10.1148/radiol.2019182726

Appendix E1

The Study Cohort and Consensus Diagnosis Procedures

The entire BIOCARD cohort consisted of 168 participants (age: 70.1 ± 8.6, 68M/100F). Of these, 24 were diagnosed as MCI (age: 74.6 ± 8.7, 10M/14F) in their most recent clinical visits and one was diagnosed as dementia (age 88, F). The remaining participants (n = 143, age: 69.1 ± 8.2, 57M/86F) were considered cognitively healthy and only these were included in this study. Approximately 82% (117/143) of the participants had a first-degree relative with dementia of the Alzheimer type and 35% (50/143) were APOE4 carriers, which is slightly higher than the rate in the general population of 26% (32). There was not a difference in sex distribution between the healthy and MCI/dementia groups (P = .35), but the MCI/dementia participants were on average older than the healthy group (P < .001).

Each study participant included in these analyses received a consensus diagnosis by the staff of the BIOCARD clinical team. This research team included: neurologists, neuropsychologists, research nurses and research assistants. During the study visit, each study participant had received a comprehensive cognitive assessment and a Clinical Dementia Rating (CDR), as well as a comprehensive medical evaluation (including a medical, neurologic and psychiatric assessment). Two sources of information were used to determine if the participant was cognitively healthy or if the participant met clinical criteria for the syndromes of MCI or dementia: (1) the CDR interview conducted with the participant and the collateral source was used to determine if there was evidence that the participant was demonstrating changes in cognition in daily life, (2) cognitive tests scores (and their comparison to established norms) were used to determine if there was evidence of significant decline in cognitive performance over time. If a participant was deemed to be impaired, the decision about the likely etiology of the syndrome was based on the medical, neurologic, and psychiatric information collected at each visit, as well as medical records obtained from the participant, where necessary. More than one etiology could be endorsed for each participant (eg, Alzheimer disease and vascular disease). One of four possible diagnostic categories was selected at each visit for each participant: (1) Healthy, (2) Mild Cognitive Impairment, (3) Impaired Not MCI or (4) Dementia. These diagnostic procedures are comparable to those implemented by the Alzheimer's Disease Centers program supported by the National Institute on Aging. The diagnostic team had no information about the results of the MRI analyses.

APOE genotype was determined by restriction endonuclease digestion of polymerase chain reaction-amplified genomic DNA (performed by Athena Diagnostics, Worcester, MA).

Application of Confirmatory Factor Analysis and Calculation of Cognitive Composite Scores

Application of Confirmatory Factor Analysis

The annual cognitive assessments consisted of a comprehensive battery of standardized neuropsychological tests covering a broad range of cognitive domains, including memory, executive function, language, visuospatial ability, attention, speed of processing and

psychomotor speed (for more details, see Albert et al [17]). This battery is comparable to the neuropsychological test battery implemented in the Alzheimer's Disease Centers program, supported by the National Institute on Aging (20). To reduce the amount of cognitive data, we applied confirmatory factor analysis (CFA) to cognitive task scores obtained from a participant's first visit to Johns Hopkins (referred to here as visit 101). These analyses included data from 265 nondemented participants with complete data on all tasks. This number excludes $n = 5$ participants with missing data on at least one task, and $n = 7$ participants with a diagnosis of dementia.

The following data processing and transformations occurred prior to model fitting. First, task scores were examined for distributional normality (skew, kurtosis). Outliers falling beyond the interquartile range and 2.5 standard deviations beyond the overall mean were replaced with values that were 2.5 standard deviations beyond the mean ($n = 25$ scores). Task scores were then converted to z-scores, and where relevant, z-scores were reverse scored so that higher scores always reflected better performance.

The CFA was based on an a priori selection and categorization of cognitive tasks by two cognitive psychologists (AS, CP), with three tasks selected for each latent variable. Model fit was evaluated with several fit indexes. The χ^2 goodness-of-fit statistic assessed the discrepancy between the sample and fitted covariance matrices; for this index, small, nonsignificant values indicate good fit (33), though several shortcomings of this statistic have been identified (34). We therefore also examined the χ^2/df ratio, for which values less than 3 are indicative of acceptable model fit (34). Model fit was also evaluated with Bentler's comparative fit index (CFI), the standardized root-mean-square residual (SRMR), and the root mean square error of approximation (RMSEA). CFI is an incremental fit index that ranges from 0–1; values > 0.95 indicate good fit (35). Both SRMR and RMSEA are absolute fit indexes; for both, lower values indicate acceptable fit (for SRMR, < 0.08; for RMSEA, < 0.06 with a corresponding nonsignificant *P* value). Nested models were compared quantitatively by calculating the change in χ^2 values across models; if the fuller, more complex model has a change in χ^2 that is significant given the loss of degrees of freedom, it was accepted as having better fit. CFA models were estimated in R (version 3.4.2; R Project for Statistical Computing) using the lavaan (latent variable analysis) package (36).

The four factor model included the following latent variables, with corresponding manifest variables shown in parentheses: (1) verbal episodic memory (Logical Memory delayed recall from the Wechsler Memory Scale-Revised (WMS-R) (8), Paired Associates immediate recall from the WMS-R (37), California Verbal Learning Test recall over trials 1–5 [38]); (2) executive function (Digit Span backward from the WMS-R (37), Trail Making Test part B (39), Digit Symbol Test from the Wechsler Adult Intelligence Scale–Revised (WAIS-R) [37]); (3) visuospatial processing (Rey-Osterreith Complex Figure copy (40), Rey Figure recall (40), Block Design subtest of the WAIS-R [41]); and (4) language (Boston Naming Test (42), Letter (FAS) Fluency (43), Category (animal) Fluency [43]). Rey Copy and Rey Recall error variances were allowed to compare given these variables reflect two measures from the same task. This four-factor model provided an acceptable fit to the data (χ^2 (df) = 93.69 (47), $P < .001$; χ^2/df = 1.99; CFI = 0.96, SRMR = 0.05, RMSEA = 0.06, $P = .15$), and provided a significantly better fit to the data than the three, two, and one factor nested models (as indicated by a significant change in χ^2; all $P < .001$, data not shown). The four-factor model was therefore selected as the final model (Fig E1).

Calculation of Cognitive Composite Scores

Summary factor scores, referred to here as cognitive composite scores, were created for the four cognitive domains listed above. To create the composite scores, task scores at each visit were converted to z-scores using the means and standard deviations from all available cognitive data at visit 101. Z-scores for Trails B were reverse scored. For each participant, for each visit, (1) individual task z-scores were weighted by their respective standardized factor loadings (see Fig E1), and (2) weighted scores were summed within each cognitive domain to create the four composite scores. If a participant did not have scores for all three tasks within a cognitive domain, the composite score was set to missing for that visit. The composite scores used in the present analyses reflect data obtained at the same visit as the MRI scan, which occurred between 2015 and 2017.

Vascular Risk Assessment

Medical history was recorded during the same visit of MRI scan. Five vascular risk factors were considered and coded as binary variables as follows: (1) hypertension (1 if recent, 0 if remote/absent), (2) hypercholesterolemia (1 if recent, 0 if remote/absent), (3) diabetes (1 if recent, 0 if remote/absent), (4) smoking (1 if smoked over 100 cigarettes in his/her life, 0 if not), and (5) body mass index (BMI; 1 if BMI > 30, 0 if not). BMI was calculated as weight in kilograms divided by height in meters squared. Vascular Risk Score (VRS) was calculated as the sum of the measures above and could range from 0 to 5.

MRI Image Sequence and Processing

TRUST MRI is based on the principle that transverse relaxation time (T_2) of the blood has a well-known and calibratable relationship with Y_v. The sequence applies venous spin labeling to isolate venous blood signal (Fig 1), which is followed by a series of T_2 preparation pulses to estimate pure blood T_2. The imaging parameters for TRUST MRI were as follows: TR = 3000 ms, TE = 3.61 ms, inversion time (TI)=1022 ms, flip angle = 90°, FOV = 220 × 220 × 5 mm^3, voxel size = 3.44 × 3.44 × 5 mm^3, four eTEs (1, 40, 80, and 160 ms) with a τ_{CPMG} of 10 ms, labeling thickness = 10 mm, and scan duration = 1.2min.

Subtraction between images with and without labeling RF pulse yields a difference image with pure venous blood signal. Regions of interest (ROIs) were drawn manually on the difference image which contains the SSS. Then the four voxels with the highest signal intensity were chosen and the averaged signal were fitted to the equation to yield blood T_2. The blood T_2 was converted to Y_v using a calibration curve, in which the hematocrit level was assumed to be 0.42 for men and 0.40 for women (44,12). Then OEF was calculated.

Additionally, a 3D T_1-weighted Magnetization-prepared-rapid-acquisition-of-gradient-echo (MPRAGE) scan was performed for brain volume quantification. The imaging parameters of MPRAGE scan were as follows: repetition time (TR)=6.7 ms, echo time (TE)=3.1 ms, shot interval = 3000 ms, flip angle = 8°, voxel size = 1 × 1 × 1.2 mm^3, number of slices = 170, sagittal orientation, and scan duration = 5min 59s. To investigate the effect of APOE4 on brain structure, we parcellated the MPRAGE image using an automatic processing tool, MRICloud (www.MRICloud.org) (22). Volumetric measures of neocortical gray matter of the frontal lobe, parietal lobe, temporal lobe, occipital lobe, as well as limbic areas including hippocampus, amygdala and entorhinal cortex were obtained. A Fluid-Attenuated Inversion Recovery (FLAIR)

image was also acquired with the following parameters: TR = 11000 ms, TE = 100 ms, flip angle = 90°, voxel size = 1 × 1 × 2 mm^3, number of slices = 69, axial orientation, and scan duration = 3min 18s.

PiB-PET Imaging Processing for Amyloid Quantification

Study participants received a bolus injection of a mean (standard deviation) 14.7 (0.8) mCi of ^{11}C-labeled Pittsburgh Compound B tracer (^{11}C-PiB). PET images were acquired with dynamic acquisition frames as follows: 4 × 15 seconds, 8 × 30 seconds, 9 × 1 minute, 2 × 3 minutes and 10 × 5 minutes for a total of 70 minutes with 33 frames. PET data were reconstructed using filtered back-projection with a ramp filter (image size = 128 × 128, pixel size = 2 × 2 mm, slice thickness = 4.25 mm), yielding a spatial resolution of approximately 4.5 mm full width at half maximum (FWHM).

Distribution volume ratio (DVR) images were generated by a method reported by Bilgel et al (23). Mean cortical DVR (cDVR) was calculated in the native space of each PET image using the simplified reference tissue model with the cerebellar gray matter as reference tissue (45). The brain parcellation map generated by the MRICloud (22) on the T_1-weighted MRI space was transformed on the DVR image by applying the transformation matrix obtained from the rigid transformation of the T1-weighted image to the 20-minute average PET image. The cDVR was calculated as the average of the DVR values of the cingulate, frontal, parietal, lateral temporal, and the lateral occipital cortical regions, excluding the sensorimotor strip. The cDVR was used as an index of cortical ^{11}C-PiB retention.

References

32. Maiti TK, Konar S, Bir S, Kalakoti P, Bollam P, Nanda A. Role of apolipoprotein E polymorphism as a prognostic marker in traumatic brain injury and neurodegenerative disease: a critical review. Neurosurg Focus 2015;39(5):E3.

33. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct Equ Modeling 1999;6(1):1–55.

34. Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. Methods Psychol Res 2003;8(2):23–74.

35. Blunch NJ. Introduction to Structural Equation Modeling using IBM SPSS Statistics and AMOS. 2nd ed. London, England: Sage, 2008.

36. Rosseel Y. lavaan: An R Package for Structural Equation Modeling. J Stat Softw 2011;48(2):1–36.

37. Wechsler DW. Wechsler Memory Scale - Revised Manual. San Antonio, Tex: Psychological Corporation, 1987.

38. Delis DC, Kramer JH, Kaplan E, Ober BA. California Verbal Learning Test. San Antonio, Tex: The Psychological Corporation, 1987.

39. Reitan RM. Validity of the Trail Making Test as an indicator of organic brain damage. Percept Mot Skills 1958;8(3):271–276.

40. Rey A. L'examen psychologique dans les cas d'encephalopathie traumatique. Arch Psychol 1941;28:286–340.

41. Wechsler DW. Wechsler Adult Intelligence Scale - Revised Manual. New York, NY: The Psychological Corporation, 1981.

42. Kaplan E, Goodglass H, Weintraub S. Boston Naming Test. Philadelphia, Pa: Lee & Febiger, 1983.

43. Benton AL, Hamsher K. Multilingual Aphasia Examination. Iowa City, Iowa: University of Iowa Press, 1976.

44. Liu P, Xu F, Lu H. Test-retest reproducibility of a rapid method to measure brain oxygen metabolism. Magn Reson Med 2013;69(3):675–681.

45. Zhou Y, Resnick SM, Ye W, et al. Using a reference tissue model with spatial constraint to quantify [11C]Pittsburgh compound B PET for early diagnosis of Alzheimer's disease. Neuroimage 2007;36(2):298–312.