although no differences were observed after matching. Rates of decline in hippocampal volume were greater in ADNI compared to MCSA for both CN and aMCI, even after matching.

**Conclusions—**Rates of decline in hippocampal volume suggest that ADNI subjects have more aggressive brain pathology than MCSA subjects, and hence may not be representative of the general population. These findings have implications for treatment trials that employ ADNI-like recruitment mechanisms and for studies validating new diagnostic criteria for AD in its various stages.

Please send correspondence to: Jennifer L. Whitwell, PhD Assistant Professor of Radiology Mayo Clinic 200 1st St SW Rochester MN 55905 Tel: 507 284 5576 Fax: 507 284 9778 whitwell.jennifer@mayo.edu.

*Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.ucla.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.ucla.edu/wp-content/uploads/how_to_apply/ADNI_Authorship_List.pdf

## INTRODUCTION

Imaging plays an important role in the study of Alzheimer's disease (AD). Imaging biomarkers can track disease progression[1], detect changes early in the mild cognitive impairment (MCI) phase[2, 3] and help predict which subjects may later develop AD[4, 5]. Imaging measures have become common outcome measures in clinical treatment trials because they may reduce sample size[6, 7]. Increasing interest in using imaging in clinical trials led to the development of the Alzheimer's Disease Neuroimaging Initiative (ADNI) which aimed to improve methods for clinical trials and validate imaging and other biomarkers[8, 9]. ADNI is an observational study of MCI and AD that used identical recruitment mechanisms as typical trials, including advertising and recruitment from memory clinics. Therefore, ADNI is based on a highly selected convenience-sample. Because ADNI data are freely available, a large number of studies are published each year using these data. However, it is unclear to what extent subjects recruited through these mechanisms are representative of the general population, and hence whether results are generalizable.

We aimed to determine whether imaging measures would differ in ADNI participants compared to the population-based cohort of the Mayo Clinic Study of Aging (MCSA). We assessed hippocampal volume and rates of decline in hippocampal volume because they are established and widely studied biomarkers of AD[6, 10, 11]. Because results could be influenced by differences in inclusion characteristics and demographics, we compared the cohorts both before and after matching for specific demographic and cognitive features.

## METHODS

### Sources of subjects and diagnostic criteria

Subjects with a clinical diagnosis of amnestic mild cognitive impairment (aMCI) and cognitively normal (CN) subjects who had been recruited into either the MCSA (and had agreed to MRI studies) or ADNI were analyzed.

The MCSA is a longitudinal epidemiologic study of normal ageing and MCI in Olmsted County, Minnesota. The recruitment mechanisms have been reported in detail previously[12]. Briefly, all Olmsted County residents aged 70–89 years on October 1, 2004 were identified using the medical records-linkage system of the Rochester Epidemiology Project[13, 14]. The population was also resampled in 2008 and 2009 in order to replenish the cohort. Subjects were randomly selected from this enumeration. Subjects received a letter of invitation giving them the opportunity to refuse participation by returning a letter of refusal. Subjects who did not return the letter then received a follow-up telephone call inviting them to participate. MRI was performed in all subjects who agreed to participate and did not have any contraindications to MRI. **S**ubjects with imaging in the MCSA have very similar demographic characteristics to those that did not undergo imaging (Table 1). Subjects were characterized as CN by consensus[12, 15], and when their age-adjusted neuropsychological test scores were consistent with normative data developed in this community[16]. Diagnostic criteria for MCI were as follows[17]: 1) cognitive concern by subject, informant (from Clinical Dementia Rating scale (CDR)[18]), nurse or physician; 2) impairment in 1 or more of the 4 cognitive domains (from cognitive battery); 3) essentially normal functional activities (using CDR and Functional Activities Questionnaire); and 4) absence of dementia (DSM-IV)[19]. Subjects were categorized as amnestic MCI (aMCI) if memory was impaired. The diagnosis of aMCI was made on clinical grounds without the use of rigid cutoffs on psychometric scores.

ADNI is a longitudinal multi-site observational study of CN, aMCI and AD (www.ADNI-info.org)[8]. Subjects were recruited using local Alzheimer's Disease Research Centers, memory clinics, newspaper ads, radio, and other public media campaigns. Diagnostic criteria for ADNI were largely the same as for the MCSA. Criteria for CN subjects were: Mini Mental State Exam (MMSE)[20] scores between 24 and 30 inclusive; no memory complaints; objective memory performance in the normal range; and a CDR score of 0 and memory box score of 0. Diagnostic criteria for aMCI were: 1) memory complaint verified by an informant; 2) objective memory impairment measured by education adjusted score on the Wechsler Memory Scale-Revised, Logical Memory II; 3) MMSE scores between 24 and 30 inclusive; 4) CDR score of 0.5 and memory box score of at least 0.5, and 5) preservation of general cognition and functional activities of daily living. Subjects enrolled for ADNI were between 55 and 90 years old. The ADNI AD subjects were not included in our analysis because the MCSA does not follow AD subjects.

Informed consent was obtained from all subjects. The MCSA was approved by the Mayo Clinic IRB, and ADNI was approved by the IRB at each individual site.

## Subject selection

We selected two samples of subjects: the first was a random sample of all available MCSA and ADNI subjects, and the second sample applied matching for demographic and cognitive variables. Both cross-sectional and longitudinal samples were selected. The first available MRI was used for the cross-sectional analysis and as baseline for the longitudinal analysis. Two serial MRI were used for the longitudinal analysis for each subject. Scan interval was approximately 12-months for ADNI and 15-months for the MCSA (the routine follow-up interval in the MCSA).

**Sample 1: Simple random sample of each cohort—**For the cross-sectional analysis, the total number of available CN subjects was 229 in ADNI and 1,283 in the MCSA. The total number of aMCI subjects available was 397 in ADNI and 179 in the MCSA. To obtain comparable sample sizes between ADNI and MCSA, we took a simple random sample of the MCSA CN subjects resulting in 229 subjects. Similarly, we took a simple random sample of the ADNI aMCI subjects resulting in 179 subjects. Because of the random sub-sampling strategy, the samples used for our analyses were representative (within sampling error) of the parent cohorts from which they were drawn (Supplemental Table e–1). For the longitudinal analysis, there were 206 ADNI CN subjects with serial scans and 686 MCSA CN subjects. There were 347 ADNI aMCI subjects with serial data and 92 MCSA aMCI subjects. Once again, to obtain comparable group sizes, we took a random sample of the MCSA CN subjects and ADNI aMCI subjects, resulting in 206 MCSA CN subjects and 92 ADNI aMCI subjects.

**Sample 2: Age, sex, education, APOE genotype, and MMSE matched samples —**In sample 2, the ADNI and MCSA subjects were frequency matched by age, sex, education, apolipoprotein (APOE) genotype, and MMSE score. All variables were dichotomized into strata: age (70–79 and 80–90 years), sex (men and women), education (6–13 and 14–20 years), and MMSE (24–28 and 29–30 for CN; 22–25 and 26–30 for aMCI). ADNI and MCSA subjects were matched with a one-to-one frequency by taking a random sample within each of the 32 strata of the larger study group to match the number of subjects in the smaller study group. CN and aMCI subjects were matched separately. Subjects that could not be matched were excluded. For the cross-sectional analysis, 212 CN subjects and 97 aMCI subjects were selected for both ADNI and MCSA. For the longitudinal analysis, 191 CN subjects and 65 aMCI subjects were selected for both ADNI and MCSA.

Subject demographics for the two cross-sectional and longitudinal samples are shown in Tables 2 and 3. The samples used for analysis differ slightly from those reported above because some subjects were excluded due to poor quality imaging.

## Image analysis

MRI acquisition protocols were very similar for MCSA and ADNI subjects, although MCSA subjects were scanned at 3T while ADNI subjects were scanned at 1.5T. ADNI collects 1.5T MRI scans in all subjects and 3T scans in only 25% of the sample; therefore ADNI 1.5T MRI scans were used for this study. To ensure that field strength did not bias our results, we compared hippocampal volumes at 1.5T and 3T in ADNI subjects that were scanned at both field strengths. Similar to a previous study[21], hippocampal measurements were comparable across field strengths (Figure 1).

MCSA subjects were imaged with a 3D magnetization prepared rapid acquisition gradient echo (MPRAGE) sequence developed at Mayo for ADNI[9]. Parameters were: sagittal plane, TR/TE/TI, 2300/3/900ms; flip angle 8°, 26cm field of view (FOV); 256×256 in-plane matrix with a phase FOV of 0.94 and slice thickness of 1.2mm. ADNI is a multi-site study and there are minor variations in the MRI protocol based on the specific hardware/software configuration on each scanner. The nominal parameters of the ADNI MPRAGE were: sagittal plane, TR/TE/TI, 2400/3/1000ms; flip angle 8°, 24cm FOV; 192×192 in-plane matrix and slice thickness of 1.2mm[9].

All images were corrected for gradient non-linearity[22] and intensity inhomogeneity[23]. Hippocampal volumes were measured using Freesurfer software version 4.5.0[24]. The cross-sectional analysis pipeline was used to calculate hippocampal volumes for the cross-sectional sample, and the longitudinal analysis pipeline was used to assess rates of hippocampal change for the longitudinal sample. Hippocampal measurements calculated using Freesurfer have been previously validated against manual measurements[25]. Total intracranial volumes (TIV) were measured using an algorithm developed in-house[26].

## Statistics

Analyses were performed in R version 2.11.0 (R Development Core Team: A Language and Environment for Statistical Computing. 2010 http://www.r-project.org) and tests of statistical significance were conducted at the two-sided alpha level of 0.05. For the cross-sectional analysis, we used hippocampal volume adjusted for TIV. We fit a linear regression model of hippocampal volume (y) versus TIV (x) in all ADNI and MCSA CN subjects with available data (n=1,480) and then used the intercept ($b_0$) and slope ($b_1$) estimates from the model to calculate hippocampal volume adjusted for TIV (HVa) as a residual [$HVa = HP - (b_0 + b_1 \times TIV)$]. For the longitudinal analysis, the annual percent decline in hippocampal volume was calculated as follows using unadjusted hippocampal volumes: (follow-up volume – baseline volume)/(baseline volume × years between scans) × 100.

Wilcoxon rank-sum/Mann-Whitney U tests were used to test differences in continuous measures between ADNI and MCSA groups, and chi-squared tests with continuity correction or Fisher's exact test were used to test differences in categorical variables. We summarized group differences in imaging measures using the probabilistic index (PI) (corresponding to the area under the receiver operator characteristic curve)[27]. The PI is a non-parametric estimate of group-wise differences or discrimination that measures the probability that the value from a randomly selected subject in one group is higher than the value from a randomly selected subject in the other group. A PI of 0.50 (or 50%) indicates no difference across groups.

# RESULTS

## Subject demographics

Differences in demographic features across MCSA and ADNI were similar for cross-sectional and longitudinal cohorts (Tables 2 and 3). In sample 1, MCSA subjects (aMCI and CN) were older, less educated, and performed worse on the MMSE than the ADNI subjects. The MCSA aMCI subjects had a lower proportion of APOE ε4 carriers than ADNI. No differences were observed in sex, education, or APOE genotype between the MCSA and ADNI subjects in sample 2. Despite frequency matching, age (cross-sectional sample only) and MMSE in the CN subjects still differed across the cohorts, although the median and interquartile ranges were similar. ADNI had a higher proportion of family history of AD, and of minorities across all samples for the CN subjects with a similar trend for aMCI in the cross-sectional sample.

## Cross-sectional results

In sample 1, hippocampal volume adjusted for TIV was significantly smaller in the MCSA CN subjects compared to the ADNI CN subjects, with no differences between the groups for the aMCI subjects (Figure 2A). After matching for age, sex, education, APOE genotype, and MMSE in sample 2, no differences in hippocampal volume adjusted for TIV were observed between MCSA and ADNI in either the CN or aMCI subjects (Figure 2B).

## Longitudinal results

In sample 1, the annual percent decline in hippocampal volume was greater in ADNI compared to the MCSA for both aMCI and CN subjects (Figure 3A). After matching for age, sex, education, APOE genotype, and MMSE in sample 2, these differences across ADNI and MCSA were still observed (Figure 3B).

# COMMENT

This study highlights demographic differences in subjects recruited into the convenience-sample ADNI cohort compared with subjects recruited into the population-based MCSA cohort, and demonstrates that imaging biomarkers from these two different recruitment mechanisms differ.

The most striking difference was that rates of decline in hippocampal volume were greater in ADNI compared to the MCSA, for both CN and aMCI subjects. This difference was observed even after matching for key demographic and cognitive variables. Increased rates of decline in hippocampal volume in CN subjects predict a faster rate of progression to dementia[28], suggesting that the ADNI CN population includes a larger proportion of subjects on the path to AD dementia. While it was somewhat unexpected that the proportion of APOE e4 carriers was not higher in the ADNI CN subjects, our findings are consistent with the unusually high proportion (50%) of ADNI controls who showed amyloid pathology as measured by Pittsburgh Compound B (PiB)[29]. By contrast, the proportion of PiB positive controls in the MCSA was only 30%[30]. The pathological diagnosis of AD was also more common in controls from a clinic versus a community setting in a previous study[31]. The ADNI CN subjects were more highly educated than the MCSA CN subjects; therefore, cognitive reserve mechanisms may have protected them from clinical decline even though they are on a steeper downward trajectory of brain atrophy. Similarly, the higher rates of atrophy suggest that the aMCI group in ADNI consists of a higher proportion of subjects with a more aggressive disease than the MCSA. Indeed, the aMCI subjects in ADNI had a higher proportion of APOE e4 carriers than those in the MCSA in sample 1. Once again, the aMCI subjects in ADNI were more highly educated than those in the MCSA suggesting that

cognitive reserve mechanisms may have protected them from decline on the MMSE and progression to a clinical diagnosis of AD.

We hypothesize that this bias in ADNI is due to the recruitment mechanism. We can speculate that CN subjects who are worried about their cognition would be more likely to attend memory clinics and be more motivated to answer advertisements for the study. Both CN and aMCI subjects with higher education are also more likely to seek medical help at a memory clinic and become involved in observational studies. These highly educated subjects could have a more aggressive underlying disease but are able to compensate cognitively. Amnestic MCI subjects recruited through a population-based study are less likely to have sought medical care at a memory clinic and may have a broader spectrum of cognitive function. In addition, an important motivator for participation in ADNI, and other convenience studies, could be the presence of a family history of dementia. Indeed, ADNI did show a higher proportion of family history compared to MCSA. Although one may assume that similar biases would be observed in the MCSA subjects that agreed to imaging, we have demonstrated that this is not the case; likely because less effort was required to agree for imaging than seek out participation in ADNI. The clinical inclusion criteria for both CN and aMCI differed slightly across the two cohorts. A diagnosis of CN in the MCSA was made by multidisciplinary consensus, which may be more conservative than the method employed in ADNI. Similarly, the diagnosis of aMCI in the MCSA is based on clinical grounds, whereas ADNI relied more on a specific cut-point on a memory test. The ADNI approach is likely to result in the recruitment of more impaired subjects. The reason that this is not reflected in the MMSE scores could be because higher education is providing cognitive reserve, and the MMSE may be insensitive to subtle cognitive impairment. The ADNI also recruited younger subjects than the MCSA, which could also have resulted in the recruitment of subjects with more aggressive disease. Rates of atrophy have been found to be greater in younger aMCI subjects[32], possibly because they have a more pure, and hence aggressive, AD pathology compared to older subjects. Older subjects are more likely to have a mixture of pathologies[33], including cerebrovascular disease[34, 35]. However, the trend for higher APOE e4 carrier frequency, younger age, and higher education in convenience samples compared to population-based samples has been observed in other cohorts[36–39], suggesting that this bias may be due to the general recruitment mechanism rather than the specific inclusion criteria employed in ADNI. Our findings suggest that CN and aMCI subjects in ADNI are not representative of the general population, and, importantly, suggest that subjects included in future pre-clinical prevention trials using the same recruitment mechanisms will also not be representative of the population. Finally, our results indicate that even rigorous demographic matching efforts are insufficient to correct for the selection bias.

The only difference observed in baseline hippocampal volumes between ADNI and MCSA was in the CN subjects in sample 1, with larger hippocampal volumes observed in ADNI. This difference is likely being driven by the younger age of the ADNI cohort, since hippocampal volume has been shown to decrease with age[40]. After matching for demographic features no differences in hippocampal volume were observed across cohorts. Cross-sectional hippocampal volumes also did not differ across ADNI and the MCSA within the aMCI subjects in sample 1, despite the observed differences in age, education, APOE genotype, and MMSE score. This could suggest that rates of decline in hippocampal volume are more sensitive markers of incident AD than cross-sectional hippocampal volume, perhaps because of the large degree of inter-subject variability in hippocampal volume. TIV also differed between MCSA and ADNI. We suspect that MCSA subjects have larger TIVs because of the northern European heritage of many Minnesotan residents, and the link between these nationalities and tall height[41].

The strengths of this study are the large numbers of subjects and the use of two samples with and without restrictive correction for major demographic or cognitive confounders. A limitation however, is that while matching was performed on the major demographic factors, it may not eliminate other potential differences, such as in other comorbidities, medication, family history, race and ethnicity, that may influence the imaging findings. ADNI had a higher proportion of minorities than MCSA. The MCSA and ADNI cohorts underwent imaging at different field strengths; however, we demonstrated excellent agreement between hippocampal volumes measured across field strengths (Figure 1). Scan intervals also differed between ADNI and MCSA, although we adjusted for these differences. While atrophy rates have been shown to accelerate over time in AD[32], the trajectory of change is likely to be approximately linear over these relatively short intervals. Lastly, while the MCSA is a population-based study there may also be some inherent participation biases[12], as is the case with any survey. The MCSA is however representative of Olmsted County in Minnesota, and of US Caucasians in general. The incidence of MCI and the demographic predictors of incident MCI in the MCSA are also similar to those reported in other population-based studies[42–44], including studies that have assessed other racial groups[45].

Overall, our findings show that subjects recruited into ADNI are not representative of the general population, and instead more closely resemble clinical populations. The imaging findings all point towards ADNI including more CN subjects that are on the path to AD dementia and more aMCI subjects that have a pure and aggressive disease phenotype. Therefore, convenience clinical series may be limited by selection biases. These findings have important implications for the design of future treatment trials. If studies that assess power calculations and sample size estimates are performed in biased convenience samples, the high rates of atrophy will lead to smaller than appropriate sample size estimates and therefore trials could be underpowered to detect treatment effects in the population. In addition, treatment trials that utilize a convenience sample will include a higher proportion of subjects with a pure and aggressive disease, and hence are more likely to detect a treatment effect. However, the magnitude of the treatment effect is likely to be less than expected when the treatment is applied to an unbiased population in which subjects are less likely to have pure AD. Care should also be taken when interpreting imaging studies from convenience samples, like ADNI. Biomarkers identified from these highly selected convenience-samples may not perfectly translate to the general population, and will need to be validated in a population-based sample. This will be particularly important for studies seeking to validate new diagnostic criteria for AD in its various stages, in which imaging biomarkers play an important role.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

Investigators on this study were funded by the NIH (U01-AG024904-01, R01-AG11378, P50-AG16574, U01-AG06786, R21-AG38736, R01-DC010367, R01-AG037491, K01-AG028573, U24-AG026395, RO1-AG15866, R01-AG034676, R01-AG023195, R01-HL70825), the DHHS/OS [90BC0009], the Dana Foundation, the Alexander Family Alzheimer's Disease Research Professorship of the Mayo Foundation, and the Robert H. and Clarice Smith and Abigail Van Buren Alzheimer s Disease Research Program of the Mayo Foundation.

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01-AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Abbott, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly and Co., Medpace, Inc., Merck and Co., Inc., Novartis AG, Pfizer Inc, F. Hoffman-La Roche, Schering-Plough, Synarc, Inc., as well as non-profit partners the Alzheimer's Association and Alzheimer's Drug Discovery

Foundation, with participation from the U.S. Food and Drug Administration. Private sector contributions to ADNI are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of California, Los Angeles. This research was also supported by NIH grants P30 AG010129, K01 AG030514, and the Dana Foundation.

## REFERENCES

1. Whitwell JL, Przybelski SA, Weigand SD, et al. 3D maps from multiple MRI illustrate changing atrophy patterns as subjects progress from mild cognitive impairment to Alzheimer's disease. Brain. 2007; 130:1777–1786. [PubMed: 17533169]
2. Whitwell JL, Shiung MM, Przybelski SA, et al. MRI patterns of atrophy associated with progression to AD in amnestic mild cognitive impairment. Neurology. 2008; 70:512–520. [PubMed: 17898323]
3. Chetelat G, Desgranges B, De La Sayette V, et al. Mapping gray matter loss with voxel-based morphometry in mild cognitive impairment. Neuroreport. 2002; 13:1939–1943. [PubMed: 12395096]
4. Jack CR Jr. Petersen RC, Xu YC, et al. Prediction of AD with MRI-based hippocampal volume in mild cognitive impairment. Neurology. 1999; 52:1397–1403. [PubMed: 10227624]
5. Ferreira LK, Diniz BS, Forlenza OV, et al. Neurostructural predictors of Alzheimer's disease: A meta-analysis of VBM studies. Neurobiology of aging. 2009
6. Jack CR Jr. Slomkowski M, Gracon S, et al. MRI as a biomarker of disease progression in a therapeutic trial of milameline for AD. Neurology. 2003; 60:253–260. [PubMed: 12552040]
7. Schott JM, Frost C, Whitwell JL, et al. Combining short interval MRI in Alzheimer's disease: Implications for therapeutic trials. Journal of neurology. 2006; 253:1147–1153. [PubMed: 16998650]
8. Mueller SG, Weiner MW, Thal LJ, et al. The Alzheimer's disease neuroimaging initiative. Neuroimaging clinics of North America. 2005; 15:869–877. xi–xii. [PubMed: 16443497]
9. Jack CR Jr. Bernstein MA, Fox NC, et al. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. J Magn Reson Imaging. 2008; 27:685–691. [PubMed: 18302232]
10. Jack CR Jr. Dickson DW, Parisi JE, et al. Antemortem MRI findings correlate with hippocampal neuropathology in typical aging and dementia. Neurology. 2002; 58:750–757. [PubMed: 11889239]
11. Fox NC, Warrington EK, Freeborough PA, et al. Presymptomatic hippocampal atrophy in Alzheimer's disease. A longitudinal MRI study. Brain. 1996; 119(Pt 6):2001–2007. [PubMed: 9010004]
12. Roberts RO, Geda YE, Knopman DS, et al. The Mayo Clinic Study of Aging: design and sampling, participation, baseline measures and sample characteristics. Neuroepidemiology. 2008; 30:58–69. [PubMed: 18259084]
13. Melton LJ 3rd. History of the Rochester Epidemiology Project. Mayo Clinic proceedings. 1996; 71:266–274. [PubMed: 8594285]
14. St Sauver JL, Grossardt BR, Yawn BP, et al. Use of a medical records linkage system to enumerate a dynamic population over time: the Rochester epidemiology project. American journal of epidemiology. 2011; 173:1059–1068. [PubMed: 21430193]
15. Petersen RC, Roberts RO, Knopman DS, et al. Prevalence of mild cognitive impairment is higher in men. The Mayo Clinic Study of Aging. Neurology. 2010; 75:889–897. [PubMed: 20820000]
16. Ivnik RJ, Malec JF, Smith GE, et al. Mayo's older americans normative studies: Updated AVLT norms for ages 56 to 97. The Clinical Neuropsychologist. 1992; 6:1744–4144.
17. Petersen RC. Mild cognitive impairment as a diagnostic entity. Journal of internal medicine. 2004; 256:183–194. [PubMed: 15324362]
18. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology. 1993; 43:2412–2414. [PubMed: 8232972]
19. American Pyschiatric Association. Diagnostic and Statistical Manual of Mental Disorders (4th Ed) (DSM IV). American Psychiatric Association; Washington DC: 1994.

20. Folstein MF, Folstein SE, McHugh PR. “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975; 12:189–198. [PubMed: 1202204]
21. Briellmann RS, Syngeniotis A, Jackson GD. Comparison of hippocampal volumetry at 1.5 tesla and at 3 tesla. Epilepsia. 2001; 42:1021–1024. [PubMed: 11554888]
22. Jovicich J, Czanner S, Greve D, et al. Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. NeuroImage. 2006; 30:436–443. [PubMed: 16300968]
23. Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE transactions on medical imaging. 1998; 17:87–97. [PubMed: 9617910]
24. Fischl B, Salat DH, Busa E, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron. 2002; 33:341–355. [PubMed: 11832223]
25. Morey RA, Petty CM, Xu Y, et al. A comparison of automated segmentation and manual tracing for quantifying hippocampal and amygdala volumes. NeuroImage. 2009; 45:855–866. [PubMed: 19162198]
26. Whitwell JL, Jack CR Jr. Boeve BF, et al. Imaging correlates of pathology in corticobasal syndrome. Neurology. 2010; 75:1879–1887. [PubMed: 21098403]
27. Acion L, Peterson JJ, Temple S, Arndt S. Probabilistic index: an intuitive non-parametric approach to measuring the size of treatment effects. Statistics in medicine. 2006; 25:591–602. [PubMed: 16143965]
28. den Heijer T, van der Lijn F, Koudstaal PJ, et al. A 10-year follow-up of hippocampal volume on magnetic resonance imaging in early dementia and cognitive decline. Brain. 2010; 133:1163–1172. [PubMed: 20375138]
29. Jagust WJ, Landau SM, Shaw LM, et al. Relationships between biomarkers in aging and dementia. Neurology. 2009; 73:1193–1199. [PubMed: 19822868]
30. Jack CR Jr. Lowe VJ, Senjem ML, et al. 11C PiB and structural MRI provide complementary information in imaging of Alzheimer's disease and amnestic mild cognitive impairment. Brain. 2008; 131:665–680. [PubMed: 18263627]
31. Schneider JA, Aggarwal NT, Barnes L, et al. The neuropathology of older persons with and without dementia from community versus clinic cohorts. J Alzheimers Dis. 2009; 18:691–701. [PubMed: 19749406]
32. Jack CR Jr. Weigand SD, Shiung MM, et al. Atrophy rates accelerate in amnestic mild cognitive impairment. Neurology. 2008; 70:1740–1752. [PubMed: 18032747]
33. Jellinger KA, Attems J. Prevalence of dementia disorders in the oldest-old: an autopsy study. Acta neuropathologica. 2010; 119:421–433. [PubMed: 20204386]
34. Jicha GA, Parisi JE, Dickson DW, et al. Neuropathologic outcome of mild cognitive impairment following progression to clinical dementia. Archives of neurology. 2006; 63:674–681. [PubMed: 16682537]
35. Schneider JA, Arvanitakis Z, Leurgans SE, Bennett DA. The neuropathology of probable Alzheimer disease and mild cognitive impairment. Annals of neurology. 2009; 66:200–208. [PubMed: 19743450]
36. Barnhart RL, van Belle G, Edland SD, et al. Geographically overlapping Alzheimer's disease registries: comparisons and implications. Journal of geriatric psychiatry and neurology. 1995; 8:203–208. [PubMed: 8561832]
37. Kokmen E, Ozsarfati Y, Beard CM, et al. Impact of referral bias on clinical and epidemiological studies of Alzheimer's disease. Journal of clinical epidemiology. 1996; 49:79–83. [PubMed: 8598515]
38. Knopman DS, Petersen RC, Rocca WA, et al. Passive case-finding for Alzheimer's disease and dementia in two U.S. communities. Alzheimers Dement. 2011; 7:53–60. [PubMed: 21255743]
39. Tsuang D, Kukull W, Sheppard L, et al. Impact of sample selection on APOE epsilon 4 allele frequency: a comparison of two Alzheimer's disease samples. Journal of the American Geriatrics Society. 1996; 44:704–707. [PubMed: 8642164]

40. Mu Q, Xie J, Wen Z, et al. A quantitative MR study of the hippocampal formation, the amygdala, and the temporal horn of the lateral ventricle in healthy subjects 40 to 90 years of age. Ajnr. 1999; 20:207–211. [PubMed: 10094339]

41. Komlos J, Breitfelder A. Are Americans shorter (partly) because they are fatter? A comparison of US and Dutch children's height and BMI values. Annals of human biology. 2007; 34:593–606. [PubMed: 18092205]

42. Caracciolo B, Palmer K, Monastero R, et al. Occurrence of cognitive impairment and dementia in the community: a 9-year-long prospective study. Neurology. 2008; 70:1778–1785. [PubMed: 18184916]

43. Luck T, Luppa M, Briel S, et al. Mild cognitive impairment: incidence and risk factors: results of the leipzig longitudinal study of the aged. Journal of the American Geriatrics Society. 2010; 58:1903–1910. [PubMed: 20840461]

44. Roberts RO, Geda YE, Knopman DS, et al. The incidence of MCI differs by subype and is higher in men: The Mayo Clinic Study of Aging. Neurology. IN PRESS.

45. Unverzagt FW, Ogunniyi A, Taler V, et al. Incidence and risk factors for cognitive impairment no dementia and mild cognitive impairment in African Americans. Alzheimer disease and associated disorders. 2011; 25:4–10. [PubMed: 20921881]

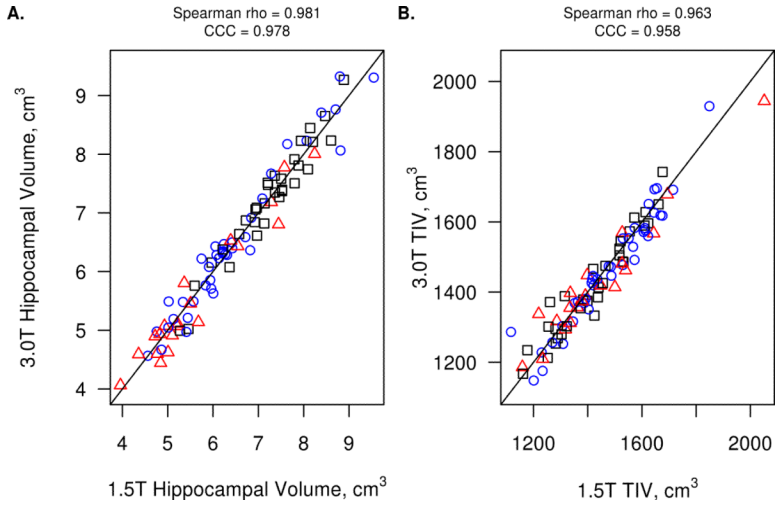


**Figure 1. A comparison of hippocampal volume measured from scans performed at 1.5T and 3.0T**

The comparison was performed using 91 subjects from ADNI that had both a 1.5T and 3.0T scan at the same visit (32 CN, 39 aMCI and 20 AD). Scatter-plots show the 3.0T vs. 1.5T hippocampal volume (Panel A) and TIV (Panel B). Different colors are used to represent each diagnostic group (CN=black, aMCI=blue, AD=red) and the identity line indicating perfect agreement is shown as a solid black line. The Spearman correlation (Spearman rho) and Lins' concordance correlation coefficient (ccc), a measure of intra-class correlation, are shown at the top of each plot. The data demonstrates an excellent agreement between 1.5T and 3.0T hippocampal volumes and TIV.

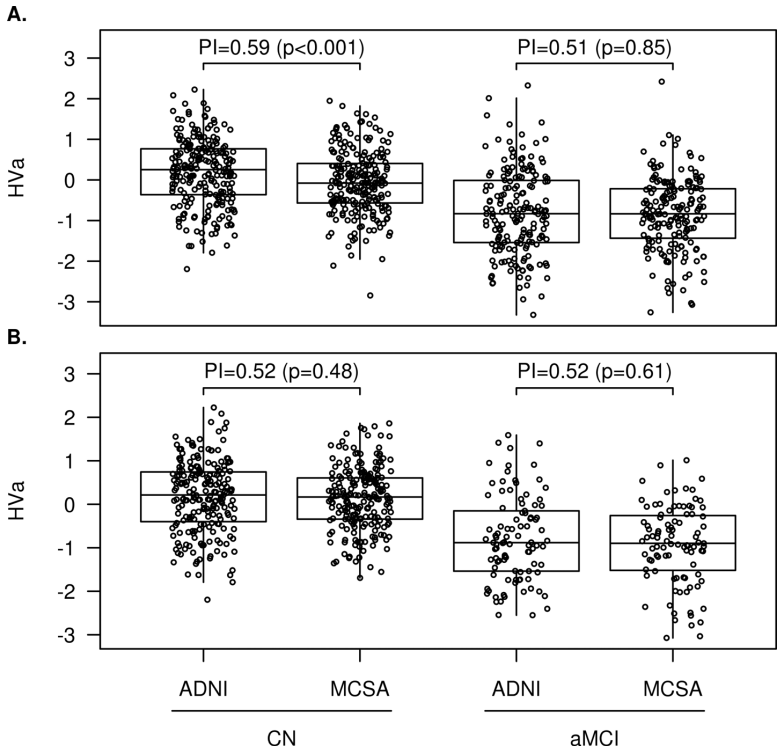


**Figure 2. Box-plots of hippocampal volumes in CN and aMCI subjects contrasting findings in the ADNI study with findings in the MCSA study**

Panel A shows the results in two simple random samples. Panel B shows the results in two samples frequency matched by age, sex, education, *APOE* genotype, and MMSE score. The boxes indicate the median and interquartile range (IQR) of the distributions while the vertical lines extending from the boxes stop at the most extreme data points within 1.5 IQRs. Because all individual points are shown, the points have been shifted randomly in the horizontal direction to avoid overlap and improve the visual display. We summarize group-wise difference using the Probabilistic Index (PI) and Wilcoxon rank-sum p values (shown in brackets). A PI of 0.50 indicates no difference across groups, whereas a PI of 0.60 indicates that 60% of the time the hippocampal volume from a random subject in ADNI is higher than the corresponding value in a random subject from the MCSA.

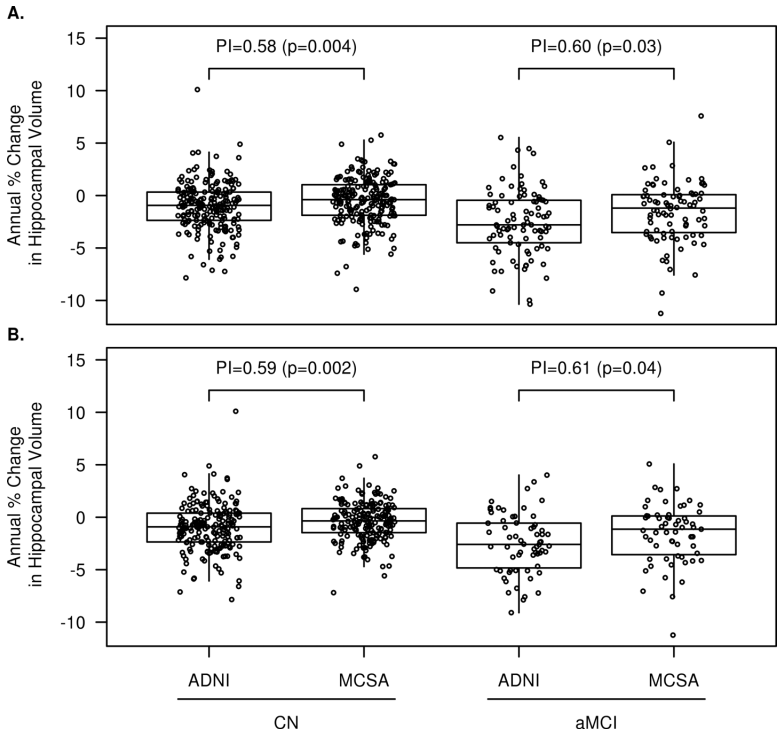


**Figure 3. Box-plots of annual percent decline in hippocampal volume in CN and aMCI subjects contrasting findings in the ADNI study with findings in the MCSA study**

Panel A shows the results in two simple random samples. Panel B shows the results in two samples frequency matched by age, sex, education, *APOE* genotype, and MMSE score. Negative values represent a decline in hippocampal volume over time. The boxes indicate the median and interquartile range (IQR) of the distributions while the vertical lines extending from the boxes stop at the most extreme data points within 1.5 IQRs. Because all individual points are shown, the points have been shifted randomly in the horizontal direction to avoid overlap and improve the visual display. We summarize group-wise difference using the Probabilistic Index (PI) and Wilcoxon rank-sum p values (shown in brackets). A PI of 0.50 indicates no difference across groups, whereas a PI of 0.60 indicates that 60% of the time the annual percent decline in hippocampal volume from a random subject in ADNI is greater than the corresponding value in a random subject from the MCSA.

## TABLE 1

Representativeness of the MCSA imaging sample

| Characteristic | MCSA No Imaging | MCSA Imaging | MCSA Combined |
|---|---|---|---|
| Number | 1112 | 1462 | 2574 |
| Age, years | 81 (75, 84) | 80 (75, 83) | 80 (75, 84) |
| Female Gender, no. (%) | 577 (51.9) | 698 (47.7) | 1275 (49.5) |
| Education, years | 13 (12, 16) | 13 (12, 16) | 13 (12, 16) |
| *APOE* ε4 positive, no. (%) | 240 (24.5) | 378 (25.9) | 618 (25.4) |
| Family history, no. (%)[ǂ] | 116 (10.7) | 189 (13.1) | 305 (12.1) |
| MMSE | 28 (26, 28) | 28 (27, 29) | 28 (27, 29) |
| Diagnosis, no. (%) | | | |
| CN | 932 (83.8) | 1283 (87.8) | 2215 (86.1) |
| aMCI | 180 (16.2) | 179 (12.2) | 359 (13.9) |

Median (inter-quartile range) shown unless otherwise noted. *APOE* ε4 genotype missing for 12% of MCSA subjects without imaging and 0.3% of MCSA subjects with imaging. Abbreviations: *APOE*, apolippoprotein E; MMSE, Mini Mental State Exam.

[ǂ]Data concerning family history of AD was not available in 44 subjects.

## TABLE 2

Descriptive characteristics of the two samples used for cross-sectional comparisons

| | CN | | aMCI | |
|---|---|---|---|---|
| **Characteristic** | **ADNI** | **MCSA** | **ADNI** | **MCSA** |
| **Sample 1: Simple Random Sample** | | | | |
| Number | 228 | 227 | 179 | 176 |
| Age, years | 76 (73, 79) | 79 (74, 83) *** | 76 (70, 80) | 81 (77, 85) *** |
| Women, no. (%) | 110 (48.2) | 116 (51.1) | 69 (38.5) | 68 (38.6) |
| Minority race, no. (%) † | 19 (8.3) | 2 (0.9) *** | 11 (6.1) | 4 (2.3) |
| Hispanic/latino, no. (%) | 2 (0.9) | 0 (0) | 7 (4.0) | 1 (0.6) |
| Education, years | 16 (14, 18) | 13 (12, 16) *** | 16 (13, 18) | 12 (12, 16) *** |
| *APOE* ε4 positive, no. (%) | 61 (26.8) | 55 (24.3) | 87 (48.6) | 63 (36.0) * |
| Family history, no. (%) | 72 (34.4) | 31 (13.8) *** | 60 (37.3) | 26 (15.5) *** |
| MMSE | 29 (29, 30) | 28 (27, 29) *** | 27 (26, 28) | 25 (24, 27) *** |
| Hippocampal volume, $cm^3$ | 7.3 (6.6, 7.9) | 7.1 (6.5, 7.5) * | 6.3 (5.6, 7.1) | 6.4 (5.7, 7.0) |
| TIV, $cm^3$ | 1437 (1322, 1544) | 1457 (1356, 1574) | 1432 (1351, 1573) | 1513 (1403, 1632) ** |
| HVa | 0.25 (−0.36, 0.77) | −0.08 (−0.57, 0.41) *** | −0.83 (−1.54, −0.01) | −0.83 (−1.43, −0.22) |
| **Sample 2: Age, Sex, Education, *APOE* Genotype, and MMSE Frequency Matched** | | | | |
| Number | 211 | 212 | 97 | 95 |
| Age, years | 76 (73, 79) | 76 (74, 79) *‡ | 80 (75, 84) | 80 (76, 84) |
| Women, no. (%) | 98 (46.4) | 98 (46.2) | 29 (29.9) | 28 (29.5) |
| Minority race, no. (%) † | 14 (6.6) | 2 (0.9) ** | 7 (7.2) | 1 (1.2) |
| Hispanic/latino, no. (%) | 2 (1.0) | 0 (0) | 1 (1.0) | 1 (1.1) |
| Education, years | 16 (14, 18) | 16 (14, 18) | 15 (12, 18) | 14 (12, 16) |
| *APOE* ε4 positive, no. (%) | 48 (22.7) | 48 (22.6) | 39 (40.2) | 38 (40.0) |
| Family history, no. (%) | 65 (33.9) | 34 (16.3) *** | 29 (33.3) | 15 (16.3) * |
| MMSE | 29 (29, 30) | 29 (29, 30) *‡ | 26 (25, 28) | 26 (24, 27) |
| Hippocampal volume, $cm^3$ | 7.3 (6.6, 7.9) | 7.3 (6.8, 7.8) | 6.4 (5.6, 7.2) | 6.4 (5.7, 7.1) |
| TIV, $cm^3$ | 1443 (1335, 1548) | 1481 (1359, 1584) | 1469 (1373, 1579) | 1543 (1434, 1644) ** |
| HVa | 0.21 (−0.40, 0.74) | 0.17 (−0.34, 0.61) | −0.88 (−1.54, −0.15) | −0.90 (−1.52, −0.26) |

Data are shown as median (inter-quartile range) unless otherwise stated. Random samples of MCSA CN and ADNI aMCI were used in sample 1 to ensure comparable group sizes. *APOE* ε4 genotype missing for 2 MCSA subjects in sample 1. MMSE scores were calculated from short test of mental status scores in the MCSA using an algorithm developed at our center. Abbreviations: *APOE*, apolipoprotein E; MMSE, Mini-Mental State Exam; TIV, total intracranial volume; HVa, hippocampal volume adjusted for TIV. Missing data: *APOE* ε4 genotype missing for 2 MCSA subjects in sample 1. MMSE was missing for 1 MCSA subject in sample 1. Race was unknown/not disclosed for 2 MCSA subjects in sample 1 and 1 MCSA subject in sample 2. Ethnicity was unknown/not disclosed for 7 subjects in sample 1 and 3 subjects in sample 2. Family history of AD was not available in 48 subjects from sample 1 and 35 subjects from sample 2. Significant differences observed across MCSA and ADNI within either CN or aMCI at

* $p<0.05$,

** $p<0.01$ and

*** p<0.001

‡ Despite frequency matching, a significant difference in MMSE scores and age was observed between ADNI and MCSA. However, the median and inter-quartile range of both MMSE and age was similar across cohorts.

† Minority race includes American Indian/Alaskan Native, Asian, and Black/African American, and more than one race.

## TABLE 3

Descriptive characteristics of the two samples used for longitudinal comparisons

| | CN | | aMCI | |
|---|---|---|---|---|
| **Characteristic** | **ADNI** | **MCSA** | **ADNI** | **MCSA** |
| **Sample 1: Simple Random Sample** | | | | |
| Number | 202 | 204 | 89 | 84 |
| Age, years | 76 (73, 79) | 78 (74, 82) *** | 74 (72, 81) | 81 (77, 84) *** |
| Women, no. (%) | 96 (47.5) | 93 (45.6) | 34 (38.2) | 28 (33.3) |
| Minority race, no. (%) ‡ | 16 (7.9) | 3 (1.5) ** | 4 (4.5) | 1 (1.2) |
| Hispanic/latino, no. (%) | 2 (1.0) | 0 (0) | 2 (2.2) | 1 (1.2) |
| Education, years | 16 (14, 18) | 14 (12, 16) *** | 16 (14, 18) | 12 (12, 16) *** |
| *APOE* ε4 positive, no. (%) | 58 (28.7) | 61 (29.9) | 47 (52.8) | 30 (35.7) * |
| Family history, no. (%) | 66 (35.3) | 33 (16.3) *** | 30 (36.6) | 14 (16.9) ** |
| MMSE | 29 (29, 30) | 28 (27, 29) *** | 27 (26, 29) | 25 (24, 27) *** |
| Annual % change in Hippocampal volume | −0.94 (−2.37, 0.32) | −0.39 (−1.87, 1.03) ** | −2.79 (−4.50, −0.45) | −1.20 (−3.48, 0.07) * |
| **Sample 2: Age, Sex, Education, *APOE* Genotype, and MMSE Frequency Matched** | | | | |
| Number | 187 | 187 | 64 | 59 |
| Age, years | 76 (73, 79) | 76 (74, 79) | 80 (77, 84) | 80 (76, 84) |
| Women, no. (%) | 86 (46.0) | 86 (46.0) | 21 (32.8) | 19 (32.2) |
| Minority race, no. (%) ‡ | 12 (6.4) | 1 (0.5) ** | 2 (3.1) | 0 (0) |
| Hispanic/latino, no. (%) | 2 (1.1) | 0 (0) | 1 (1.6) | 1 (1.7) |
| Education, years | 16 (14, 18) | 16 (14, 18) | 14 (12, 17) | 14 (12, 16) |
| *APOE* ε4 positive, no. (%) | 47 (25.1) | 46 (24.6) | 21 (32.8) | 20 (33.9) |
| Family history, no. (%) | 57 (33.1) | 27 (14.6) *** | 23 (38.3) | 11 (19.0) * |
| MMSE | 29 (29, 30) | 29 (29, 29) **† | 26 (25, 27) | 26 (24, 27) |
| Annual % change in | −0.92 (−2.36, 0.39) | −0.35 (−1.47, 0.82) ** | −2.59 (−4.75, −0.56) | −1.14 (−3.56, 0.12) * |

Data are shown as median (inter-quartile range) unless otherwise stated. Random samples of MCSA CN and ADNI aMCI were used in sample 1 to ensure comparable group sizes. MMSE scores were calculated from short test of mental status scores in the MCSA using an algorithm developed at our center. Abbreviations: *APOE*, apolipoprotein E; MMSE, Mini-Mental State Exam; TIV, total intracranial volume; IQR, inter-quartile range. Missing data: Race is unknown/not disclosed for 2 MCSA subjects in sample 1. Ethnicity is unknown/not disclosed for 5 subjects in sample 1 and 3 ADNI subjects in sample 2. MMSE is missing for 2 MCSA subjects in sample 1. Family history of AD was not available for 24 subjects in sample 1 and 22 subjects in sample 2. Significant differences observed across MCSA and ADNI within either CN or aMCI at

* p<0.05,

** p<0.01 and

*** p<0.001

† Despite frequency matching by MMSE strata, a significant difference in MMSE scores was observed between ADNI and MCSA. However, the median and inter-quartile range was similar across cohorts

‡ Minority race includes American Indian/Alaskan Native, Asian, and Black/African American, and more than one race.