Supplemental Materials

## Sample Recruitment and Clinical Characterization

Participants were evaluated and followed within the research program of the University of California at Davis Alzheimer’s Disease Center (UCD ADC). About two thirds of this sample were recruited through community-based recruitment protocols designed to enhance both the racial and ethnic diversity and the spectrum of cognitive dysfunction of the sample with an emphasis on normal cognition and MCI. Recruiters utilized various outreach methods such as soliciting in a community hospital lobby, a community survey, health fairs or word of mouth. Community recruitment methods have been previously described (Hinton et al. 2010). The other one third of the sample initially sought a clinical evaluation at the UCD ADC and subsequently were recruited for this study. These individuals predominantly had a clinical diagnosis of MCI. Enrollment began in 2001 and a rolling enrollment design was used to build and maintain the cohort. Inclusion criteria included age 60 or older at their first examination and ability to speak English or Spanish. Exclusion criteria included unstable major medical illness, major primary psychiatric disorder, and substance abuse or dependence in the last five years. Participants received clinical evaluations through the UC Davis Alzheimer’s Disease Center on a roughly annual basis that included diagnosis, based on standard diagnostic criteria, of normal cognition versus mild cognitive impairment (MCI) versus dementia as well as etiologic diagnosis.

## Cognitive Assessment

Item response theory and confirmatory factor analysis methods were used to construct Spanish and English Neuropsychological Assessment Scales (SENAS) measures that are psychometrically matched across domains in terms of level of reliability across the ability continuum. Importantly, these measures do not have floor and ceiling effect and are normally distributed. The episodic memory score is derived from a multi-trial word-list-learning test (Mungas et al. 2004). The semantic memory measure is a composite of highly correlated verbal (object-naming) and nonverbal (picture association) tasks. The executive function composite is constructed from component tasks of category fluency, phonemic (letter) fluency, and working memory (digit-span backward, visual-span backward, list sorting). Spatial ability was measured using the SENAS Spatial Localization scale which assesses ability to perceive and reproduce two-dimensional spatial relationships that are increasingly complex. Language of test administration was determined by an algorithm that combined information regarding each participant’s language preference in several specific contexts (e.g., conversing at home, listening to radio or television, conversing outside the home, preferred language for reading). Administration procedures, measure development and psychometric characteristics of the SENAS battery are described in detail elsewhere (Mungas et al. 2004).

## Clinical Diagnosis

All participants received multidisciplinary diagnostic evaluations at baseline and at approximately annual intervals following the baseline evaluation. Baseline and follow-up evaluations followed the same protocol with a detailed medical history, physical and neurological exam, and clinical neuropsychological assessment. A physician fluent in Spanish examined subjects who spoke only Spanish. A family member or other informant was interviewed to obtain information about cognitive and independent functioning. Clinical neuropsychological tests were different from the cognitive measures used in analyses in this study to estimate reserve and longitudinal cognitive trajectories. Routine dementia work-up laboratory tests were obtained at the baseline evaluation and when clinically indicated at the time of follow-up evaluations.

Diagnosis of cognitive syndrome (Normal, mild cognitive impairment (MCI), Dementia) and, for individuals with dementia, underlying etiology, was made in a multidisciplinary consensus conference following standardized criteria and methods. Dementia was diagnosed using DSM-III-R (Association 1987) criteria for dementia modified such that dementia could be diagnosed in the absence of memory impairment if there was significant impairment of two or more other cognitive domains. MCI was diagnosed according to standard clinical criteria and was further sub-typed according to current Alzheimer’s Disease Centers Uniform Data Set guidelines (Morris 2006). Normal cognitive function was diagnosed if there was no clinically significant cognitive impairment. All diagnoses were made blind to the neuropsychological tests that were analyzed in this study.

At baseline, there were 49 single domain amnestic MCI cases, 29 multiple domain amnestic MCI, 18 single domain non-amnestic MCI, and 11 multiple domain non-amnestic MCI. Etiologic diagnosis for individuals with dementia at baseline was probable Alzheimer’s disease (AD) for 15, possible AD for 5, and mixed AD and cerebrovascular disease for 1. Seventy Two (67.3%) of individuals with a baseline diagnosis of MCI converted to dementia by the final assessment. Forty Five (24.1%) individuals who were cognitively normal at baseline converted to MCI and 29 (15.5%) converted to dementia over the course of follow-up. These results show substantial clinical progression over the course of follow-up and highlight the heterogeneity of clinical trajectories in this sample.

## Baseline Brain Measurements

Structural MRI images were processed to remove the skull using an atlas-based method (Aljabar et al. 2007, 2009) followed by human quality control to provide generally minor cleanup if needed. MRI brain images were then nonlinearly registered to a minimal deformation template (MDT) synthetic brain image (Kochunov et al. 2001) adapted for age range of 60 and above; the registration was performed by a cubic B-spline deformation (Rueckert et al. 2006). Gray, white and CSF tissues segmentation was initiated automatically by reverse transforming the MDT segmentation onto native structural MRIs using the computed registration parameters. This was the start of an iterative maximal likelihood estimation of tissue classes based upon alternating voxel class assignment followed by tissue class parameter estimation until convergence. The class likelihood priors included terms designed to enhance accuracy at likely tissue boundaries (Fletcher et al. 2012). Finally, native lobar gray matter volumes were computed by reverse transforming MDT lobar ROIs into native space using the B-spline registration parameters.

## Gray Matter Change

TBM generates deformation fields by registering brain scans at differing time points and using these to estimate local volume changes between the scans (Ashburner and Friston 2000) via the log-transformed determinant of the Jacobian matrix for the deformation field at each voxel. The Jacobian determinant gives the local volume change (Hua et al. 2008) factor in ranges from 0-1 for contractions and > 1 for expansions. Because these generate skewed distributions, the log-transform is used, giving a zero mean symmetric distribution (Yanovsky et al. 2009) with negative values indicating contradictions and positive value expansions. Log-Jacobians represent approximate percentage volume changes for small magnitudes of change. Native hippocampal mask segmentation was accomplished using a multi-atlas matching algorithm (Aljabar et al. 2009) from atlas image hippocampi carefully segmented according to the EADC-ADNI harmonized hippocampal protocol (Frisoni and Jack 2015, 2011) followed by human analyst quality control as needed. All log-Jacobian measures of change rates were normalized to represent change over an interscan interval of two years, under the assumption that atrophy rates are roughly log-linear.

## Modeling of Cognition Random Intercepts and Slopes

We compared a series of models to determine whether cognitive intercepts and slopes could be summarized by second order factors. These were unconditional models that did not include covariates or independent variables. The initial model included correlated intercept and slope random effects for each of the four cognitive outcomes, but we then evaluated whether second order latent variables (one with intercepts as indicators, one with slopes) explained the correlations among the random effects. The second order factors were identified by fixing one loading to 1.0 and freely estimating the other loadings. We compared the fits of models with 0, 1, and 2 second order factors using comparative fit indices including the Akaike Information Criterion (AIC) (Akaike 1987), the Bayesian Information Criterion (BIC) (Schwarz 1978), and the Sample Size Adjusted Bayesian Information Criterion (aBIC) (Sclove 1987). These indices differ in the relative weighting of model fit and model parsimony with AIC valuing parsimony the least and BIC the most. Lower values on all indices indicate better model fit. The best fit was obtained with the model that had a global slope second order factor but individual intercept random effects.

## References

Akaike, H. 1987. “Factor Analysis and AIC.” *Psychometrika* 52: 317–32.

Aljabar, P., R. a Heckemann, a Hammers, J. V. Hajnal, and D. Rueckert. 2009. “Multi-Atlas Based Segmentation of Brain Images: Atlas Selection and Its Effect on Accuracy.” *Neuroimage* 46: 726–38. <https://doi.org/10.1016/j.neuroimage.2009.02.018>.

Aljabar, P., R. Heckemann, A. Hammers, J. V. Hajnal, and D. Rueckert. 2007. “Classifier Selection Strategies for Label Fusion Using Large Atlas Databases.” *Med Image Comput Comput Assist Interv* 10 (Pt 1): 523–31.

Ashburner, J., and K. J. Friston. 2000. “Voxel-Based Morphometry–the Methods.” *Neuroimage* 11 (6 Pt 1): 805–21. <https://doi.org/10.1006/nimg.2000.0582>.

Association, American Psychiatric. 1987. *Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised*. Washington, DC: American Psychiatric Association.

Fletcher, E., B. Singh, D. Harvey, O. Carmichael, and C. DeCarli. 2012. “Adaptive Image Segmentation for Robust Measurement of Longitudinal Brain Tissue Change.” *Conf Proc IEEE Eng Med Biol Soc* 2012: 5319–22. <https://doi.org/10.1109/EMBC.2012.6347195>.

Frisoni, Giovanni B., and Clifford R. Jack. 2011. “Harmonization of Magnetic Resonance-Based Manual Hippocampal Segmentation: A Mandatory Step for Wide Clinical Use.” *Alzheimer’s & Dementia: The Journal of the Alzheimer’s Association* 7 (2): 171–74. <https://doi.org/10.1016/j.jalz.2010.06.007>.

———. 2015. “HarP: The EADC-ADNI Harmonized Protocol for Manual Hippocampal Segmentation. A Standard of Reference from a Global Working Group.” *Alzheimer’s & Dementia: The Journal of the Alzheimer’s Association* 11 (2): 107–10. <https://doi.org/10.1016/j.jalz.2014.05.1761>.

Hinton, L., K. Carter, B. R. Reed, L. Beckett, E. Lara, C. DeCarli, and D. Mungas. 2010. “Recruitment of a Community-Based Cohort for Research on Diversity and Risk of Dementia.” *Alzheimer’s Disease and Associated Disorders* 24 (3): 234–41. <https://doi.org/10.1097/WAD.0b013e3181c1ee01>.

Hua, X., A. Leow, N. Parikshak, S. Lee, M. Chiang, A. Toga, C. Jackjr, M. Weiner, and P. Thompson. 2008. “Tensor-Based Morphometry as a Neuroimaging Biomarker for Alzheimer’s Disease: An MRI Study of 676 AD, MCI, and Normal Subjects.” *Neuroimage* 43 (3): 458–69. <https://doi.org/10.1016/j.neuroimage.2008.07.013>.

Kochunov, P., J. L. Lancaster, P. Thompson, R. Woods, J. Mazziotta, J. Hardies, and P. Fox. 2001. “Regional Spatial Normalization: Toward an Optimal Target.” *J Comput Assist Tomogr* 25 (5): 805–16.

Morris, J. C. 2006. “Mild Cognitive Impairment Is Early-Stage Alzheimer Disease: Time to Revise Diagnostic Criteria.” *Archives of Neurology* 63 (1): 15–16. [https://doi.org/63/1/15 [pii] 10.1001/archneur.63.1.15](https://doi.org/63/1/15%20%5Bpii%5D%2010.1001/archneur.63.1.15).

Mungas, D., B. R. Reed, P. K. Crane, M. N. Haan, and H. Gonzalez. 2004. “Spanish and English Neuropsychological Assessment Scales (SENAS): Further Development and Psychometric Characteristics.” *Psychol Assess* 16 (4): 347–59. <https://doi.org/10.1037/1040-3590.16.4.347>.

Rueckert, Daniel, Paul Aljabar, Rolf A. Heckemann, Joseph V. Hajnal, and Alexander Hammers. 2006. “Diffeomorphic Registration Using B-Splines.” In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006.*, edited by R. Larsen, M. Nielsen, and J. Sporring, 4191:702–9. Lecture Notes in Computer Science-MICCAI. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.

Schwarz, G. 1978. “Estimating the Dimension of a Model.” *Annals of Statistics* 6: 461–64.

Sclove, S. L. 1987. “Application of Model-Selection Criteria to Some Problems in Multivariate Analysis.” *Psychometrika* 52: 333–43.

Yanovsky, Igor, Alex D. Leow, Suh Lee, Stanley J. Osher, and Paul M. Thompson. 2009. “Comparing Registration Methods for Mapping Brain Change Using Tensor-Based Morphometry.” *Medical Image Analysis* 13 (5): 679–700. <https://doi.org/10.1016/j.media.2009.06.002>.

##### 

Table 1. Correlations of SENAS cognitive intercept and slope random effects.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | vm\_int | sem\_int | exec\_int | spat\_int | vm\_sl | sem\_sl | exec\_sl | spat\_sl |
| **vm\_int** |  |  |  |  |  |  |  |  |
| **sem\_int** | 0.484 |  |  |  |  |  |  |  |
| **exec\_int** | 0.732 | 0.667 |  |  |  |  |  |  |
| **spat\_int** | 0.476 | 0.763 | 0.662 |  |  |  |  |  |
| **vm\_sl** | 0.525 | 0.041 | 0.225 | -0.026 |  |  |  |  |
| **sem\_sl** | 0.597 | 0.09 | 0.289 | 0.013 | 0.952 |  |  |  |
| **exec\_sl** | 0.586 | 0.03 | 0.271 | -0.023 | 0.978 | 0.987 |  |  |
| **spat\_sl** | 0.491 | -0.072 | 0.192 | -0.146 | 0.954 | 0.974 | 0.977 |  |

##### 

Table 2. Fit indices of alternate models to characterized covariance among cognitive intercepts and slopes.

|  |  |  |  |
| --- | --- | --- | --- |
| Model | AIC | BIC | aBIC |
| Separate Intercepts - Separate Slopes | 14105 | 14512 | 14287 |
| Global Intercept - Separate Slopes | 14231 | 14523 | 14361 |
| **Separate Intercepts - Global Slope** | **14104** | **14368** | **14222** |
| Global Intercept - Global Slope | 14227 | 14457 | 14330 |

Note: AIC = Akaike Information Criterion, BIC = Bayesian Information Criterion, aBIC = Sample Size Adjusted Bayesian Information Criterion.