***White matter lesions: spatial heterogeneity, links to risk factors, cognition, genetics, atrophy***

## Online Supplementary Materials

### Supplementary section e-1: Data assessment and laboratory work in SHIP

The Institute for Community Medicine at the Medical Faculty of the University of Greifswald2 led SHIP, a population-based prospective cohort representative of the German northern east of Pomerania. SHIP started at baseline with the cohort SHIP-0 between 1997 and 2001. From 2008 to 2013 the second follow-up examination SHIP-2 was carried out. At the same time with SHIP-2 a new sample from the same region and recruitment criteria was drawn and similar examinations were undertaken between 2008 and 2012 (SHIP-Trend). SHIP-2 and SHIP-Trend included whole-body MRI scans and neuroimaging components.

Clinical data were collected by a computer assisted face-to-face interview. Smoking was divided in three categories, specifically: current smoking, former smoking, and never smoked. Having completed the interview, participants underwent medical examinations: including the measurement of height and weight (continuous variable). Waist circumference was measured in cm (continuous variable). After a 5-min resting period, blood pressure was measured three times on the right arm of seated subjects using a digital blood pressure monitor (HEM-705CP, Omron, Tokyo, Japan), with each reading being followed by a further resting period of 3 min. Cuffs were applied according to the circumference of the participant's arm. The mean of the second and third measurements (mm  Hg) was used for the analyses (continuous variables). All subjects were informed to bring in their packing containers of all medication they had taken during the last 7 days, as well as their drug prescription sheets. Every compound was recorded. We used antihypertensive, antidiabetic and lipid-lowering drugs, as indicators for cardivascular risk factors in the general population. High-density lipoprotein cholesterol concentrations were measured photometrically (Hitachi 704, Roche, Mannheim, Germany). Total cholesterol, low-density lipoprotein and high-density lipoprotein were measured as dimensional scores.

### Supplementary section e-2: Polygenic risk score calculation

For the calculation of the AD polygenic risk score, we used 19 single nucleotide polymorphisms found in the study of Lambert et al. that included 17 008 cases and 37 154 controls3. Then the AD polygenic risk score was calculated for SHIP participants with the available genotyping (n=1837) as the sum of the number of alleles of every single nucleotide polymorphism weighted by the logarithm of the corresponding odds ratio4. Higher polygenic risk score is directly proportional to the risk for AD.

### Supplementary section e-3: Determination of participants with possible cognitive impairment in SHIP

In SHIP two cognitive tests were obtained: the verbal learning and memory test (the German version for California verbal learning and memory test 5) for the sub-cohort SHIP-2 (n=730) and the Nurnberg age inventory for the sub-cohort SHIP-Trend (n=1,637), respectively. Nurnberg age inventory is a German test developed to measure the cognition abilities during brain aging 6. The verbal learning and memory test and Nurnberg age inventory consist of subtests including immediate and delayed memory tests. Accordingly, we labeled subjects with possible cognitive impairment to their age-adjusted cognitive scores (n=178)4,7,8.

### Supplementary section e-4: Image acquisition

In SHIP, a single 1.5 Tesla Siemens MRI scanner (Magnetom Avanto, Siemens Healthineers, Erlangen, Germany) was used to obtain the T1-weighted images with an axial MPRAGE sequence, with the following parameters: 1×1 mm in-plane spatial resolution, slice thickness=1.0 mm (flip angle 15°), TE=3.4 ms and TR=1900 ms. Additionally we used an axial T2-FLAIR with the following acquisition parameters: 0.9×0.9 mm in-plane spatial resolution, slice thickness=3.0 mm, TE=3250 ms and TR=50000 ms. Imaging parameters for the BLSA are detailed in 1. MRI was acquired on two similarly configured Philips 3T scanners (n=705 and n=42 observations). T1-weighted images with a sagittal MPRAGE sequence were acquired with the following parameters: 1×1 mm in-plane spatial resolution, slice thickness=1.20 mm (flip angle 8°), TE=3.1 ms and TR=6.5 ms (TR=6.8 ms for the second scanner). A transverse T2-FLAIR sequence was acquired with the following acquisition parameters: 0.8×0.8 mm in-plane spatial resolution, slice thickness=3 mm (slice thickness=4.4 mm for the second scanner), TE=68 ms and TR=11000 ms.

### Supplementary section e-5: MR images pre-processing

WMH were segmented using a multi-modal supervised learning-based segmentation method9, applied on co-registered FLAIR and T1 images for each subject. The T1 image of each subject was non-linearly aligned to a standardized brain atlas, and regional volumetric maps, named RAVENS maps, were calculated for WMH. WMH segmentation and the generation of WMH RAVENS maps have been described in detail previously7. All WMH segmentation results were visually inspected for high quality in SHIP by M.H. and in BLSA by G.E. Structural T1-weighted MRI images were preprocessed using the protocols of the ENIGMA consortium, in which the cortical parcellations were performed with the available FreeSurfer software package (version 5.3) 10. A cortical surface model was calculated for every subject, and cortical thickness was measured across the entire cortical mantle 11. Cortical surfaces were inflated and normalized to the fsaverage5 template via a spherical registration 12. The spatially normalized cortical thickness maps were smoothed using an isotropic Gaussian filter kernel with full width at half maximum size of 20 mm. In order to ensure the high quality of the data, standardized ENIGMA protocols (http://enigma.ini.usc.edu/protocols/imaging-protocols/) were implemented to scrutinize the thirty-four cortical gray matter structures that were labeled per hemisphere using the Desikan-Killiany atlas 13 along with the two whole-hemisphere measures.

### Supplementary section e-6: Identification of WMH components via non-negative matrix factorization

Capitalizing on recent advances in multivariate analysis methods, we used non-negative matrix factorization to uncover WMH components (WMHC). Non-negative matrix factorization decomposes high dimensional imaging data into a lower-dimensional space by summarizing the complex multivariate patterns of co-variation in the data with a set of non-negative components. Each component encodes a distinct pattern by positively weighting elements that co-vary, thus aggregating variance and approximating the data. Each data sample is approximated by a subject-specific, additive combination of non-negative components. The subject-specific coefficients encode the effect of the associated components in reconstructing the data and allow for subsequent group analyses. Even though the coefficients are not the original WMH volumes, they are average volumes, and hence have the unit milliliter (ML). Importantly, the non-negativity constraints result in a cluster-based representation characterized by advantageous specificity, which is comparable to traditional region of interests analysis, and improves generalizability14. The non-negative matrix factorization method does not explicitly constrain the components to be spatially distinct. However, it has the property that it results in final components that have significantly higher sparsity (e.g. 0 weights in many voxels) and that are highly distinct, in comparison to more traditional data-driven decomposition approaches like the principle component analysis and the independent component analysis. This property essentially comes from the non-negativity constraint in the formulation. A detailed qualitative and quantitative comparison of the sparsity of the resulting components for non-negative matrix factorization, principle component analysis and the independent component analysis on neuroimaging data is given in Sotiras et al14. The non-negative matrix factorization method was applied to the SHIP sample to derive a set of WMH components. Subject-specific coefficients that correspond to these components were then calculated for both the SHIP and the BLSA samples.

The non-negative matrix factorization method estimates a predefined number of components (*k*), thus providing a decomposition of the data in different levels of resolution. The choice of *k* is important and it should be done such that the components capture the data variability while the noise is discarded. Accordingly Sotiras et al. proposed a selection strategy based on the variation of the reconstruction error with *k*, by detecting the inflection point of the slope of the reconstruction error 14.

We applied this strategy to determine *k* for the WMH data by applying the decomposition with *k*=2,3, .. 16 for white matter hyperintensities. We did not observe convergence of the reconstruction error, a measure previously used for determination of optimal number of components in similar data-driven approaches for cortical structures14. This could be explained by the highly skewed distribution of total white matter hyperintensities within the sample, as well as by the lower spatial specificity of white matter hyperintensities compared to cortical structures. For this reason, we applied a strategy adapted to WMH, which aims to find the largest spatial coverage with the lowest number of components. As seen in Figure.e-2, the spatial coverage of the components has shown a peak with *k*=4, and had lower change for larger *k*.

We assessed the robustness of the final components through reproducibility experiments using random half-splits of the SHIP sample with similar age and sex distributions. The components derived from the two independent SHIP splits were largely similar, as shown in Supplementary Figure e-3. We quantified the extent to which these components are overlapping between the two splits by calculating the inner product of matching component pairs from the two splits. Please note that our components are unit norm, so their inner product is equivalent to cosine similarity with a maximum value of 1, for the maximum overlap between them. We present the results of this analysis in Supplementary Table e-1. The median overlap value was 88%, showing highly reproducible results. Finally, We have applied our structural covariance approach in the BLSA baseline sample (n=307). The derived components were highly similar, with a median overlap value of 66% between components derived from SHIP vs. BLSA (Supplementary Table e-2). We believe that these results suggest high reproducibility from a completely independent sample with a different mean age. The derived patterns from BLSA were shown in Supplementary Figure e-4. Notably, the highest overlap score was obtained for the frontal component.

### Supplementary section e-7: Regression models used for association with vascular risk factors, AD genetics and cognition

To study association of WMHC with age we used linear regression models with WMHC as outcome and age decade as dummy variable. Categorization in a dummy variable included the upper edge. The linear regression model is adjusted for the same factors included in the subsequent risk factors analysis (i. e. smoking, education, blood pressure, glycosylated hemoglobin, waist circumference, antihypertensive medication use, anti-diabetic medication use, lipid-lowering medication use, total cholesterol, low-density lipoprotein, high-density lipoprotein and intima-media thickness). We applied linear regression models, which included WMHC as outcomes, and age square, age, sex, smoking, education, blood pressure, glycosylated hemoglobin, waist circumference, antihypertensive medication use, anti-diabetic medication use, lipid-lowering medication use, total cholesterol, low-density lipoprotein, high-density lipoprotein and intima-media thickness4,7 as predictors adjusting for study sub-cohort to assess the cross-sectional association in the whole SHIP sample (n=1836). Similarly, we investigated cross-sectional associations of the four WMHC with the AD polygenic risk score in the whole SHIP sample with available genotyping (n=985) and in those older than 65 years (n=189). In those models WMHC were the outcomes and AD polygenic risk score was the predictor. Regression models for AD genetics using the whole age range SHIP sample were adjusted for age, age2, sex, education and SHIP sub-cohort as explained in our previous analysis on effect of genetic factors on aging brain8. Finally, we used independent linear mixed-effects models to investigate associations of longitudinal change in WMHC with cognition in BLSA (n=307 and 747 observations). In those models cognitive scores were the outcomes and WMHC were predictors and these models were adjusted for age, age2, sex, education level8 and scanner type. All models for risk factors analysis, association with the polygenic risk score in SHIP and longitudinal associations with cognitive testing in BLSA applied Bonferroni multiple comparison correction. As the main outcome variables were the four WMH components, the Bonferroni correction was equivalent to multiplying the actual P value by four in the case of vascular risk factors or AD genetics and by sixteen in the case of cognition. We have used the significance threshold P=0.05 on the corrected P values across the paper.

### Supplementary section e-8: Assessment of the significance of differences in WMHC age trend-lines and trajectories

We assessed the significance of pairwise differences in the age trends of WMH components, through permutation tests15, using 10000 permutations for the cross-sectional analyses; illustrated in Figure.1 for SHIP and Figure.e-6 for BLSA at baseline, and 1000 permutations for the longitudinal analyses. The data matrix *D* included *k=4* measurements (corresponding to the coefficients of the four WMHC) for n subjects (or time points for the longitudinal data). At each permutation experiment, the component assignments, i.e. the indexes of *D* were randomly permuted and age trends were computed for each component using the same regression model used for the actual data. The distance between the two trend lines for each component pair is estimated by calculating the mean absolute distance between the trend lines at *m=1000* uniformly sampled points on the *x* axis, i.e. age. The significance (P values) of the difference between the trend lines of two components is calculated by comparing the distance calculated for actually observed values against those obtained with permuted indexes. Please note that for the longitudinal data the permutation of the index values was done by preserving longitudinal dependencies, such that a component was permuted to the same one for all time points of a subject.

### Supplementary section e-9: A note on potential WMH staging captured with the four white matter hyperintensities components

We performed additional analyses to further explore staging captured with the four WMHC. We followed an approach similar to the approach used in Grothe et al. to investigate the in vivo staging of regional amyloid deposition16. We first categorized each WMHC component for each subject as “present/absent” in terms of WMH (in analogy to “amyloid positive/negative”). We used a constant threshold t for the categorization (t is calculated as the median of total WMH volume for the baseline sample divided by the number of components, i.e. 4 here). WMH components with volumes higher than t were assigned 1 (present), and vice versa. Figure 2 shows the results of this analysis for SHIP and baseline BLSA samples. Each row of the matrix corresponds to a participant and each column to one of the four WMHC. The data shows some hierarchical nesting across participants, both in BLSA and SHIP, starting with WMHC-fron., and suggesting a staging proportion in explaining the data, with the same order of components, except WMHC-dors. and WMHC-deep (note however that these two components have similar counts of “WMH present” subjects). Besides this consistent pattern of staging, we also observe a considerable number of subjects with distribution profiles that don’t conform to the model, which may indicate different etiologies of WMHC.

### Supplementary section e-10: A note on the overlap between the cortical thickness maps related to each WMH component

We have evaluated the overlap between the cortical thickness-maps associated with the four WMH components using the dice coefficient metric. The dice coefficient demonstrates the degree of which two regions are overlapping (Table.e-4). We believe that the proportion of overlap in cortical gray matter atrophy associations is mainly due to the partial correlations between the four WMH components. Please note that, as well as the overlaps, we also observe regional differences in WMH specific cortical atrophy associations. Table.e-4 is showing that cortical regions associated with the frontal WMH component are overlapping to a lesser extent with regions associated with other components (median dice = 0.5211). This is consistent with the results of our analyses; the frontal component appears first and therefore cortical damage related to WMH in the frontal lobe could exist when other WMH components start to appear.

## Supplementary Tables and Figures

Table e-1: Overlap between WMH components calculated independently from the two splits in the SHIP sample. The amount of overlap is measured by the inner product of matching component pairs.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **WMHC-post.**  **(posterior periventricular)** | **WMHC-fron.**  **(frontal periventricular)** | **WMHC-dors.**  **(dorsal periventricular)** | **WMHC-deep**  **(deep white matter)** |
| **Similarity index** | 0.9275 | 0.9759 | 0.7949 | 0.7383 |

Table e-2: Overlap between matching WMH components calculated independently from the SHIP sample and from the BLSA baseline sample. The amount of overlap is measured by the inner product of matching component pairs.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **WMHC-post.**  **(posterior periventricular)** | **WMHC-fron.**  **(frontal periventricular)** | **WMHC-dors.**  **(dorsal periventricular)** | **WMHC-deep**  **(deep white matter)** |
| **Similarity index** | 0.6640 | 0.9164 | 0.6585 | 0.4047 |

Table e-3: Results of the permutation tests for assessing the differences between age trends of WMH component pairs. Bonferroni corrected P-values are reported for each WMHC pair.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | WMHC-post. (posterior periventricular) | WMHC-fron. (frontal periventricular) | WMHC-dors. (dorsal periventricular) | WMHC-deep (deep white matter) |
| SHIP cross-sectional sample (n=1836) | WMHC-post. (posterior periventricular) | X | **0.0004\*** | **0.0016\*** | **0.0004\*** |
| WMHC-fron. (frontal periventricular) | **0.0004\*** | X | **0.0004\*** | **0.0004\*** |
| WMHC-dors. (dorsal periventricular) | **0.0016\*** | **0.0004\*** | X | 0.8000 |
| WMHC-deep (deep white matter) | **0.0004\*** | **0.0004\*** | 0.8000 | X |
|  |  |  |  |  |  |
| BLSA cross-sectional sample at baseline (n=307) | WMHC-post. (posterior periventricular) | X | **0.0004\*** | **0.00016\*** | **0.0004\*** |
| WMHC-fron. (frontal periventricular) | **0.0004\*** | X | **0.0004\*** | **0.0004\*** |
| WMHC-dors. (dorsal periventricular) | **0.0006\*** | **0.0004\*** | X | 0.2370 |
| WMHC-deep (deep white matter) | **0.0004\*** | **0.0004\*** | 0.2370 | X |
|  |  |  |  |  |  |
| BLSA longitudinal sample (n=307 and 747 observations) | WMHC-post. (posterior periventricular) | X | **0.004\*** | **0.004\*** | **0.004\*** |
| WMHC-fron. (frontal periventricular) | **0.004\*** | X | **0.004\*** | **0.004\*** |
| WMHC-dors. (dorsal periventricular) | **0.004\*** | **0.004\*** | X | **0.008\*** |
| WMHC-deep (deep white matter) | **0.004\*** | **0.004\*** | **0.008\*** | X |

\*Significant at level <0.05

Table e-4: Overlap between GM cortical atrophy regions associated with each of the four WMH components. The amount of overlap between two regions is measured using their Dice coefficient.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **WMHC-post.**  **(posterior periventricular)** | **WMHC-fron.**  **(frontal periventricular)** | **WMHC-dors.**  **(dorsal periventricular)** | **WMHC-deep**  **(deep white matter)** |
| **WMHC-post.**  **(posterior periventricular)** | X | 0.4958 | 0.6815 | 0.7889 |
| **WMHC-fron.**  **(frontal periventricular** | 0.4958 | X | 0.6530 | 0.5211 |
| **WMHC-dors.**  **(dorsal periventricular)** | 0.6815 | 0.6530 | X | 0.7099 |
| **WMHC-deep**  **(deep white matter)** | 0.7889 | 0.5211 | 0.7099 | X |

Table e-5: Multiple regression models for WMHC in a cross-sectional association with risk factors in the SHIP sample included in this study after exclusion of possible impaired individuals (n=1,658). A linear regression model was used independently for each WMHC, using all risk factors jointly as predictors and the WMHC values as the outcome.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Factor |  |  |  |  |  |  |  |  |
|  | WMHC-post.  (posterior periventricular) | | WMHC-fron.  (frontal periventricular) | | WMHC-dors.  (dorsal periventricular) | | WMHC-deep  (deep white matter) | |
|  | Estimate (S.E.) | P-Value (Bonferroni corrected) | Estimate (S.E.) | P-Value (Bonferroni corrected) | Estimate (S.E.) | P-Value (Bonferroni corrected) | Estimate (S.E.) | P-Value (Bonferroni corrected) |
|  |  |  |  |  |  |  |  |  |
| **Sex, Female** | -0.104(0.073) | 0.628 | 0.140(0.064) | 0.116 | -0.035(0.057) | 1.000 | 0.050(0.051) | 1.000 |
|  |  |  |  |  |  |  |  |  |
| **Systolic blood pressure, mm Hg** | 0.001(0.002) | 1.000 | 0.005(0.002) | **0.016\*** | 0.002(0.002) | 1.000 | 0.002(0.001) | 0.564 |
|  |  |  |  |  |  |  |  |  |
| **Intima-Media Thickness §, mm** | 0.200(0.155) | 0.784 | 0.111(0.135) | 1.000 | 0.273(0.121) | 0.096 | 0.129(0.107) | 0.912 |
|  |  |  |  |  |  |  |  |  |
| **Cigarette smoking** |  |  |  |  |  |  |  |  |
| Ex-smoker | 0.205(0.068) | **0.012\*** | 0.047(0.059) | 1.000 | 0.125(0.053) | 0.076 | 0.106(0.047) | 0.096 |
| Current smoker | 0.205(0.080) | **0.040\*** | 0.133(0.069) | 0.224 | 0.102(0.062) | 0.412 | 0.119(0.055) | 0.128 |
|  |  |  |  |  |  |  |  |  |
| **Antihypertensive drugs** | 0.119(0.076) | 0.468 | 0.149(0.066) | 0.096 | 0.076(0.059) | 1.000 | 0.077(0.052) | 0.560 |
|  |  |  |  |  |  |  |  |  |
| **Lipid lowering drugs** | 0.146(0.111) | 0.756 | 0.046(0.097) | 1.000 | 0.092(0.087) | 1.000 | 0.139(0.077) | 0.069 |
|  |  |  |  |  |  |  |  |  |

\* Significance at level p < 0.05, S.E: Standard Error, White matter hyperintensities Component (WMHC), § Measure was available for 1,644 subjects. Models are adjusted for age2, age, cholesterol ratio (i.e. high-density lipoprotein / low-density lipoprotein), glycosylated hemoglobin (HbA1c), waist circumference, education, physical activity, antidiabetic drugs and SHIP sub cohorts

Table e-6: Longitudinal associations between WMHC and cognitive scores in the BLSA cohort. A linear regression model was used independently for each WMHC (as predictor) and for each cognitive test (as the outcome).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cohort | Outcome |  |  |  |  |  |  |  |  |
|  |  | WMHC-post.  (posterior periventricular) | | WMHC-fron.  (frontal periventricular) | | WMHC-dors.  (dorsal periventricular) | | WMHC-deep  (deep white matter) | |
|  |  | Estimate (S.E.) | P-Value (Bonferroni corrected) | Estimate (S.E.) | P-Value (Bonferroni corrected) | Estimate (S.E.) | P-Value (Bonferroni corrected) | Estimate (S.E.) | P-Value (Bonferroni corrected) |
| **BLSAa(n=307, 747 observations)** |  |  |  |  |  |  |  |  |  |
| **TMT-A** | 0.132(0.288) | 1.000 | 0.516(0.372) | 1.000 | 0.475(0.342) | 1.000 | 0.252(0.498) | 1.000 |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| **TMT-B** | 2.294(0.911) | 0.192 | 2.549(1.271) | 0.720 | 3.332(1.021) | **0.016\*** | 3.723(1.379) | 0.112 |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| **BVRT** | 0.165(0.103) | 1.000 | 0.284(0.139) | 0.672 | 0.317(0.116) | 0.112 | 0.426(0.159) | 0.144 |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| **CVLT&** | 0.192(0.150) | 1.000 | 0.414(0.212) | 0.832 | 0.287(0.176) | 1.000 | 0.463(0.239) | 0.848 |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

\* Significance at level P<0.05

& Sum of short- and long-delay free recall

a Linear mixed-effects models for longitudinal relationships are adjusted for, age, age2, sex, education and BLSA scanner number

S.E: Standard Error, Baltimore Longitudinal Study of Aging (BLSA), White Matter Hyperintensities Component (WMHC), Trail Making Test (TMT), Benton Visual Retention Test (BVRT), California Verbal Learning Test (CVLT)

Table e-7: Longitudinal associations between WMHC and cognitive scores in cognitively normal subjects in BLSA. A linear regression model was used independently for each WMHC (as predictor) and for each cognitive test (as the outcome).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cohort | Outcome |  |  |  |  |  |  |  |  |
|  |  | WMHC-post.  (posterior periventricular) | | WMHC-fron.  (frontal periventricular) | | WMHC-dors.  (dorsal periventricular) | | WMHC-deep  (deep white matter) | |
|  |  | Estimate (S.E.) | P-Value (Bonferroni corrected) | Estimate (S.E.) | P-Value (Bonferroni corrected) | Estimate (S.E.) | P-Value (Bonferroni corrected) | Estimate (S.E.) | P-Value (Bonferroni corrected) |
|  |  |  |  |  |  |  |  |  |  |
| **TMT-A** | 0.151 (0.289) | 1.000 | 0.547 (0.374) | 1.000 | 0.486 (0.347) | 0.648 | 0.290 (0.504) | 1.000 |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| **TMT-B** | 2.247 (0.919) | 0.240 | 2.558 (1.279) | 0.736 | 3.330 (1.027) | **0.016\*** | 3.600 (1.471) | 0.240 |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| **BVRT** | 0.149 (0.103) | 1.000 | 0.258 (0.139) | 1.000 | 0.278 (0.117) | 0.288 | 0.394 (0.159) | 0.224 |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| **CVLT&** | -0.123 (0.147) | 1.000 | -0.299 (0.192) | 1.000 | -0.164 (0.175) | 1.000 | -0.353 (0.236) | 1.000 |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

\* Significance at level p < 0.05.

& Sum of short- and long-delay free recall

a Linear mixed-effects models for longitudinal relationships were adjusted for, age, age2, sex, education and BLSA scanner number.

S.E: Standard Error, Baltimore Longitudinal Study of Aging (BLSA), White Matter Hyperintensities Component (WMHC), Trail Making Test (TMT), Benton Visual Retention Test

Table e-8: Pairwise Pearson’s correlation coefficients between WMH components and and between each component and the total WMH volume in SHIP.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **WMHC-post.**  **(posterior periventricular)** | **WMHC-fron.**  **(frontal periventricular)** | **WMHC-dors.**  **(dorsal periventricular)** | **WMHC-deep**  **(deep white matter)** | **Total WMHC volume** |
| **WMHC-post.**  **(posterior periventricular)** | X | 0.628\* | 0.874\* | 0.844\* | 0.965\* |
| **WMHC-fron.**  **(frontal periventricular** | 0.628\* | X | 0.581\* | 0.581\* | 0.759\* |
| **WMHC-dors.**  **(dorsal periventricular)** | 0.874\* | 0.581\* | X | 0.782\* | 0.924\* |
| **WMHC-deep**  **(deep white matter)** | 0.844\* | 0.581\* | 0.782\* | X | 0.886\* |
| **Total WMHC volume** | 0.965\* | 0.759\* | 0.924\* | 0.886\* | X |

\* Significant at p<0.001

Abbreviations: WM hyperintensities (WMH)

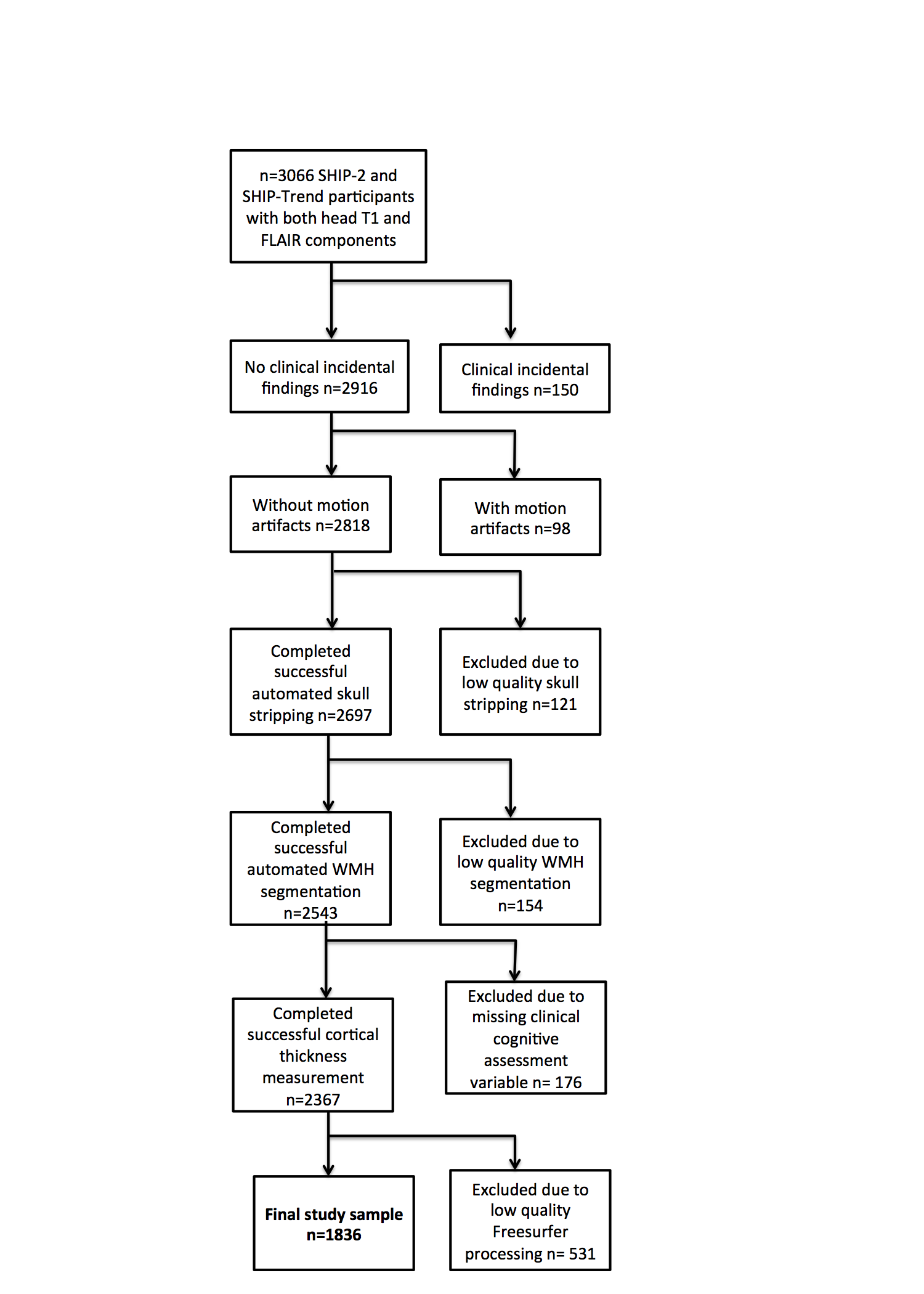


Figure e-1: Flowchart showing final SHIP sample included in this study.

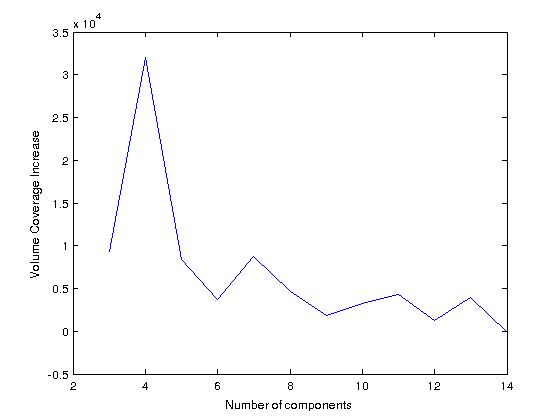


Figure e-2: Change in total white matter hyperintensities volume coverage as a function of number of white matter hyperintensities components. The figure demonstrates that with more components larger volume is covered. However the increase is less important after four components, which was determined as the optimal number for maximal spatial coverage with the lowest number of components. Notably, adding more components could lead to more spatial overlap between them.

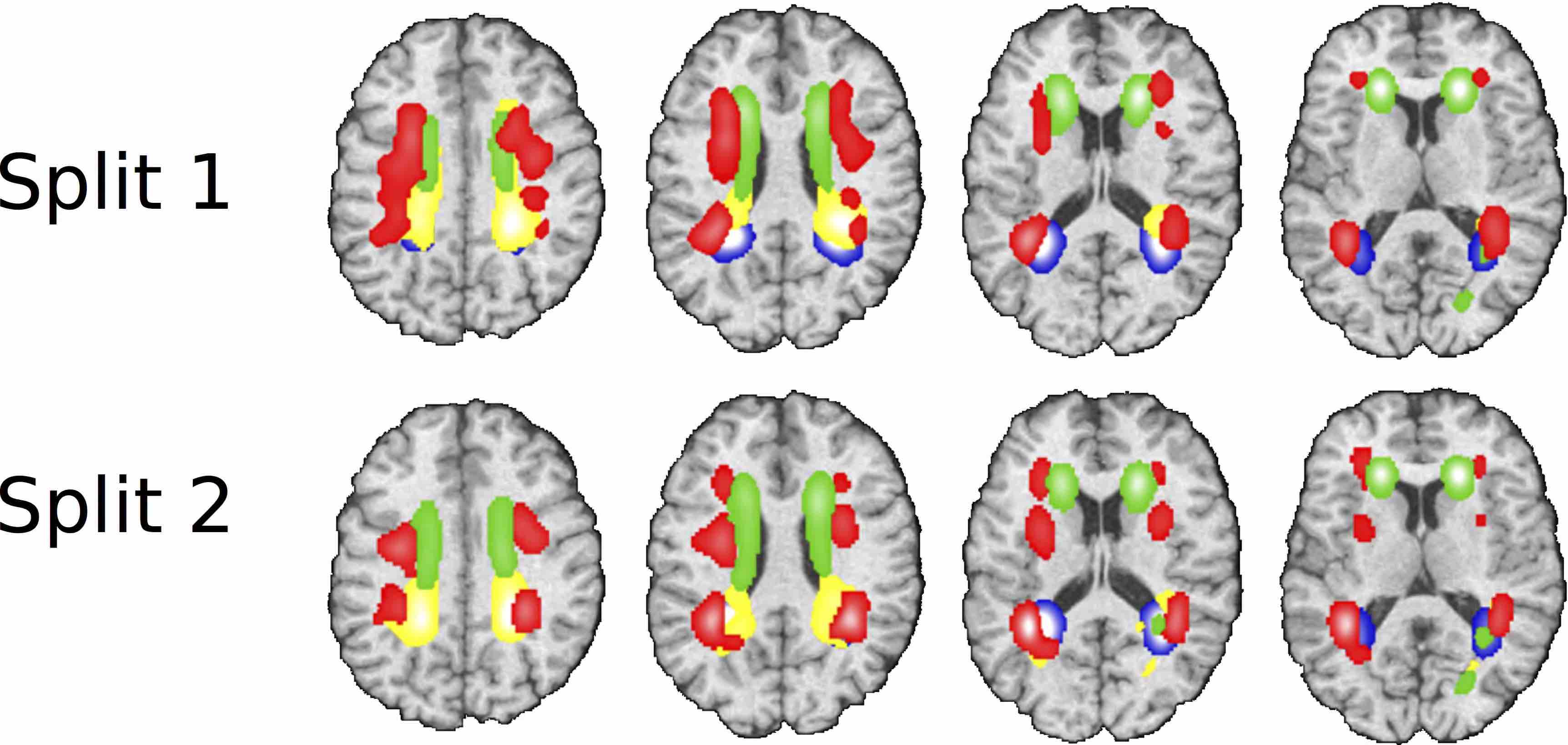


Figure e-3: The decomposition of white matter hyperintensities into four components showed highly reproducible results in two random splits (n=918 for both splits) of the SHIP population with similar age and sex distribution. The four WMH components are posterior periventricular (WMHC-post., blue), frontal periventricular (WMHC-fron., green), dorsal periventricular (WMHC-dors., yellow) and deep WM (WMHC-deep, red).

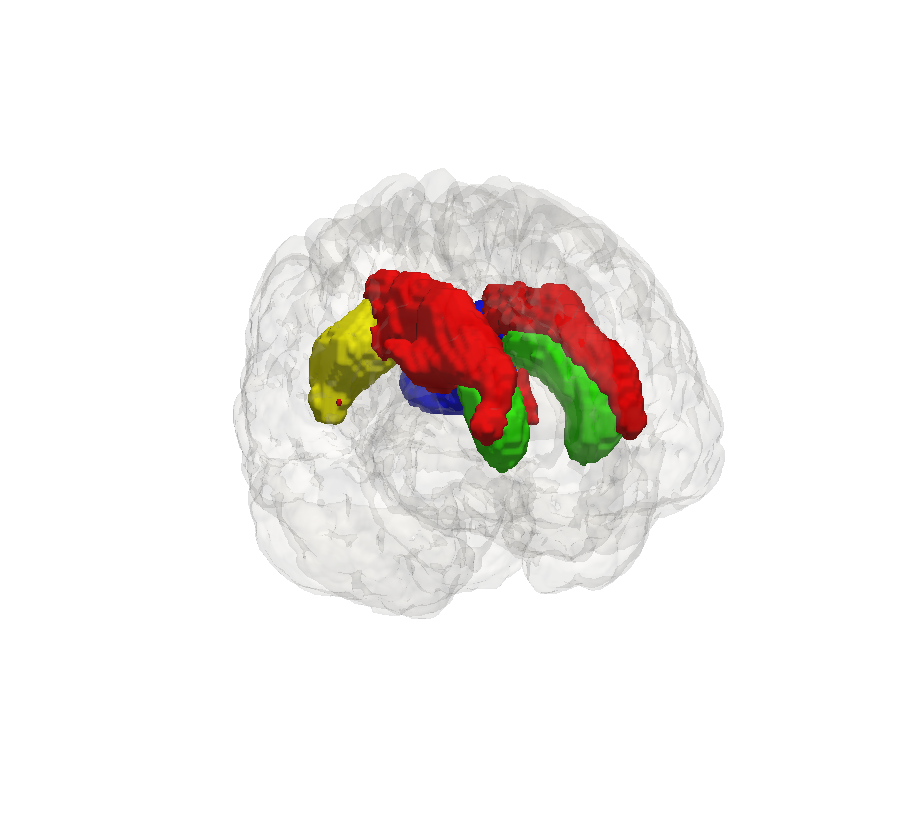


Figure e-4: The decomposition of white matter hyperintensities into four components in the BLSA population (n=307) showed reproducible results to some extent compared to the decomposition in the SHIP sample. Please note that age, sex and education differ between the two samples but we still get median overlap close to 66% in the reconstruction of the four components. The four WMH components are posterior periventricular (WMHC-post., blue), frontal periventricular (WMHC-fron., green), dorsal periventricular (WMHC-dors., yellow) and deep WM (WMHC-deep, red)

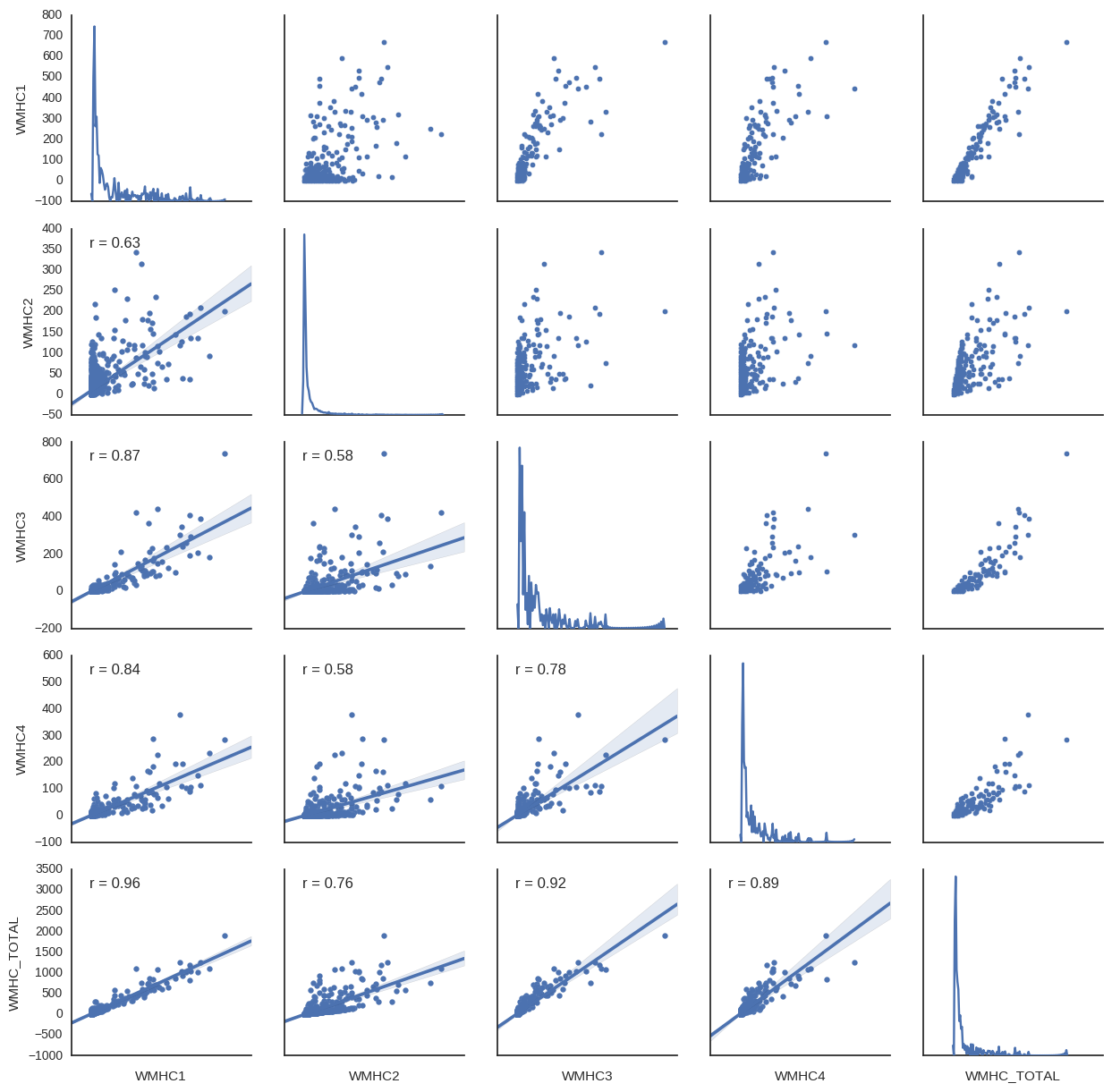


Figure e-5 Pairwise correlations between components, and between each component and the total WMH in SHIP. While the component pairs have high correlations, this can partly be explained by the skewed distribution of the data, i.e. with a small proportion of the subjects with significantly high WMH values, and we observe that the relationship is not linear between each pair of components.

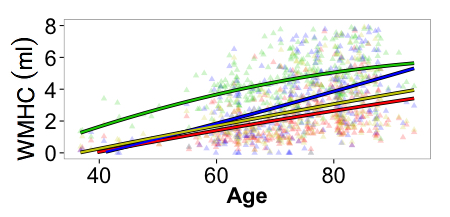


Figure e-6: White matter hyperintensities components (WMHC) coefficients plotted as function of age for the BLSA sample at baseline (n=307). The four WMH components are periventricular posterior (WMHC-post., blue), periventricular frontal (WMHC-fron., green), periventricular dorsal (WMHC-dors., yellow) and deep WM (WMHC-deep, red).

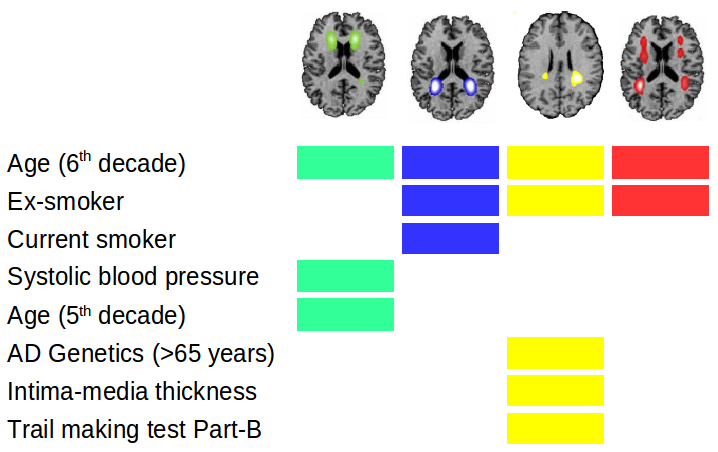


Figure e-7 Summary associations of vascular risk factors, AD polygenic risk score and cognitive tests with white matter hyperintensities components (Bonferroni corrected P <0.05) based upon regression models explained in Tables 3, 4 and e-6.

Abbreviations: White Matter Hyperintensities Component (WMHC), Trail Making Test (TMT)

## References

1. Venkatraman VK, Gonzalez CE, Landman B, et al. Region of interest correction factors improve reliability of diffusion imaging measures within and across scanners and field strengths. *NeuroImage*. 2015;119:406-416. doi:10.1016/j.neuroimage.2015.06.078

2. Völzke H, Alte D, Schmidt CO, et al. Cohort Profile: The Study of Health in Pomerania. *Int J Epidemiol*. 2010. doi:10.1093/ije/dyp394

3. Lambert J-C, Ibrahim-Verbaas CA, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer’s disease. *Nat Genet*. 2013;45(12):1452-1458.

4. Habes M, Janowitz D, Erus G, et al. Advanced Brain Aging: relationship with epidemiologic and genetic risk factors, and overlap with Alzheimer disease atrophy patterns. *Transl Psychiatry*. 2016;6:e775.

5. Woods SP, Delis DC, Scott JC, Kramer JH, Holdnack JA. The California Verbal Learning Test – second edition: Test-retest reliability, practice effects, and reliable change indices for the standard and alternate forms. *Arch Clin Neuropsychol*. 2006;21(5):413-420. doi:http://dx.doi.org/10.1016/j.acn.2006.06.002

6. Fleischmann UM, Oswald WD. *Nürnberger-Alters-Inventar: NAI:[Testkarten Mit Testvorlagen, 2 Sätze Testbogen (Version A-E), Auswertungsfolien Und Kurzmanual].* Univ. Erlangen-Nürnberg, Inst. für Psychologie II; 1990.

7. Habes M, Erus G, Toledo JB, et al. White matter hyperintensities and imaging patterns of brain aging in the general population. *Brain*. 2016;139(Pt 4):1164-1179. doi:10.1093/brain/aww008

8. Habes M, Toledo JB, Resnick SM, et al. Relationship between APOE Genotype and Structural MRI Measures throughout Adulthood in the Study of Health in Pomerania Population-Based Cohort. *AJNR Am J Neuroradiol*. May 2016. doi:10.3174/ajnr.A4805

9. Lao Z, Shen D, Liu D, et al. Computer-Assisted Segmentation of White Matter Lesions in 3D {MR} Images Using Support Vector Machine. *Acad Radiol*. 2008;15(3):300-313. doi:http://dx.doi.org/10.1016/j.acra.2007.10.012

10. Fischl B, Salat DH, Busa E, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*. 2002;33(3):341-355.

11. Fischl B, Dale AM. Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc Natl Acad Sci U S A*. 2000;97(20):11050-11055. doi:10.1073/pnas.200033797

12. Fischl B, Sereno MI, Tootell RB, Dale AM. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum Brain Mapp*. 1999;8(4):272-284.

13. Desikan RS, Segonne F, Fischl B, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*. 2006;31(3):968-980. doi:10.1016/j.neuroimage.2006.01.021

14. Sotiras A, Resnick SM, Davatzikos C. Finding imaging patterns of structural covariance via Non-Negative Matrix Factorization. *NeuroImage*. 2015;108:1-16. doi:10.1016/j.neuroimage.2014.11.045

15. Good PI. *Permutation, Parametric, and Bootstrap Tests of Hypotheses (Springer Series in Statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc.; 2004.

16. Grothe MJ, Barthel H, Sepulcre J, Dyrba M, Sabri O, Teipel SJ. In vivo staging of regional amyloid deposition. *Neurology*. 2017;89(20):2031-2038. doi:10.1212/WNL.0000000000004643