White Matter Lesion Progression

Genome-Wide Search for Genetic Influences

Edith Hofer, PhD; Margherita Cavaliere, MD; Joshua C. Bis, PhD; Charles DeCarli, MD; Myriam Fornage, PhD; Sigurdur Sigurdsson, MSc; Velandai Srikanth, PhD; Stella Trompet, PhD; Benjamin F.J. Verhaaren, MD, PhD; Christiane Wolf, PhD; Qiong Yang, PhD; Hieab H.H. Adams, MSc; Philippe Amouyel, MD, PhD; Alexa Beiser, PhD; Brendan M. Buckley, MD, PhD; Michele Callisaya, PhD; Ganesh Chauhan, PhD; Anton J.M. de Craen, PhD; Carole Dufouil, PhD; Cornelia M. van Duijn, PhD; Ian Ford, PhD; Paul Freudenberg, MSc; Rebecca F. Gottesman, MD, PhD; Vilmundur Gudnason, MD, PhD; Gerardo Heiss, MD, PhD; Albert Hofman, MD, PhD; Thomas Lumley, PhD; Oliver Martinez, PhD; Bernard Mazoyer, MD, PhD; Chris Moran, MB BCH; Wiro J. Niessen, PhD; Thanh Phan, PhD; Bruce M. Psaty, MD, MPH, PhD; Claudia L. Satizabal, PhD; Naveed Sattar, MD, PhD; Sabrina Schilling, MSc; Dean K. Shibata, MD; P. Eline Slagboom, PhD; Albert Smith, PhD; David J. Stott, MD, PhD; Kent D. Taylor, PhD; Russell Thomson, PhD; Anna M. Töglhofer, MSc; Christophe Tzourio, MD, PhD; Mark van Buchem, MD, PhD; Jing Wang, PhD; Rudi G.J. Westendorp, MD, PhD; B. Gwen Windham, MD; Meike W. Vernooij, MD, PhD; Alex Zijdenbos, PhD; Richard Beare, PhD; Stéphanie Debette, MD, PhD; M. Arfan Ikram, MD, PhD; J. Wouter Jukema, MD, PhD; Lenore J. Launer, PhD; W. T. Longstreth Jr, MD; Thomas H. Mosley, PhD; Sudha Seshadri, MD; Helena Schmidt, MD, PhD; Reinhold Schmidt, MD; on behalf of the Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium

Received March 2, 2015; final revision received July 15, 2015; accepted August 21, 2015.

From the Department of Neurology (E.H., M.C., H.S., R.S.) and Institute for Medical Informatics, Statistics and Documentation (E.H.), Medical University of Graz, Graz, Austria; Cardiovascular Health Research Unit, Departments of Medicine (J.C.B., B.M.P.), Radiology (D.K.S.), and Neurology and Epidemiology (W.T.L.), University of Washington, Seattle; Department of Neurology and Center for Neuroscience, University of California at Davis, Sacramento (C.D.C., O.M.); Brown Foundation Institute of Molecular Medicine (M.F.) and Human Genetics Center (M.F.), University of Texas Health Science Center, Houston; Icelandic Heart Association, Kopavogur, Iceland (S.S., V.G., A.S.); Stroke and Ageing Research Group, Department of Medicine, Southern Clinical School, Monash University, Melbourne, Victoria, Australia (V.S., M.C., C.M., T.P., R.B.); Departments of Cardiology (S.T., J.W.J.), Gerontology and Geriatrics (S.T., A.J.M.C., R.G.J.W.), Molecular Epidemiology (P.E.S.), and Radiology (M.B.), Leiden University Medical Center, Leiden, The Netherlands; Departments of Epidemiology (B.F.J.V., H.H.H.A., C.M.D., A.H., M.W.V., M.A.I.), Radiology (B.F.J.V., H.H.H.A., M.W.V., M.A.I.), and Medical Informatics (W.I.N.), Erasmus Medical Center, Rotterdam, The Netherlands; INSERM U897 (C.W., G.C., D.D., S.S., D.R.), CNRS-CEA UMR5296 (B.M.), and INSERM U708 (C.T.), Université Bordeaux Segalen, Bordeaux, France; Department of Statistical Genetics, Max Planck Institute of Psychiatry, Munich, Germany (C.W.); Department of Biostatistics, Boston University School of Public Health, MA (Q.Y., A.B., J.W.); INSERM U744, Pasteur Institute, Lille, France (P.A.); Department of Pharmacology and Therapeutics, University College Cork, Cork, Ireland (B.M.B.); INSERM U1161 and Lariboisière Hospital, Paris 7 University, Paris, France (G.C., S.S., S.D.); Centre for Medical Systems Biology, Leiden, The Netherlands (C.M.D.); Robertson Center for Biostatistics, University of Glasgow, Glasgow, United Kingdom (I.F.); Research Unit for Genetic Epidemiology, Institute of Molecular Biology and Biochemistry, Center of Molecular Medicine, Medical University of Graz, Graz, Austria (P.F., A.M.T., H.S.); Department of Neurology, Johns Hopkins University, Baltimore, MD (R.F.G.); University of Iceland, Reykjavik, Iceland (V.G., A.S.); Department of Epidemiology, University of North Carolina, Chapel Hill (G.H.); Department of Statistics, University of Auckland, Auckland, New Zealand (T.L.); Faculty of Applied Sciences, Delft University, Delft, The Netherlands (W.I.N.); Group Health Research Institute, Group Health Cooperative, Seattle, WA (B.M.P.); Department of Neurology, Boston University School of Medicine, MA (C.L.S., S.D., S.S.); Institute of Cardiovascular and Medical Sciences, Faculty of Medicine, Glasgow, United Kingdom (N.S., D.J.S.); Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute, CA (K.D.T.); Department of Pediatrics, Harbor-UCLA Medical Center, Torrance, CA (K.D.T.); Menzies Research Institute, Hobart, Tasmania, Australia (R.T.); Department of Public Health, University of Copenhagen, Copenhagen, Denmark (R.G.I.W.); Department of Medicine, University of Mississippi, Jackson (B.G.W., T.H.M.); Biospective Inc, Montréal, Quebec, Canada (A.Z.); Developmental Imaging, Murdoch Childrens Research Institute, Royal Childrens Hospital, Parkville, Victoria, Australia (R.B.); Durre Center for Cardiogenetic Research, Amsterdam, The Netherlands (J.W.J.); Interuniversity Cardiology Institute of the Netherlands, Utrecht, The Netherlands (J.W.J.); and National Institute on Aging, Bethesda, MD (L.J.L.).

The online-only Data Supplement is available with this article at http://stroke.ahajournals.orglookup/suppl/doi:10.1161/STROKEAHA.115.009252.

Correspondence to Reinhold Schmidt, MD, Division of Neurogeriatrics, Department of Neurology, Medical University of Graz, Auenbruggerplatz 22, 8036 Graz, Austria. E-mail reinhold.schmidt@medunigraz.at

© 2015 American Heart Association, Inc.

Stroke is available at http://stroke.ahajournals.org DOI: 10.1161/STROKEAHA.115.009252

http://stroke.ahajournals.org/lookup/suppl/doi:10.1161/STROKEAHA.115.009252
Background and Purpose—White matter lesion (WML) progression on magnetic resonance imaging is related to cognitive decline and stroke, but its determinants besides baseline WML burden are largely unknown. Here, we estimated heritability of WML progression, and sought common genetic variants associated with WML progression in elderly participants from the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium.

Methods—Heritability of WML progression was calculated in the Framingham Heart Study. The genome-wide association study included 7773 elderly participants from 10 cohorts. To assess the relative contribution of genetic factors to progression of WML, we compared in 7 cohorts risk models including demographics, vascular risk factors plus single-nucleotide polymorphisms that have been shown to be associated cross-sectionally with WML in the current and previous association studies.

Results—A total of 1085 subjects showed WML progression. The heritability estimate for WML progression was low at 6.5%, and no single-nucleotide polymorphisms achieved genome-wide significance (P<5×10⁻⁸). Four loci were suggestive (P<1×10⁻⁵) of an association with WML progression: 10q24.32 (rs10883817, P=1.46×10⁻⁷); 12q13.13 (rs4761974, P=8.71×10⁻⁷); 20p12.1 (rs6135309, P=3.69×10⁻⁶); and 4p15.31 (rs7664442, P=2.26×10⁻⁶). Variants that have been previously related to WML explained only 0.8% to 11.7% more of the variance in WML progression than age, vascular risk factors, and baseline WML burden.

Conclusions—Common genetic factors contribute little to the progression of age-related WML in middle-aged and older adults. Future research on determinants of WML progression should focus more on environmental, lifestyle, or host-related biological factors. (Stroke. 2015;46:3048-3057. DOI: 10.1161/STROKEAHA.111.009252.)

Key Words: aging ▪ biological factors ▪ cerebral small vessel diseases ▪ magnetic resonance imaging ▪ white matter lesions

The pathogenesis of white matter lesions (WMLs) on magnetic resonance imaging (MRI) is still incompletely understood. WML burden was shown to be highly heritable.1–3 Age and hypertension are the main known risk factors for WML but explain only a small proportion of lesion presence and burden.4 Before the era of genome-wide association studies (GWAS), candidate gene studies investigated variants in 19 genes and found associations between WML extent and polymorphisms in the apolipoprotein E (APOE), the methylentetrahydrofolate reductase, the angiotensin-converting enzyme, and the angiotensinogen genes.5,6 Moreover, genetic variants in the NOTCH3 gene may not only play a role in CADASIL (cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy), a monogenic cerebral small vessel disease, but are also likely to be involved in the pathogenesis of age-related WML burden.7 In 2011, the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium8 performed the first GWAS on WML burden in the general population.9 The CHARGE investigators identified 6 single-nucleotide polymorphisms (SNPs) mapping to a locus on chromosome 17q25 to be related to WML burden. The findings of the discovery meta-analyses had been confirmed in an independent sample of 1607 Aging Gene-Environment Susceptibility Reykjavik Study (AGES-Reykjavik) participants and in 1417 and 1677 elderly white participants from the Three-City-Dijon Study (3C-Dijon)9 and the Rotterdam Study III (RS III),10 as well as in a study of WML burden in persons with a clinical ischemic stroke.11 The association was also confirmed in an Asian population including 1190 Japanese persons with a mean age of 66 years.12 The region on chromosome 17 is ~100-kb long and harbors several genes with diverse functions such as the 2 tripartite motif-containing genes (TRIM65 and TRIM47) the WW domain-binding protein 2 gene (WB2P), the mitochondrial ribosomal protein L38 gene (MRPL38), the Fas-binding factor 1 gene (FBF1), the acyl-coenzyme A oxidase 1 gene (ACOX1), and the Caenorhabditis elegans homolog (UNC13D) gene. Although genetic factors may play an important role in the occurrence of WML in middle-aged to older adults, whether these genes or others influence the further progression of WML is unknown. Using data on WML progression from all the cohorts currently available within the CHARGE consortium, we examined the heritability of WML progression and performed a meta-analysis of GWAS data in 7773 individuals of European descent from 10 cohorts to identify common SNPs that influence the risk for WML progression. We also assessed the relative contribution of genetic factors in predicting WML progression beyond information that can be obtained from baseline clinical and MRI data alone.

Materials and Methods

Study Population

Study participants were from 10 prospective cohort studies collaborating in the CHARGE Consortium:13 the Atherosclerosis Risk in Communities (ARIC) study,14 the Austrian Stroke Prevention Study (ASPS),15 the Cardiovascular Health Study (CHS),17 the Framingham Heart Study (FHS),18,19 the Prospective Study of Pravastatin in the Elderly at Risk (PROSPER),20,21 the RS II, the RS III,22 the Tasmanian Study of Cognition and Gait (TASCOG),23 and the 3C-Dijon study.24 All participating studies agreed on phenotype harmonization, covariate selection, prespecified analytic plans for within-study analyses, and meta-analysis of results. Each study secured approval from Institutional Review Boards, and all participants provided written informed consent for study participation, MRI scanning, and use of DNA for genetic research.

Participants were eligible for this study if they had genotyping, serial MRI, and lacked a history of transient ischemic attacks, strokes, dementia, or any combination of these conditions. All the individuals in this analysis were whites of European descent. The number of participants and their characteristics in each cohort are shown in Table I in the online-only Data Supplement.
WML Progression Assessment

In each study, eligible participants were invited to undergo serial MRI scans, which were performed and interpreted in a standardized fashion without knowledge of demographic, clinical, or genetic information (Section II in the online-only Data Supplement). Except for ARIC and CHS whose readers used a 10-point scale, readers in the other cohorts measured the volume of WML on each MRI scan. WML progression was defined as absent or present (Section III in the online-only Data Supplement). In brief, WML progression was considered to be present if visual rating increased by at least 1 grade between baseline and follow-up in ARIC and CHS, or if WML volume increased by at least 1 SD of the study-specific mean of volume change in AGES-Reykjavik, ASPS, FHS, PROSPER, RS II, RS III, TASCOG, and 3C-Dijon. WML regression was rare in all studies and was not considered separately in these analyses.

Genotyping

The consortium was formed after individual studies had finalized their GWAS platforms, which differed across studies. All studies used their genotype data to impute to the 2.5 million nonmonomorphic, autosomal SNPs described in HapMap European population panel. Extensive quality control analyses were performed in each cohort. Details on the genotyping, quality control, and imputation efforts are described in Tables II–IV in the online-only Data Supplement.

Heritability of WML Progression

Heritability of WML progression was calculated based on family structure in FHS, which is the only family study among the participating cohorts. Calculations were based on annual WML change and 2 models were assessed. The first model adjusted for age and sex, whereas the second model additionally adjusted for WML volume at baseline. The ratio of the genetic variance to the phenotypic variance in FHS was determined using variance component models in the Sequential Oligogenic Linkage Analysis Routines (SOLAR) software.25

Genome-Wide Association Analyses

For the GWAS, each study fitted an additive genetic model with a 1 degree-of-freedom trend test relating genotype dosage, 0 to 2 copies of the minor allele, and presence or absence of WML progression. We used logistic regression models to calculate regression estimates with corresponding SEs. Initial analyses were adjusted for age, sex, interval between scans as well as principal components of population structure if appropriate. Subsequent analyses included additional adjustment for WML volume at baseline. In addition, ARIC and CHS also adjusted for study site, and FHS, for familial structure. We then conducted a meta-analysis of logistic regression estimates and SEs using a fixed effects inverse-variance weighting approach with genomic control to correct for population stratification if appropriate. We also conducted a meta-analysis of genome-wide association studies without adjustment for WML volume at baseline. The meta-analysis regression coefficients $\beta$ of the SNPs were then compared between the investigational subsets using a z test.

Annotation

Key SNPs were functionally annotated with SNPnexus27 and ANNOVAR.28 The SNAP29 web application of the Broad Institute was used to determine linkage disequilibrium (LD). Regional association plots were generated with LocusZoom.30 Furthermore, we used the NCBI Genotype-Tissue Expression eQTL Browser31 to check if a given SNP was associated with a quantitative gene expression trait.

Performance of WML Progression Risk Models

To estimate the relative importance of genetic factors for progression of WML, we compared the explained variance between 3 models in ASPS, CHS, FHS, PROSPER, RS II, RS III, and 3C-Dijon. Model 1 served as the reference and included age, sex, hypertension, diabetes mellitus, current smoking, and the time interval between scans; model 2 included model 1 variables and those SNPs with the lowest $P$ value at each highly suggestive locus according to the present GWAS meta-analysis with adjustment for WML burden at baseline. We also included a missense SNP identified by annotation analyses. Model 3 included all model 2 variables plus genetic polymorphisms that were shown to be associated with WML in previous publications.3,7,9,32–43

Results

The analyses included 7773 participants from 10 cohort studies (Table I in the online-only Data Supplement). A total of 4342 (56%) study participants were women, 4030 (52%) were hypertensive, and 690 (8.9%) were diabetic. The mean systolic blood pressure ranged between 121.9 and 157.8 mm Hg, and the mean diastolic blood pressure ranged between 69.6 and 86.5 mm Hg. The mean time interval between the baseline and follow-up MRI was shortest in AGES-Reykjavik with 29.1 months and longest in ARIC with 123.6 months. Overall, WML progression was observed in 1085 (14.0%) study participants with the study-specific mean annual volume increase ranging between 0.2 cm$^3$ and 1.4 cm$^3$. The association between mean WML volume at baseline and mean annual volume increase in our cohorts was moderate (Pearson correlation coefficient $r=0.59$), but did not reach statistical significance ($P=0.095$). The relationship for each cohort is displayed in Figure III in the online-only Data Supplement.

Heritability of WML Progression

After adjustment for age and sex, the heritability estimate for WML progression observed in 1368 FHS individuals was 6.5%. Additional adjustment for WML burden at baseline resulted in a heritability estimate of 4.7%.

Genome-Wide Associations of WML Progression

Figure 1 illustrates the meta-analysis results of genome-wide association analyses on WML progression without and with adjustment for WML burden at baseline. In both the meta-analyses, no SNPs achieved genome-wide significance ($P<5\times10^{-8}$). Meta-analysis without adjustment for WML burden at baseline revealed 45 SNPs in 7 loci on 7 chromosomes with a highly suggestive association with WML progression.
at $P<1\times10^{-5}$ (Table V in the online-only Data Supplement). After additional adjustment for WML burden at baseline, 8 highly suggestive associations were identified (Table VI in the online-only Data Supplement). These SNPs are located in 6 loci on 6 chromosomes. There is an overlap of 4 suggestive loci (1q41, 5q12.1, 12q13.13, and 13q34) between the 2 meta-analysis results. In the meta-analysis without adjustment for WML burden at baseline 35 of 45 suggestive SNPs in total were located at 10q24.32. The SNP rs10883817 had the lowest $P$ value ($1.46\times10^{-6}$) at this locus, with an odds ratio (OR) of 1.27, a mean minor allele frequency of 0.41, and a mean imputation quality score of 0.98. The suggestive variants at locus 10q24.32 were in LD ($r^2$ between 0.336 and 1) and reside either in introns of genes AS3MT, CNNM2, NT5C2, in the intergenic regions between these genes, or in exons of CNNM2 and NT5C2 (Figure 2A).

At the second locus of interest, 12q13.13, we identified 3 highly suggestive variants in the analysis without adjustment for WML burden and only 1 with adjustment. In the meta-analysis without adjustment for baseline WML burden, rs4761974 had the lowest $P$ value ($8.71\times10^{-7}$) with an OR of 0.5. After additional adjustment for baseline WML volume the $P$ value of the association changed to $2.94\times10^{-6}$ (OR=0.5). The SNP had a mean minor allele frequency of 0.057 and a mean imputation quality score of 0.98. Suggestive SNPs at this locus were in high LD ($r^2=0.642$) and reside in introns or in close proximity to the gene SLC4A8 (Figure 2B).

At the third locus 20p12.1 in the analysis without adjustment for WML burden at baseline, the top SNP was rs6135309 with a $P$ value of $3.69\times10^{-6}$, an OR of 0.5, a mean minor allele frequency of 0.29, and a mean imputation quality score of 0.97. Both suggestive variants at this locus were in LD ($r^2=1$) and reside in introns of MACROD2 (Figure 2C).

The fourth locus suggestive of an association with WML progression was located on chromosome 4p15.31 in the analysis with adjustment for WML burden at baseline. The top SNP at this locus was rs7664442 ($P=2.26\times10^{-6}$; OR=0.65), which had a mean minor allele frequency of 0.079 and a mean imputation quality score of 0.96. These 3 SNPs at 4p15.31 were in perfect LD ($r^2=1$) and are located in the intergenic region near GBA3 (Figure 2D).

Functional annotation of all highly suggestive SNPs and SNPs in LD with them was performed. The 2 suggestive exonic variants at locus 10q24.32, rs2275271 ($P=2.26\times10^{-6}$; OR=0.79), and rs3740387 ($P=5.10\times10^{-6}$; OR=1.25) were synonymous. Two other suggestive SNPs at this locus have been reported in previous GWASs investigating schizophrenia46–48 (rs7897654; $P=5.26\times10^{-4}$; OR=0.79 and rs7914558;...
No association between suggestive SNPs and quantitative gene expression traits could be determined.

A comparison of the effect size of suggestive SNPs on WML progression between studies with short- versus long-term follow-up failed to demonstrate significantly increased effect sizes in studies with long-term follow-up. There were also no significant differences in the effect sizes of suggestive SNPs on WML progression between younger and older cohorts (data not shown).

**Performance of WML progression risk models**

As can be seen from Table, model 1, which included demographics and risk factors, explained between 2.3% and 14.9%...
of the variance in WML progression in the various cohorts. Among studies that assessed WML progression by volume change, 17.2% to 43.3% of the WML progression variance was explained if baseline WML was included (model 1). In CHS, which used a visual rating scale for assessment of WML progression, only 4.5% of the variance could be explained. Highly suggestive SNPs from the current GWAS and all genetic variants previously described as being related to WML accounted for additional 1.1% to 8.6% (model 2) and 0.8% to 11.7% (model 3) of the variance of WML progression beyond age, sex, vascular risk factors, and baseline WML volume.

**Discussion**

In this first GWAS on WML progression, our data indicate that genetic factors contribute only little to WML progression in the general elderly population. This conclusion relies on 3 major findings. First, the family-based heritability estimate for WML progression was only 6.5% versus a heritability of 55% for baseline WML burden in the same sample. Second, our genome-wide analysis yielded no associations at a genome-wide significance level. Third, risk prediction models including highly suggestive SNPs according to the present GWAS meta-analysis plus all genetic polymorphisms that have been shown to be associated with WML burden in previous literature explained only between 0.8% and 11.7% more of the variance of WML progression than that explained by baseline clinical and MRI data. These results for WML progression oppose previous cross-sectional studies in which the contribution of genetic factors for WML burden was substantial. This finding is puzzling because cross-sectionally assessed WML burden and WML progression are both quantitative measures with presumably the same biological basis. In this context it is of note, however, that in our study initial burden of WML predicted annual WML volume increase moderately, but this relationship was not significant. Moreover, for both, brain volume and cognitive functions, similar results of limited genetic variance in longitudinal measures have been shown. In line with our findings on WMLs, these investigations also found that measures of brain atrophy and cognitive performance assessed cross-sectionally are highly heritarily but predominantly environmental factors account for the rate of change of these cerebral phenotypes over time.

We identified 4 suggestive loci for WML progression on chromosomes 10q24.32, 12q13.13, 20p12.1, and 4p15.31. None of the SNPs in these loci reached genome-wide significance; however, some of the neighboring genes within these loci have been related to neuropsychiatric or vascular diseases. Locus 10q24.32 includes AS3MT (arsenic [+3 oxidation state] methyltransferase), CNNM2 (Cyclin M2), and NT5C2 (5’-nucleotidase, cytosolic II). These genes have previously been identified to be associated with schizophrenia and blood pressure in GWAS and replication studies. CNNM2 and NT5C2 were associated with coronary artery disease, and CNNM2 was additionally associated with intracranial aneurysm. Moreover, a suggestive variant in an intron of AS3MT, rs10748835 (P = 2.79 × 10⁻⁶; OR = 1.25), is known to modify cognitive function in persons with low-level arsenic exposure.

The locus at 12q13.13 includes SLCA4A8, a sodium bicarbonate cotransporter of the designated solute carrier family 4. It transports sodium and bicarbonate ions across the cell membrane. The gene is highly expressed in all major regions of the brain and is involved in pH regulation in human neurons.

<table>
<thead>
<tr>
<th>Table. Performance of WML Progression Risk Models</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model</strong></td>
</tr>
<tr>
<td>ASPS Without WML burden adjustment</td>
</tr>
<tr>
<td>With WML burden adjustment</td>
</tr>
<tr>
<td>CHS Without WML burden adjustment</td>
</tr>
<tr>
<td>With WML burden adjustment</td>
</tr>
<tr>
<td>FHS Without WML burden adjustment</td>
</tr>
<tr>
<td>With WML burden adjustment</td>
</tr>
<tr>
<td>PROSPER Without WML burden adjustment</td>
</tr>
<tr>
<td>With WML burden adjustment</td>
</tr>
<tr>
<td>RS II Without WML burden adjustment</td>
</tr>
<tr>
<td>With WML burden adjustment</td>
</tr>
<tr>
<td>RS III Without WML burden adjustment</td>
</tr>
<tr>
<td>With WML burden adjustment</td>
</tr>
<tr>
<td>3C-Dijon Without WML burden adjustment</td>
</tr>
<tr>
<td>With WML burden adjustment</td>
</tr>
</tbody>
</table>

Data are Nagelkerke $R^2$. Model 1: WML progression = age + sex + time + hypertension + diabetes mellitus + current smoking* without or with adjustment for WML burden at baseline. Model 2: WML progression = age + sex + time + hypertension + diabetes mellitus + current smoking + GWAS SNPs without or with adjustment for WML burden at baseline. Model 3: WML progression = age + sex + time + hypertension + diabetes mellitus + current smoking + GWAS SNPs + previously reported SNPs + APOE4 without or with adjustment for WML burden at baseline. 3C-Dijon indicates three-City Dijon study; ASPS, Austrian Stroke Prevention Study; CHS, Cardiovascular Health Study; FHS, Framingham Heart Study; GWAS, genome-wide association study; PROSPER, Prospective Study of Pravastatin in the Elderly at Risk; RS II, Rotterdam Study II; SNP, single-nucleotide polymorphism; and WML, white matter lesion.

*Frequency of current smoking ranged between 5.9% in 3C-Dijon and 24.5% in RS III.
MACROD2 (MACRO domain containing 2) at locus 20p12.1 is a protein-coding gene that has been found to be associated with MRI-defined brain infarcts in a previous GWAS.64 This is particularly intriguing as recent studies have suggested new lacunes are most likely to develop in areas of white matter progression.65 MACROD2 has also been associated with autistic traits.66

Although we included a large number of participants from 10 cohort studies with longitudinal assessment of WML change, we cannot exclude the possibility that we missed significant associations of WML progression to SNPs with small effect sizes or low-risk allele frequencies. Moreover, 3 other factors need to be considered when interpreting our results. First, we used a binary phenotype based on visual rating or cut-off values of volumetric change to define WML progression. This conservative definition decreased the likelihood of false-positive ratings but, among those with WML at baseline, might have led to an underestimation of WML progression. Second, the average time period between scans among contributing cohorts was 54 months, and despite the fact that single studies had followed their participants for >10 years, it is conceivable that this period was too short to reveal the full impact of genetic factors on WML progression. Our findings that the effect sizes of suggestive GWAS SNPs on WML progression were not significantly larger in cohorts with long-term follow-up versus cohorts with short-term follow-up and in older versus younger cohorts do not support this assumption. Third, we by design concentrated on elderly people. Genetic factors may play a larger role for white matter progression in younger populations.

Our study findings have important implications for future research on age-related white matter changes. They suggest that, although the contribution of genetic factors seems to be large during the initiating phase of white matter damage, the propagating phase of WML seems to be mainly influenced by nongenetic determinants. With the exception of high blood pressure, these nongenetic risk factors for WML progression remain largely unknown.67,68 On the basis of our data, we need to intensify the search for potentially modifiable environmental and lifestyle factors that influence the progression of age-related white matter changes and the associated morbidity. Moreover, although this study focused on the influence of genetics on age-related WML progression, the effects of heritability on injury-induced WML progression are an interesting direction for future studies.

Acknowledgments

Atherosclerosis Risk in Communities (ARICs) thank the staff and participants of the ARIC study for their important contributions. Austrian Stroke Prevention Study (ASPS) thank the staff and the participants of the ASPS for their valuable contributions. 3C-Dijon thank the staff and the participants of the 3C Study for their important contributions. 3C-Dijon thank Anne Boland (Centre National de Génotypage, Institut de Génomique, CEA) for her technical help in preparing the DNA samples for analyses. Study concept design was performed by Dr R. Schmidt, Dr H. Schmidt, Dr S. Seshadri, and Dr DeCarli; data analysis was carried out by Drs Hofer and Cavaleri; article was prepared by Drs Hofer, Cavaleri, R. Schmidt, and H. Schmidt; cohort contributions (alphabetical order); Study concept design was performed by Aging Gene-Environment Susceptibility Reykjavik Study (AGES-Reykjavik): Drs Launer and Gudnason; ARIC: Drs Fornage and Mosley; ASPS: Drs H. Schmidt and R. Schmidt; Cardiovascular Health Study (CHS): Drs Psaty and Longstreth; Framingham Heart Study (FHS): Dr Seshadri; Prospective Study of Pravastatin in the Elderly at Risk (PROSPER): Drs Buckley, Ford, Stott, Sattar, Westendorp, and Jukema; Rotterdam Study (RS): Dr Ikram; Tasmanian Study of Cognition and Gait (TASCOG): Dr Srikanth; 3C-Dijon: Drs Dufouil and Tzourio; Phenotype data acquisition/QC was performed by AGES-Reykjavik: S. Sigurdsson, Drs van Buchem, and Zijlendbo; ARIC: Drs Shibata, Windham, Gottesman, Heiss, and Mosley; ASPS: Drs H. Schmidt and R. Schmidt; CHS: Drs Psaty and Longstreth; FHS: Drs Beiser and DeCarli; PROSPER: Drs Buckley, Ford, Stott, Sattar, Westendorp, and Jukema; RS: Drs Vernooij, Niessen, and Ikram; TASCOG: Dr Srikanth, Dr Phan, Dr Callisaya, C. Moran, and Dr Beare; 3C-Dijon: Drs Dufouil, Mazoyer, and Tzourio; Genotype data acquisition/QC was performed by AGES-Reykjavik: Dr Smith; ARIC: Dr Fornage; ASPS: Dr Schmidt and F. Freudenberger; CHS: Drs Bis, Lumley, and Taylor; FHS: Dr Wang and Yang; PROSPER: Drs Trompet, Jukema, Slagboom, and Craen; RS: Drs Hofman and van Duijn; TASCOG: Dr Thomson; 3C-Dijon: Drs Wolf, Debette, Chauhan, and Amouyel; Data analysis was carried out by AGES-Reykjavik: Dr Smith; ARIC: Dr Fornage; ASPS: Dr Schmidt and F. Freudenberger; CHS: Drs Bis, Lumley, and Taylor; FHS: Dr Wang and Yang; PROSPER: Drs Trompet, Jukema, Slagboom, and Craen; RS: Drs Hofman and van Duijn; TASCOG: Dr Thomson; 3C-Dijon: Drs Wolf, Debette, Chauhan, and Amouyel; Critical revision of article for Important Intellectual content was done by all the authors.

Sources of Funding

Aging Gene-Environment Susceptibility-Reykjavik Study: The research has been funded by the National Institute on Aging (NIA) contract N01-AG-12100 with contributions from the National Eye Institute, the National Institute on Deafness and Other Communication Disorders and the National Heart, Lung, and Blood Institute (NHLBI), the NIA Intramural Research Program, Hjartavermd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament). Atherosclerosis Risk in Communities Study: The research is carried out as a collaborative study supported by NHLBI contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN26820110009C, HHSN26820110010C, HHSN26820110011C, and HHSN268201100012C), R01HL087641, R01HL59367, and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health (NIH) contract HHSN268200625226C. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the NIH and NIH Roadmap for Medical Research. Funds for this project were also supported by grant HL093029 to Dr Fornage. Austrian Stroke Prevention Study (ASPS): The research reported in this article was funded by the Austrian Science Fund (Fonds zur Förderung der wissenschaftlichen Forschung) grant number P20545-P05 and P13180. The Medical University of Graz supports the databank of the ASPS. Cardiovascular Health Study (CHS): The research was supported by NHLBI contracts HHSN2682012000036C, HHSN268200800007C, N01HC5222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086, and N01HC15103; and NHLBI grants U01HL082925, R01HL087652, R01HL105756, R01HL103612, and R01HL120393 with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through R21AG023629 from the NIA. A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. The protocol of genotyping data was supported, in part, by the National Center for Advancing Translational Sciences, University of California Los Angeles Clinical and Translational Science Institute grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The content is solely the responsibility of the author(s) and does not necessarily represent the official views of the NIH. Framingham Heart Study (FHS): From the FHS of the NHLBI of the NIH and Boston University School of Medicine. This work was supported by the NHLBI’s FHS (contract no. N01-HC-25195) and its
Disclosures

Dr Amouyal has received personal fees from Servier, Hoffman Laroche, Total, Genoscreen, Alzprotect, and Fondation Plan Alzheimer. Dr Dufouil has received payment for lectures from the American Academy of Neurology. Dr Lumley has received support for travel to the NIH/NIHBL, Dr Niessen is a co-founder and scientific director of Quantib BV and owns stock/stock options from Quantib BV. Dr Phan is a member of the Advisory Board for Genzyme on Fabry Disease and has received payment for lectures from Bayer, Pfizer, Genzyme, and Boehringer Ingelheim. Dr Psaty has provided service on the data and safety monitoring board for a clinical trial of a device by the manufacturer (Zoll LifeCor) and service on a Steering Committee for the Yale Open Data Access Project funded by Johnson & Johnson. Dr R. Schmidt has received consulting fees from Axon Neurosciences, Avraham Pharmaceuticals, and Pfizer. Dr Verhaaren has received support for travel from Nederlandse Hartstichting, 2009B102. Dr Zijdenbos has received consulting fees and support for travel from the National Institute on Aging and has been employed or is currently employed at Prodeema Medical, Montreal Neurological Institute, and Biospective Inc. The other authors report no conflicts.

References


SUPPLEMENTAL MATERIAL

White matter lesion progression: A genome-wide search for genetic influences


Corresponding author: Reinhold Schmidt, Medical University of Graz, Department of Neurology, Division of Neurogeriatrics, Auenbruggerplatz 22, 8036 Graz, Austria
reinhold.schmidt@medunigraz.at

Contents:

Section I: Study descriptions
  Aging Gene-Environment Susceptibility - Reykjavik Study (AGES-Reykjavik)
  The Atherosclerosis Risk in Communities Study (ARIC)
  The Austrian Stroke Prevention Study (ASPS)
  The Cardiovascular Health Study (CHS)
  Framingham Heart Study (FHS)
  The Prospective Study of Pravastatin in the Elderly at Risk (PROSPER)
  Rotterdam Study (RS II and RS III)
  Tasmanian Study of Cognition and Gait (TASCOG)
  The Three-City Dijon study (3C-Dijon)

Section II: Magnetic resonance imaging

Section III: Phenotype harmonization

Supplemental Table I: Characteristics of study participants
Supplemental Table II: Genotyping quality control
Supplemental Table III: SNP-level QC
Supplemental Table IV: Imputation
Supplemental Table V: GWAS results without adjustment for WMH load at baseline
Supplemental Table VI: GWAS results with adjustment for WMH load at baseline
Supplemental Figure I: Quantile-quantile (QQ) plots (analyses without adjustment for WMH volume at baseline)
Supplemental Figure II: Quantile-quantile (QQ) plots (analyses with adjustment for WMH volume at baseline)
Supplemental Figure III: Scatterplot (mean WML volume at baseline versus mean annual WML volume increase)

Supplemental References
Section I: Study descriptions.

Aging Gene-Environment Susceptibility - Reykjavik Study (AGES-Reykjavik)
The AGES-Reykjavik Study is a single center prospective cohort study based on the Reykjavik Study. The Reykjavik Study was initiated in 1967 by the Icelandic Heart Association to study cardiovascular disease and risk factors. The cohort included men and women born between 1907 and 1935 who lived in Reykjavik at the 1967 baseline examination. Re-examination of surviving members of the cohort was initiated in 2002 as part of the AGES-Reykjavik Study. The AGES-Reykjavik Study is designed to investigate aging using a multifaceted comprehensive approach that includes detailed measures of brain function and structure. All cohort members were European Caucasians. The study design has been described previously. \(^1\) Briefly, as part of a comprehensive examination, all participants answered a questionnaire, underwent a clinical examination and had blood drawn. All consenting participants without contraindications were offered a brain MRI on a dedicated machine in the study center: a total of 5003 participants had an MRI. \(^2\) Of these, 3664 were genotyped at the Laboratory of Neurogenetics, Intramural Research Program, NIA, Bethesda, Maryland, and 3219 participants passed QC criteria for genotyping. Of these, 224 had complete genotyping and MRI data with assessment of white matter hyperintensity burden at baseline and follow up; 21 participants with a history of transient ischemic attack or stroke were excluded, leaving 203 for these analyses.

Atherosclerosis Risk in Communities Study (ARIC)
The ARIC study is a prospective population-based study of atherosclerosis and clinical atherosclerotic diseases in 15,792 men and women, including 11,478 white participants, drawn from 4 United States communities (Suburban Minneapolis, Minnesota; Washington County, Maryland; Forsyth County, North Carolina; and Jackson, Mississippi). In the first 3 communities, the sample reflects the demographic composition of the community. In Jackson, only black residents were enrolled. Participants were between age 45 and 64 years at their baseline examination in 1987-1989. \(^3\) Vascular risk factors and outcomes, including transient ischemic attack, stroke and dementia, were determined in a standard fashion. \(^4\) During the first 2 years (1993-1994) of the third ARIC examination, participants aged 55 and older from the Forsyth County and Jackson sites were invited to undergo cranial MRI. This subgroup of individuals with MRI scanning represents a random sample of the full cohort because examination dates were allocated at baseline through randomly selected induction cycles. All subjects without MRI contra-indications who completed the visit 3 MRI exam were invited for a 2nd MRI scan between 2004-2006. A total of 989 participants, including 487 white participants, completed an MRI scan at baseline and follow up. Among these, 419 participants had genome-wide genotype data available for these analyses.

The Austrian Stroke Prevention Study (ASPS)
The ASPS study is a single center prospective follow-up study on the effects of vascular risk factors on brain structure and function in the normal elderly population of the city of Graz, Austria. The procedure of recruitment and diagnostic work-up of study participants has been described previously. \(^5, 6\) A total of 2007 participants were randomly selected from the official community register stratified by gender and 5 year age groups. Individuals were excluded from the study if they had a history of neuropsychiatric disease, including previous stroke, transient ischemic attacks, and dementia, or an abnormal neurologic examination determined on the basis of a structured clinical interview and a physical and neurologic examination. During 2 study periods between September 1991 and March 1994 and between January 1999 and December 2003 an extended diagnostic work-up including MRI and neuropsychological testing was done in 1076 individuals aged 45 to 85 years randomly selected from the entire cohort: 509 from the first period and 567 from the second. In 1992, blood was drawn from all study participants for DNA
extraction. They were all European Caucasians. Genotyping was performed in 996 participants, and the 443 who also underwent MRI scanning at baseline and follow up with assessment of white matter hyperintensity burden were available for these analyses.

The Cardiovascular Health Study (CHS)
The CHS is a population-based cohort study of risk factors for vascular disease in adults 65 years or older conducted across 4 field centers in the United States: Sacramento County, California; Washington County, Maryland; Forsyth County, North Carolina; and Pittsburgh, Allegheny County, Pennsylvania. The original predominantly white cohort of 5,201 persons was recruited in 1989-1990 from a random sample of people on Medicare eligibility lists. An additional 687 African-Americans were enrolled in 1992-1993, for a total sample of 5,888. Vascular risk factors and outcomes, including transient ischemic attack, stroke and dementia, were determined in a standard fashion. DNA was extracted from blood samples drawn on all participants who consented to genetic testing at their baseline examination in 1989-90 or 1992-1993. In 2007-2008, genotyping was performed at the General Clinical Research Center's Phenotyping/Genotyping Laboratory at Cedars-Sinai on 3980 CHS participants who were free of cardiovascular disease at baseline and who had DNA available for genotyping. Because most other cohorts were predominantly white, the African-American participants were excluded from this analysis to limit the potential for false positive associations due to population stratification. Among these white participants, genotyping was attempted in 3,397 participants and was successful in 3,295 persons. Of these participants, 1223 had MRI scans performed with assessment of white matter hyperintensity burden at baseline and follow up and were available for these analyses.

Framingham Heart Study (FHS)
The FHS is a three-generation, single-site, community-based, prospective cohort study that was initiated in 1948 to investigate risk factors for cardiovascular disease including stroke. It now comprises 3 generations of participants: the original cohort followed since 1948 (Original); their offspring and spouses of the offspring, followed since 1971 (Offspring); and children from the largest offspring families enrolled in 2000 (Gen 3). The Original cohort enrolled 5209 men and women who comprised two-thirds of the adult population then residing in Framingham, MA, USA. Survivors continue to receive biennial examinations. The Offspring cohort comprises 5,124 persons (including 3,514 biological offspring) who have been examined approximately once every 4 years. Participants in the first two generations were invited to undergo an initial brain MRI in 1999-2005. Brain MRI in Gen 3 only began in 2009 and is not included in these analyses as they had only completed one brain MRI scan by 2013. The population of Framingham was virtually entirely white in 1948 when the Original cohort was recruited. Vascular risk factors and outcomes, including transient ischemic attack, stroke and dementia, were identified prospectively since 1948 through an ongoing system of FHS clinic and local hospital surveillance. Participants had DNA extracted and provided consent for genotyping in the 1990s. Genotyping was performed at Affymetrix (Santa Clara, CA) through an NHLBI funded SNP-Health Association Resource (SHARE) project. Genotyping was attempted in 5,293 participants from Original and Offspring cohorts, and 4,519 persons met QC criteria. Failures (call rate <97%, extreme heterozygosity or high Mendelian error rate) were largely restricted to persons with whole-genome amplified DNA and DNA extracted from stored serum samples. Of these 4,519 persons 4,116 were alive in 1999 when the MRI study began. Of these, 1,570 participants from the Original and Offspring cohorts have undergone two cranial MRI with measurement of white matter hyperintensity burden during the periods of 1999-2005 and 2001-2007 respectively. Of these, 25 participants were excluded for stroke or TIA, 5 for decreased white matter hyperintensity volume in second visit and 164 for lack of GWAS data or covariates. The remaining 1376 participants constitute the FHS sample for this study.
The Prospective Study of Pravastatin in the Elderly at Risk (PROSPER)

PROSPER is a prospective, randomized, double-blind placebo-controlled trial to investigate the effect of pravastatin on cardiovascular and cerebrovascular events in elderly individuals who have vascular disease or who are at high risk to develop vascular disease. The study is multicentric and between December 1997 and May 1999 individuals were enrolled and screened in Glasgow (Scotland), Cork (Ireland) and Leiden (The Netherlands). A total of 5804 elderly men and women aged between 72 and 80 years were included in the study if they suffered from coronary, cerebral or peripheral vascular disease or if they had a high risk of developing such a disease because of smoking, hypertension or diabetes. Subjects with poor cognitive function (mini mental state examination score < 24) were excluded from the study. A whole genome wide screening has been performed in the sequential PHASE project with the use of the Illumina 660K beadchip. Of 5,763 subjects DNA was available for genotyping. Genotyping was performed with the Illumina 660K beadchip, after QC (call rate <95%) 5,244 subjects and 557,192 SNPs were left for analysis. These SNPs were imputed to 2.5 million SNPs based on the HAPMAP built 36 with MACH imputation software.

Rotterdam Study (RS II and RS III)

The Rotterdam Study is a population-based cohort study among inhabitants of a district of Rotterdam (Ommoord), The Netherlands, and aims to examine the determinants of disease and health in the elderly with a focus on neurogeriatric, cardiovascular, bone, and eye disease. In 1990-1993, 7,983 persons participated and were re-examined every 3 to 4 years (Rotterdam Study I). In 2000-2001 the cohort was expanded by 3,011 persons who had not yet been part of the Rotterdam Study (Rotterdam Study II). All participants had DNA extracted at their first visit. Genotyping was attempted in participants with high-quality extracted DNA in 2007-2008. Genotyping was done at the Human Genotyping Facility, Genetic Laboratory Department of Internal Medicine, Erasmus MC, Rotterdam, the Netherlands. In 2005-2006, 895 non-demented persons of the 3011 participants from the Rotterdam Study II were randomly selected to undergo cranial MRI scanning. Both quantitative assessment of white matter hyperintensity volume and genome-wide genotype data was available in 591 individuals. Once participants enter the Rotterdam Study, history of transient ischemic attack and stroke is obtained through self-report and medical records. Subsequently they are continuously monitored for major events, including transient ischemic attack and stroke, by automated linkage of the general practitioners’ records and hospital discharge files with the study database. After exclusion of 53 participants with prevalent stroke or TIA at the time of MRI, 538 participants from the Rotterdam Study II were available for the present analysis. For RS III, 3122 persons received two MRI scans between 2006 and 2013, of which 2925 remained after exclusion of prevalent stroke or dementia at the baseline scan, or missing information on these diagnoses. Genotyping was performed in 1782 persons of this sample.

Tasmanian Study of Cognition and Gait (TASCOG)

TASCOG is a study of cerebrovascular mechanisms underlying gait, balance and cognition in a population-based sample of Tasmanian people aged at least 60 years. Individuals aged 60–86 years (n = 395) living in Southern Tasmania, Australia, were randomly selected from the electoral roll between 2006 and 2008 to participate in the study. Individuals were excluded if they lived in a nursing home, had a contraindication for magnetic resonance scanning (MRI) or were unable to walk without a gait aid. Participants underwent brain MRI scans and genotyping. DNA was extracted from peripheral blood samples by proteinase K digestion following cell lysis, then phenol-chloroform purification. DNA was genotyped using Illumina Hap370CNV chips at the Diamantina Institute and Institute of Molecular Biosciences, University of Queensland, Australia, for 370 participants, and call rates were greater than 97% for all samples. Genotypes for 22
individuals were excluded, either because they were closely related to other individuals, they were outliers in a population ancestry analysis or their sex predicted from genotypes did not match sex as recorded in the database. Among the 348 remaining participants with available genome-wide data, 231 participants had baseline and follow up MRI. After exclusion of 1 participant with dementia and 10 with infarcts on MRI and 3 with insufficient MRI image quality, 217 individuals were available for the present analysis.

The Three-City Dijon Study (3C-Dijon)
The 3C is a cohort study conducted in three French cities (Bordeaux, Dijon, Montpellier), comprising 9,294 participants, designed to estimate the risk of dementia and cognitive impairment attributable to vascular factors. Eligibility criteria included living in the city and being registered on the electoral rolls in 1999, 65 years or older, and not institutionalized. The study protocol was approved by the Ethical Committee of the University Hospital of Kremlin-Bicêtre and each participant signed an informed consent. Data reported in this article were obtained in Dijon (3C-Dijon study), where 4,931 individuals were recruited (1999–2001). The overall design of the 3C-Dijon study is detailed elsewhere. Participants aged less than 80 years and enrolled between June 1999 and September 2000 (n=2,763) were invited to undergo a brain MRI. Although 2,285 subjects agreed to participate (82.7%), because of financial limitations, 1,924 MRI scans were performed, of which 1,795 had WMH data. After 4 years, 1,328 individuals had a repeated MRI and WMH data available. Of these, 1,164 participants had genome-wide genetic data and WMH volume measures at baseline and follow up. After exclusion of individuals with prevalent dementia (N=4) or stroke (N=47), 1,113 participants were available for analyses. DNA samples of 3C-Dijon participants were genotyped at the Centre National de Génotypage, Evry, France (www.cng.fr).
Section II: Magnetic resonance imaging.

In all of the studies, the MRI scans were performed in eligible participants in a standardized fashion and interpreted without knowledge of demographic or clinical information. The field strength of the scanners used ranged from 0.5T to 1.5 T. T1- and T2-weighted scans in the axial plane were obtained for all participants. These were complemented by either scans obtained with fluid attenuation inversion recovery or proton density sequences to allow better separation of white matter hyperintensities and cerebrospinal fluid.

AGES-Reykjavik
Images were acquired on a 1.5 T Signa TwinSpeed system (General Electric Medical Systems, Waukesha, Wisconsin, USA). The image protocol consisted of the following pulse sequences: a T2*-weighted gradient echo type echo planar sequence; a proton density/T2-weighted fast spin echo sequence; and a fluid attenuated inversion recovery (FLAIR) sequence. These sequences were acquired with 3 mm thick interleaved slices. Images were also acquired with a T1-weighted three-dimensional spoiled gradient echo sequence with slice thickness being 1.5 mm.

ARIC
Both MRI examinations used General Electric (General Electric Medical Systems) or Picker (Picker Medical Systems) 1.5-Tesla scanners. The scanning protocol included a series of sagittal T1-weighted scans and axial proton-density, T2-weighted and T1-weighted scans with 5 mm thickness and no interslice gaps. At the first MRI examination, WMH severity was graded from proton density-weighted images using a 0 to 9 scale (semi-quantitative) developed in CHS. At the second MRI examination, WMH severity was measured, both, using a semi-quantitative white matter grade and a semi-automated volumetric analysis. WMH volume was determined from axial fluid-attenuated inversion recovery (FLAIR) images.

ASPS
MRI was performed on 1.5-Tesla whole body imaging systems (Gyroscan S 15 and ACS, Philips Medical Systems, Eindhoven, The Netherlands) using axial proton-density and T2-weighted sequences. Additionally, T1-weighted images were acquired in the sagittal plane. For all images, slice thickness was 5 mm with no interslice distance.

CHS
Magnetic resonance imaging was performed on General Electric or Picker 1.5-Tesla scanners at 3 field centers and on a 0.35-Tesla Toshiba scanner at the fourth. The scanning protocol included a series of sagittal T1-weighted scans and axial proton-density, T2-weighted and T1-weighted scans with 5 mm thickness and no interslice gaps. Both ARIC and CHS used the same protocols for scanning and for interpretation.

FHS
Participants were evaluated with a 1 or 1.5-Tesla Siemens Magnetom scanner. 3D T1 and double echo proton density (PD) and T2 coronal images acquired in 4-mm contiguous slices were performed.

PROSPER
All imaging was performed on an MR system operating at field strength of 1.5 Tesla (Philips Medical Systems, Best, the Netherlands). Gray and white matter volumes were calculated by SIENAX technique. In short, SIENAX starts by extracting brain and skull images from the single whole-head input data. The brain image is then affine-registered to Montreal Neurological Institute (MNI) 152 space (by using the skull image to determine the registration scaling), done primarily to obtain the volumetric scaling factor to be used as normalization for head size. Next,
tissue-type segmentation with partial volume estimation is carried out to calculate total volume of brain tissue (including separate estimates of volumes of gray matter, white matter, peripheral gray matter, and ventricular cerebrospinal fluid).

RS II and RS III
In 1995-1996 participants originating from the Rotterdam Study I underwent MRI of the brain on a 1.5-Tesla Siemens Vision scanner. The protocol included axial T1-weighted, T2-weighted, and proton-density scans with slice thickness of 5mm and a high-resolution 3D inversion recovery HASTE scan with slice thickness of 1.25 mm. The slice thickness was 5 or 6 mm with an interslice gap of 20%. In 2005-2006, participants originating from the Rotterdam Study II and III underwent MRI of the brain including axial T1-weighted, proton-density, and FLAIR sequences on a 1.5-Tesla GE Healthcare scanner. The slice thickness was 1.6 mm on the T1-weighted, and proton-density weighted sequences and 2.5 mm on the FLAIR sequence. All slices were contiguous.

TASCOG
MRI scans were performed using a GE 1.5 Tesla scanner, with the following sequences; High resolution T1-weighted spoiled gradient echo (SPGR) MRI scans [TR 35ms, TE 7ms, flip angle 35°, field of view 24 cm, voxel size = 1mm3] comprising 120 contiguous slices; Axial 3-dimensional T-2 weighted fast spin echo images (TR = 4300ms; TE = 106ms; NEX = 1; turbo factor = 48; voxel size = 3 mm3); Axial FLAIR (fluid attenuated inversion recovery) sequence (TR 8802, TE 125, TI 2200, 3mm contiguous thickness).

3C-Dijon
MRI acquisition was performed on a 1.5-Tesla Magnetom scanner (Siemens, Erlangen, Germany) using T1-, T2- and proton density (PD) weighted sequences. A fast multislice, double-echo, T2-weighted, 2-dimensional axial acquisition was used, with 4-mm-thick slices, and 0.4mm between slice spacing.
Section III: Phenotype harmonization.

In AGES-Reykjavik, ASPS, FHS, PROSPER, RS II, RS III, TASCOG and the 3C-Dijon Study, white matter hyperintensity (WMH) volume was estimated by volumetric measurement using custom-written computer programs based on an automatic segmentation algorithm or a semi-automatic segmentation analysis involving operator-guided removal of non-brain elements. In ARIC WMH volume was estimated on a semi-quantitative scale by visual comparison with reference templates that successively increased from barely detectable white matter lesions to extensive, confluent abnormalities, as previously described. In CHS MRI scans were evaluated without knowledge of any clinical information, neuroradiologists at the reading center estimated white matter using a 10-point system, from 0 to 9 (most abnormal), using a library of templates as previously described. White matter lesion (WML) progression was defined as absent or present. In all studies using visual rating of white matter lesion progression we used their definition for progression, while in all studies with volumetric assessment of white matter lesion progression was dichotomized based on volume change below or above one standard deviation of the study specific mean of WMH volume change.

AGES-Reykjavik
Total white matter lesion volume was computed automatically with an algorithm based on the Montreal Neurological Institute pipeline. The AGES-Reykjavik/Montreal Neurological Institute pipeline has been modified to accommodate full brain coverage including cerebellum and brainstem, multispectral images (T1-weighted three-dimensional spoiled gradient echo sequence, FLAIR, and proton density/T2-weighted fast spin echo sequences), high throughput, and minimal editing. Progression was defined as volume change above 2.99 cm$^3$.

ARIC
At the first MRI examination, WMHs were estimated as the relative total volume of periventricular and subcortical white matter signal abnormality on proton density–weighted axial images by visual comparison with eight templates that successively increased from barely detectable white matter changes (Grade 1) to extensive, confluent changes (Grade 8). Individuals with no white matter changes received Grade 0, and those with changes worse than Grade 8 received Grade 9. Unlike the volumetric measurements, the visual grading is inherently normalized for brain volume/size. Study participant’s brain images are compared with the reference standards after zooming them to approximately the same apparent size. Hence, grading of WMH is performed based on the same brain volume/size across all participants. At the second MRI examination, scans were scored for WMH using both a semi-quantitative scale by visual comparison with reference templates and a quantitative scale based on a semi-automated volumetric analysis. An automated algorithm was used to segment each of the FLAIR images into voxels assigned to one of three categories based on signal intensity - normal brain, CSF, or leukoaraiosis. The leukoaraiosis maps were manually edited to exclude infarcts and other lesions. The mean absolute error and test-retest coefficient of variation for this method are 6.6% and 1.4%, respectively, for leukoaraiosis volume. Total intracranial volume was manually measured from T1-weighted sagittal images. Progression was defined based on the semi-quantitative grades by subtracting grade on follow-up scan from grade on initial scan, which had been read in a side-by-side fashion. WMH progression was coded present when the change score was one or more. In addition, WMH volume at the first MRI exam was estimated using a prediction equation ($R^2=0.80$) relating volume from visual grades. The prediction equation was created using actual data from the 2nd MRI exam (visual grades and quantitative volumes). Visual grades from the first exam were then entered into the equation to calculate estimated volumes at the first exam. Change in WMH volume was calculated by subtracting estimated WMH volume at the first exam from WMH volume at the 2nd exam.
ASPS
Lesion load measurements were done on proton density–weighted images on an UltraSPARC workstation (Sun Microsystems) using DISPImage. Using a hard copy with all lesions outlined as a reference, a trained technician outlined all lesions on the computer image with use of a semiautomated segmentation algorithm provided by the DISPImage program. The total lesion volume was calculated by multiplying the total lesion area by slice thickness. Progression was defined as volume change above 0.79 cm$^3$.

CHS
Briefly, without knowledge of any clinical information, neuroradiologists at the reading center estimated white matter grades using a 10-point system, from 0 to 9 (most abnormal), using a library of templates. A change score was calculated by subtracting grade on follow-up scan from grade on initial scan, which had been read in a side-by-side fashion. WMH progression was coded present when the change score was one or more.$^{34}$

FHS
All MR images were transferred to the centralized reading center at the University of California–Davis Medical Center and analyses were performed on QUANTA 6.2, a custom-designed image analysis package operating on a Sun Microsystems Ultra 5 workstation. Semiautomated analysis of pixel distributions based on mathematical modeling of MRI pixel intensity histograms for cerebrospinal fluid (CSF) and brain matter (white matter and gray matter) were used to determine the optimal threshold of pixel intensity to best distinguish CSF from brain matter based on methods published previously.$^{37}$ For segmentation of WMH from other brain tissues, the first and second echo images from T2 sequences were summed and a log-normal distribution was fitted to the summed data. A segmentation threshold for WMH was determined as 3.5 standard deviations (SD) in pixel intensity greater than the mean of the fitted distribution of brain parenchyma.$^{38}$ Progression was defined as volume change above 1.07 cm$^3$.

PROSPER
The algorithm FIRST (FMRIB’s Integrated Registration and Segmentation Tool) was applied to estimate the volume of hippocampus, nucleus accumbens, globus pallidus, amygdala, putamen, caudate nucleus and thalamus. FIRST is part of FSL (FMRIB’s Software Library) and performs both registration and segmentation of the mentioned subcortical regions. To assess cerebral microbleeds, all MRI scans were read in consensus by two experienced raters who were blinded to subjects’ clinical history. Cerebral microbleeds were defined as focal areas of signal loss on T2*-weighted gradient echo pulse sequence (“blooming effect”) that are invisible or smaller on T2-weighted MRI. For each subject the number and location (cortical, subcortical, and infratentorial) of the cerebral microbleeds were recorded. Segmentation of white matter hyperintensities volume was performed automatically using software for Neuro-Image Processing in Experimental Research (SNIPER), an in-house developed program for image processing.$^{39}$ This segmentation was based on the T2-weighted and FLAIR images. Cerebral infarcts were defined as parenchymal defects seen on FLAIR images with the same signal intensity as CSF and a surrounding rim of high signal intensity following a vascular distribution.

RS II and RS III
For WMH quantification a custom-designed fully automated white matter lesion segmentation method was used.$^{40}$ In scans from Rotterdam Study I, cerebrospinal fluid, gray matter, white matter and white matter hyperintensities were segmented by a manual-training-based k-nearest neighbour classifier on multi-modal magnetic resonance imaging data. In scans from Rotterdam Study II and III, an atlas-based k-nearest neighbour classifier was used to segment cerebrospinal fluid, grey matter, and white matter. This classifier was trained by registering brain
atlases to the subject. White matter hyperintensities were subsequently segmented by using the resulting gray matter segmentation to automatically find a white matter hyperintensity threshold in a fluid-attenuated inversion recovery scan. False-positive lesions were removed by ensuring that the lesions are within the white matter. In all scans, visual checks were performed and if needed any segmentation errors manually corrected. Progression was defined as volume change above 2.11 cm$^3$.

**TASCOG**

All scans were registered to a standard 152-brain Montreal Neurological Institute (MNI) template in stereotaxic coordinate space. Using T1 sequences and methods based on statistical parametric mapping software (SPM5), brain tissue was classified as gray matter, white matter, or cerebrospinal fluid. Fully automated morphological segmentation with adaptive boosting classification was applied for both baseline and follow-up scans, using FLAIR and T1- and T2-weighted scans to identify WMLs. The WML volumes estimated using this approach closely corresponded with expert manual segmentation (intraclass correlation (ICC) 0.90, 95% confidence interval (CI) = 0.80–0.95, n = 30). Two stroke experts determined the presence and number of MRI infarcts by consensus, with infarct defined as a hypointensity ≥3mm in diameter on T1-weighted and FLAIR images, with a surrounding hyperintense rim on FLAIR taking care not to misclassify perivascular spaces as infarcts. Progression was defined as volume change above 1.89 cm$^3$.

**3C-Dijon**

Raw data were transferred for analysis and storage to the MRI study center (Department of Neurofunctional Imaging, Caen). Fully automated image processing software was developed to detect, measure, and localize WMH. Measurement of WMH volume progression was performed using a tailored algorithm as previously described. This method uses a common space (white matter mask volume) specific to each participant to accurately quantify the evolution of WMH, taking into account differences in head size, and allows quantification of already existing WMH at baseline and identification of new WMH on follow-up images. Progression was defined as volume change above 0.65 cm$^3$. 
## Supplemental Table I. Characteristics of study participants.

<table>
<thead>
<tr>
<th></th>
<th>AGES-Reykjavik</th>
<th>ARIC*</th>
<th>ASPS</th>
<th>CHS*</th>
<th>FHS</th>
<th>PROSPER</th>
<th>RS II</th>
<th>RS III</th>
<th>TASCOG</th>
<th>3C-Dijon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects, N</td>
<td>203</td>
<td>419</td>
<td>443</td>
<td>1223</td>
<td>1376</td>
<td>459</td>
<td>538</td>
<td>1782</td>
<td>217</td>
<td>1113</td>
</tr>
<tr>
<td>Number of Women, N (%)</td>
<td>118 (58.0)</td>
<td>256 (61.1)</td>
<td>232 (52.3)</td>
<td>744 (54.1)</td>
<td>202 (44.0)</td>
<td>270 (50.2)</td>
<td>984 (55.2)</td>
<td>94 (43.3)</td>
<td>694 (62.3)</td>
<td></td>
</tr>
<tr>
<td>Age at MRI, mean±SD</td>
<td>75.1±5.0</td>
<td>62.3±4.2</td>
<td>63.4±7.5</td>
<td>74.1±4.3</td>
<td>62.4±10.2</td>
<td>75.0±3.2</td>
<td>67.1±5.2</td>
<td>56.4±5.8</td>
<td>70.8±6.7</td>
<td>72.3±4.0</td>
</tr>
<tr>
<td>Age at MRI, range</td>
<td>68-90</td>
<td>52-71</td>
<td>46-83</td>
<td>67-92</td>
<td>42-92</td>
<td>70-83</td>
<td>61-88</td>
<td>46-84</td>
<td>61-86</td>
<td>65-82</td>
</tr>
<tr>
<td>Hypertension, N (%)</td>
<td>160 (78.8)</td>
<td>121 (29.5)</td>
<td>294 (66.4)</td>
<td>598 (48.9)</td>
<td>454 (33.0)</td>
<td>293 (63.8)</td>
<td>332 (61.9)</td>
<td>784 (44.0)</td>
<td>160 (73.7)</td>
<td>834 (74.9)</td>
</tr>
<tr>
<td>Systolic blood pressure, mean±SD</td>
<td>144.3±19.3</td>
<td>121.9±17.4</td>
<td>140.4±21.8</td>
<td>132.6±19.6</td>
<td>126.1±17.7</td>
<td>157.8±21.7</td>
<td>142.9±18.5</td>
<td>131.6±18.4</td>
<td>143.2±23.5</td>
<td>148.3±22.8</td>
</tr>
<tr>
<td>Diastolic blood pressure, mean±SD</td>
<td>75.9±9.5</td>
<td>69.6±9.7</td>
<td>86.5±10.0</td>
<td>70.2±10.0</td>
<td>73.3±9.5</td>
<td>86.0±11.1</td>
<td>80.8±10.4</td>
<td>82.2±10.6</td>
<td>80.5±13.3</td>
<td>84.9±11.3</td>
</tr>
<tr>
<td>Diabetes, N (%)</td>
<td>25 (12.3)</td>
<td>32 (7.8)</td>
<td>25 (5.6)</td>
<td>112 (9.2)</td>
<td>138 (10.0)</td>
<td>74 (16.1)</td>
<td>49 (9.1)</td>
<td>115 (6.5)</td>
<td>29 (13.7)</td>
<td>91 (8.2)</td>
</tr>
<tr>
<td>Interval between scans in months, mean±SD</td>
<td>29.1±9.4</td>
<td>123.6±8.4</td>
<td>62.3±22.1</td>
<td>60.1±7.5</td>
<td>70.2±17.7</td>
<td>33.4±1.5</td>
<td>41.6±2.0</td>
<td>48.5±4.5</td>
<td>29.2±5.2</td>
<td>43.1±4.1</td>
</tr>
<tr>
<td>WML at baseline in cm³, mean±SD</td>
<td>17.6±18.2</td>
<td>8.4±4.6**</td>
<td>2.2±4.5</td>
<td>-</td>
<td>1.6±3.3</td>
<td>5.3±9.9</td>
<td>5.9±8.2</td>
<td>3.0±3.8</td>
<td>9.5±7.9</td>
<td>5.4±4.6</td>
</tr>
<tr>
<td>WML at baseline in cm³, median (range)</td>
<td>11.7 (0.5-116.5)</td>
<td>6.3 (3.2-34.3)**</td>
<td>0.5 (0.0-32.5)</td>
<td>-</td>
<td>0.6 (0.0-44.0)</td>
<td>1.6 (0.0-91.3)</td>
<td>3.5 (0.6-103.4)</td>
<td>2.0 (0.2-63.0)</td>
<td>7.6 (1.5-62.2)</td>
<td>4.0 (0.5-46.9)</td>
</tr>
<tr>
<td>Annual WML volume increase in cm³, mean±SD</td>
<td>1.2±1.8</td>
<td>0.3±0.7**</td>
<td>0.2±0.6</td>
<td>-</td>
<td>0.3±0.8</td>
<td>0.8±1.2</td>
<td>1.4±0.7</td>
<td>0.2±0.4</td>
<td>0.7±1.1</td>
<td>0.2±0.5</td>
</tr>
<tr>
<td>WML progression, N (%)</td>
<td>22 (10.8)</td>
<td>74 (18.0)</td>
<td>38 (8.6)</td>
<td>355 (29.0)</td>
<td>118 (8.6)</td>
<td>51 (11.1)</td>
<td>70 (13.0)</td>
<td>133 (7.5)</td>
<td>63 (29.0)</td>
<td>161 (14.5)</td>
</tr>
</tbody>
</table>

*ARIC and CHS used a semi-quantitative grade based on 8 templates (10-point scale).

**WML at baseline was estimated from a prediction equation that related volume from visual grades. The prediction equation was developed from visual grades and quantitative volumes measured on the participants at a later visit. Details of the methods are described in Gottesman et al. 2010. 

<table>
<thead>
<tr>
<th>AGES-Reykjavik</th>
<th>ARIC</th>
<th>ASPS</th>
<th>CHS</th>
<th>FHS</th>
<th>PROSPER</th>
<th>RS II</th>
<th>RS III</th>
<th>TASCOG</th>
<th>3C-Dijon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genotyping platforms, SNP panel</td>
<td>Illumina HumanHap 370CNV Duo BeadChip®</td>
<td>Illumina Human610-Quad BeadChip®</td>
<td>Illumina Human610-CNV Duo BeadChip®</td>
<td>Illumina 660-Quad BeadChips</td>
<td>Illumina HumanHap550 Duo BeadChip® and Illumina Human 610 Quad BeadChip®</td>
<td>Illumina Human 610 Quad BeadChip®</td>
<td>Illumina Human 610 Quad BeadChip®</td>
<td>Illumina Human610-Quad® BeadChips</td>
<td></td>
</tr>
<tr>
<td>Genotyping center</td>
<td>National Institutes on Aging, National Institutes of Health</td>
<td>Broad Institute</td>
<td>Cedars-Sinai’s General Clinical Research Center’s Phenotyping/Genotyping Laboratory</td>
<td>Affymetrix (Santa Clara)</td>
<td>Human Genotyping Facility, Genetic Laboratory Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands</td>
<td>Human Genotyping Facility, Genetic Laboratory Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands</td>
<td>Human Genotyping Facility, Genetic Laboratory Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands</td>
<td>Diamantina Institute and Institute of Molecular Biosciences, University of Queensland, Australia, Centre National de Génotypage, Evry, France</td>
<td></td>
</tr>
<tr>
<td>Genotyping calling algorithm</td>
<td>Illumina Bead Studio</td>
<td>Birdseed</td>
<td>Illumina Bead Studio</td>
<td>Illumina Bead Studio</td>
<td>Illumina Bead Studio</td>
<td>Illumina Genome Studio</td>
<td>Illumina Genome Studio</td>
<td>Illumina GenCall algorithm</td>
<td>Illumina Bead Studio</td>
</tr>
<tr>
<td>Call rate filter</td>
<td>&gt;97%</td>
<td>&gt;95%</td>
<td>&gt;98%</td>
<td>&gt;95%</td>
<td>&gt;97%</td>
<td>&gt;95%</td>
<td>&gt;97.5%</td>
<td>&gt;97.5%</td>
<td>&gt;97%</td>
</tr>
<tr>
<td>Exclusion on ethnicity</td>
<td>Sample is European Caucasian</td>
<td>Non-Whites not included</td>
<td>Non-European ancestry not included</td>
<td>African American participants were excluded</td>
<td>Sample is European</td>
<td>Non-European ancestry not included</td>
<td>Non-European ancestry not included</td>
<td>Non-European ancestry not included</td>
<td>African-Americans not included</td>
</tr>
<tr>
<td>Other pre-imputation exclusions</td>
<td>sample failures, genotyped sex different from recorded sex, discordance with prior genotyping</td>
<td>sample failures, genotyped sex different from recorded sex, discordance with prior genotyping</td>
<td>sample failures, genotyped sex different from recorded sex, discordance with prior genotyping</td>
<td>sample failures, genotyped sex different from recorded sex, discordance based on pedigree analyses</td>
<td>sample failures, genotyped sex different from recorded sex, excess inter/intra-hetero-zygosity, related individuals</td>
<td>genotyped sex different from recorded sex, excess inter/intra-hetero-zygosity, related individuals</td>
<td>genotyped sex different from recorded sex, excess inter/intra-hetero-zygosity</td>
<td>genotyped sex different from recorded sex, related individuals, outliers in a population ancestry analysis</td>
<td>sample failures, genotyped sex different from recorded sex</td>
</tr>
</tbody>
</table>
**Supplemental Table III. SNP-level QC.**

<table>
<thead>
<tr>
<th></th>
<th>AGES-Reykjavik</th>
<th>ARIC</th>
<th>ASPS</th>
<th>CHS</th>
<th>FHS</th>
<th>PROSPER</th>
<th>RS II</th>
<th>RS III</th>
<th>TASCOG</th>
<th>3C-Dijon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-imputation MAF filter</td>
<td>&lt; 1%</td>
<td>&lt;1%</td>
<td>&lt; 1%</td>
<td>N/A</td>
<td>&lt; 1%</td>
<td>&lt; 1%</td>
<td>&lt; 1%</td>
<td>&lt; 1%</td>
<td>&lt; 0.5%</td>
<td>&lt; 1%</td>
</tr>
<tr>
<td>Pre-imputation HWE filter</td>
<td>P&lt;1x10^{-6}</td>
<td>p&lt;1x10^{-5}</td>
<td>p&lt;1x10^{-6}</td>
<td>P&lt;1 x 10^{-6}</td>
<td>p&lt;1x10^{-6}</td>
<td>p&lt;1x10^{-6}</td>
<td>p&lt;1x10^{-6}</td>
<td>p&lt;1x10^{-7}</td>
<td>p&lt;1x10^{-6}</td>
<td></td>
</tr>
<tr>
<td>Pre-imputation Call frequency filter</td>
<td>&lt; 98%</td>
<td>&lt; 98%</td>
<td>&lt; 98%</td>
<td>&lt;97%</td>
<td>&lt; 97%</td>
<td>&lt; 97.5%</td>
<td>&lt; 97.5%</td>
<td>&lt; 97.5%</td>
<td>&lt;97%</td>
<td>&lt; 98%</td>
</tr>
<tr>
<td>Other pre-imputation SNP filters</td>
<td>No observed hetero-zygotes, Not present in HapMap reference panel</td>
<td>No observed hetero-zygotes, Not present in HapMap reference panel</td>
<td>No observed hetero-zygotes, Not present in HapMap reference panel</td>
<td>No observed hetero-zygotes, missing from dbSNP, and &gt;1 duplicate or Mendelian inconsistency</td>
<td>No observed hetero-zygotes, missing from HapMap reference panel</td>
<td>No observed hetero-zygotes, Not present in HapMap reference panel</td>
<td>No observed hetero-zygotes, Not present in HapMap reference panel</td>
<td>None</td>
<td>Not present in HapMap reference panel</td>
<td></td>
</tr>
<tr>
<td>N autosomal SNPs for imputation</td>
<td>326,034</td>
<td>665,450</td>
<td>550,635</td>
<td>306,655</td>
<td>378,163</td>
<td>557,192</td>
<td>466,389</td>
<td>500,547</td>
<td>313,702</td>
<td>537,029</td>
</tr>
</tbody>
</table>

PLINK: http://pngu.mgh.harvard.edu/~purcell/plink/
## Supplemental Table IV. Imputation.

<table>
<thead>
<tr>
<th></th>
<th>AGES-Reykjavik</th>
<th>ARIC</th>
<th>ASPS</th>
<th>CHS</th>
<th>FHS</th>
<th>PROSPER</th>
<th>RS II</th>
<th>RS III</th>
<th>TASCOG</th>
<th>3C-Dijon</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Imputation software</strong></td>
<td>MACH</td>
<td>MACH v1.0.16</td>
<td>MACH v1.0.15</td>
<td>BIM-BAM15</td>
<td>MACH</td>
<td>MACH 1.0.16</td>
<td>MACH 1.0.16</td>
<td>MACH 1.0.16</td>
<td>MACH 1.0.16</td>
<td>MACH</td>
</tr>
<tr>
<td><strong>Imputation reference panel</strong></td>
<td>HapMap release 22, Build 36</td>
<td>HapMap CEU, Build 35</td>
<td>HapMap release 22, Build 36</td>
<td>HapMap, release 22, Build 36</td>
<td>HapMap release 22, Build 36</td>
<td>HapMap release 22, Build 36</td>
<td>HapMap release 22, Build 36</td>
<td>HapMap, release 22, Build 36</td>
<td>HapMap, release 22, Build 36</td>
<td></td>
</tr>
<tr>
<td><strong>Imputation quality metrics</strong></td>
<td>$R^2$</td>
<td>$R^2$</td>
<td>$R^2$</td>
<td>Observed to expected variance ratio</td>
<td>$R^2$</td>
<td>$R^2$</td>
<td>$R^2$</td>
<td>$R^2$</td>
<td>Squared correlation between imputed and true genotypes</td>
<td>$R^2$</td>
</tr>
<tr>
<td><strong>Other imputation quality filters</strong></td>
<td>N/A</td>
<td>Mono-morphic HWE $p&lt;1\times10^{-6}$ Callrate $&lt;95%$</td>
<td>Quality</td>
<td>N/A</td>
<td>Quality</td>
<td>Quality</td>
<td>Quality</td>
<td>Quality</td>
<td>N/A</td>
<td>Quality</td>
</tr>
<tr>
<td><strong>N genotyped + imputed SNPs for analysis</strong></td>
<td>2,532,729</td>
<td>2,536,406</td>
<td>2,543,887</td>
<td>2,543,887</td>
<td>2,540,223</td>
<td>2,543,887</td>
<td>2,543,887</td>
<td>2,543,887</td>
<td>2,543,149</td>
<td>2,368,202</td>
</tr>
</tbody>
</table>
Supplemental Table V. GWAS meta-analysis result without adjustment for WMH load at baseline

(filtered MAF% 5%, RSQ 0.3).

<table>
<thead>
<tr>
<th>SNP</th>
<th>Locus</th>
<th>Position</th>
<th>Coded Allele</th>
<th>Non-coded Allele</th>
<th>Freq coded allele</th>
<th>Odds ratio</th>
<th>P value</th>
<th>Direction</th>
<th>Function</th>
<th>Closest Gene</th>
<th>Distance to gene (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs4761974</td>
<td>12q13.13</td>
<td>50166712</td>
<td>a</td>
<td>t</td>
<td>0.943</td>
<td>0.50</td>
<td>8.71E-07</td>
<td>---</td>
<td>intronic</td>
<td>SLC4A8</td>
<td>0</td>
</tr>
<tr>
<td>rs10883817</td>
<td>10q24.32</td>
<td>104745421</td>
<td>a</td>
<td>g</td>
<td>0.406</td>
<td>1.27</td>
<td>1.46E-06</td>
<td>+++++++</td>
<td>intronic</td>
<td>CNNM2</td>
<td>0</td>
</tr>
<tr>
<td>rs12255047</td>
<td>10q24.32</td>
<td>104750742</td>
<td>a</td>
<td>g</td>
<td>0.405</td>
<td>1.27</td>
<td>1.63E-06</td>
<td>++++++++</td>
<td>intronic</td>
<td>CNNM2</td>
<td>0</td>
</tr>
<tr>
<td>rs7071373</td>
<td>10q24.32</td>
<td>104761912</td>
<td>a</td>
<td>g</td>
<td>0.59</td>
<td>0.79</td>
<td>1.95E-06</td>
<td>-------</td>
<td>intronic</td>
<td>CNNM2</td>
<td>0</td>
</tr>
<tr>
<td>rs1890185</td>
<td>10q24.32</td>
<td>104738449</td>
<td>a</td>
<td>c</td>
<td>0.593</td>
<td>0.79</td>
<td>2.02E-06</td>
<td>-------</td>
<td>intronic</td>
<td>CNNM2</td>
<td>0</td>
</tr>
<tr>
<td>rs7914558</td>
<td>10q24.32</td>
<td>104765898</td>
<td>a</td>
<td>g</td>
<td>0.417</td>
<td>1.27</td>
<td>2.14E-06</td>
<td>++++++++</td>
<td>intronic</td>
<td>CNNM2</td>
<td>0</td>
</tr>
<tr>
<td>rs10509757</td>
<td>10q24.32</td>
<td>104742950</td>
<td>a</td>
<td>g</td>
<td>0.409</td>
<td>1.26</td>
<td>2.16E-06</td>
<td>++++++++</td>
<td>intronic</td>
<td>CNNM2</td>
<td>0</td>
</tr>
<tr>
<td>rs111191518</td>
<td>10q24.32</td>
<td>104772843</td>
<td>c</td>
<td>g</td>
<td>0.594</td>
<td>0.79</td>
<td>2.23E-06</td>
<td>-------</td>
<td>intronic</td>
<td>CNNM2</td>
<td>0</td>
</tr>
<tr>
<td>rs11118844</td>
<td>1q41</td>
<td>219976911</td>
<td>a</td>
<td>g</td>
<td>0.855</td>
<td>1.46</td>
<td>2.48E-06</td>
<td>++++++++</td>
<td>UTR5</td>
<td>DUSP10</td>
<td>0</td>
</tr>
<tr>
<td>rs1538204</td>
<td>10q24.32</td>
<td>104723883</td>
<td>a</td>
<td>t</td>
<td>0.409</td>
<td>1.26</td>
<td>2.63E-06</td>
<td>++++++++</td>
<td>intronic</td>
<td>CNNM2</td>
<td>0</td>
</tr>
<tr>
<td>rs7096269</td>
<td>10q24.32</td>
<td>104726711</td>
<td>a</td>
<td>g</td>
<td>0.591</td>
<td>0.79</td>
<td>2.72E-06</td>
<td>-------</td>
<td>intronic</td>
<td>CNNM2</td>
<td>0</td>
</tr>
<tr>
<td>rs10748835</td>
<td>10q24.32</td>
<td>104650246</td>
<td>a</td>
<td>g</td>
<td>0.411</td>
<td>1.25</td>
<td>2.79E-06</td>
<td>++++++++</td>
<td>intronic</td>
<td>AS3MT</td>
<td>0</td>
</tr>
<tr>
<td>rs10883805</td>
<td>10q24.32</td>
<td>104696841</td>
<td>t</td>
<td>c</td>
<td>0.592</td>
<td>0.79</td>
<td>2.86E-06</td>
<td>-------</td>
<td>intronic</td>
<td>CNNM2</td>
<td>0</td>
</tr>
<tr>
<td>rs7092200</td>
<td>10q24.32</td>
<td>104834862</td>
<td>t</td>
<td>c</td>
<td>0.57</td>
<td>0.80</td>
<td>3.11E-06</td>
<td>-------</td>
<td>intergenic</td>
<td>NT5C2</td>
<td>2902</td>
</tr>
<tr>
<td>rs4761964</td>
<td>12q13.13</td>
<td>50106221</td>
<td>a</td>
<td>g</td>
<td>0.938</td>
<td>0.60</td>
<td>3.13E-06</td>
<td>---</td>
<td>intronic</td>
<td>SLC4A8</td>
<td>0</td>
</tr>
<tr>
<td>rs10748836</td>
<td>10q24.32</td>
<td>104683907</td>
<td>a</td>
<td>g</td>
<td>0.405</td>
<td>1.26</td>
<td>3.22E-06</td>
<td>++++++++</td>
<td>intronic</td>
<td>CNNM2</td>
<td>0</td>
</tr>
<tr>
<td>rs111191438</td>
<td>10q24.32</td>
<td>104627854</td>
<td>c</td>
<td>g</td>
<td>0.411</td>
<td>1.26</td>
<td>3.28E-06</td>
<td>++++++++</td>
<td>intronic</td>
<td>AS3MT</td>
<td>0</td>
</tr>
<tr>
<td>rs2297786</td>
<td>10q24.32</td>
<td>104669968</td>
<td>t</td>
<td>c</td>
<td>0.407</td>
<td>1.25</td>
<td>3.33E-06</td>
<td>++++++++</td>
<td>intronic</td>
<td>CNNM2</td>
<td>0</td>
</tr>
<tr>
<td>rs10786733</td>
<td>10q24.32</td>
<td>104784937</td>
<td>a</td>
<td>g</td>
<td>0.406</td>
<td>1.26</td>
<td>3.37E-06</td>
<td>++++++++</td>
<td>intronic</td>
<td>CNNM2</td>
<td>0</td>
</tr>
<tr>
<td>rs10786719</td>
<td>10q24.32</td>
<td>10467982</td>
<td>a</td>
<td>g</td>
<td>0.588</td>
<td>0.80</td>
<td>3.56E-06</td>
<td>-------</td>
<td>intronic</td>
<td>AS3MT</td>
<td>0</td>
</tr>
<tr>
<td>rs6135309</td>
<td>20p12.1</td>
<td>14850522</td>
<td>t</td>
<td>c</td>
<td>0.71</td>
<td>0.78</td>
<td>3.69E-06</td>
<td>+++++++</td>
<td>intergenic</td>
<td>MACROD2</td>
<td>0</td>
</tr>
<tr>
<td>rs4532960</td>
<td>10q24.32</td>
<td>104657936</td>
<td>t</td>
<td>c</td>
<td>0.41</td>
<td>1.25</td>
<td>3.81E-06</td>
<td>++++++++</td>
<td>intergenic</td>
<td>AS3MT</td>
<td>5751</td>
</tr>
<tr>
<td>rs2391776</td>
<td>13q34</td>
<td>109041426</td>
<td>t</td>
<td>c</td>
<td>0.857</td>
<td>0.58</td>
<td>3.82E-06</td>
<td>---</td>
<td>intergenic</td>
<td>IRS2</td>
<td>162759</td>
</tr>
<tr>
<td>rs1591915</td>
<td>10q24.32</td>
<td>104636839</td>
<td>t</td>
<td>c</td>
<td>0.41</td>
<td>1.25</td>
<td>4.17E-06</td>
<td>++++++++</td>
<td>intronic</td>
<td>AS3MT</td>
<td>0</td>
</tr>
<tr>
<td>rs2275271</td>
<td>10q24.32</td>
<td>104804152</td>
<td>t</td>
<td>c</td>
<td>0.59</td>
<td>0.79</td>
<td>4.37E-06</td>
<td>-------</td>
<td>exonic</td>
<td>CNNM2</td>
<td>0</td>
</tr>
<tr>
<td>rs10883824</td>
<td>10q24.32</td>
<td>104802887</td>
<td>a</td>
<td>g</td>
<td>0.59</td>
<td>0.79</td>
<td>4.65E-06</td>
<td>-------</td>
<td>interonic</td>
<td>CNNM2</td>
<td>0</td>
</tr>
<tr>
<td>rs4917994</td>
<td>10q24.32</td>
<td>104801699</td>
<td>t</td>
<td>c</td>
<td>0.41</td>
<td>1.26</td>
<td>4.68E-06</td>
<td>++++++++</td>
<td>interonic</td>
<td>CNNM2</td>
<td>0</td>
</tr>
<tr>
<td>rs10883795</td>
<td>10q24.32</td>
<td>104644567</td>
<td>t</td>
<td>c</td>
<td>0.70</td>
<td>0.79</td>
<td>4.73E-06</td>
<td>-------</td>
<td>interonic</td>
<td>AS3MT</td>
<td>0</td>
</tr>
<tr>
<td>rs943035</td>
<td>10q24.32</td>
<td>104829142</td>
<td>t</td>
<td>c</td>
<td>0.60</td>
<td>0.80</td>
<td>4.96E-06</td>
<td>-------</td>
<td>downstream</td>
<td>CNNM2</td>
<td>0</td>
</tr>
<tr>
<td>rs3740387</td>
<td>10q24.32</td>
<td>104839458</td>
<td>a</td>
<td>g</td>
<td>0.40</td>
<td>1.25</td>
<td>5.10E-06</td>
<td>++++++++</td>
<td>exonic</td>
<td>NT5C2</td>
<td>0</td>
</tr>
<tr>
<td>rs1046778</td>
<td>10q24.32</td>
<td>104651474</td>
<td>t</td>
<td>c</td>
<td>0.70</td>
<td>0.79</td>
<td>5.19E-06</td>
<td>-------</td>
<td>UTR3</td>
<td>AS3MT</td>
<td>0</td>
</tr>
<tr>
<td>rs7897654</td>
<td>10q24.32</td>
<td>104652448</td>
<td>t</td>
<td>c</td>
<td>0.70</td>
<td>0.79</td>
<td>5.26E-06</td>
<td>-------</td>
<td>downstream</td>
<td>AS3MT</td>
<td>0</td>
</tr>
<tr>
<td>rs943036</td>
<td>10q24.32</td>
<td>104826037</td>
<td>t</td>
<td>c</td>
<td>0.60</td>
<td>0.80</td>
<td>5.29E-06</td>
<td>-------</td>
<td>interonic</td>
<td>CNNM2</td>
<td>0</td>
</tr>
<tr>
<td>rs10883826</td>
<td>10q24.32</td>
<td>104820809</td>
<td>a</td>
<td>g</td>
<td>0.6</td>
<td>0.79</td>
<td>5.34E-06</td>
<td>-------</td>
<td>interonic</td>
<td>CNNM2</td>
<td>0</td>
</tr>
<tr>
<td>rs1926034</td>
<td>10q24.32</td>
<td>104819092</td>
<td>a</td>
<td>g</td>
<td>0.38</td>
<td>1.27</td>
<td>5.85E-06</td>
<td>++++++++</td>
<td>interonic</td>
<td>CNNM2</td>
<td>0</td>
</tr>
<tr>
<td>rs7100709</td>
<td>10q24.32</td>
<td>104639719</td>
<td>a</td>
<td>c</td>
<td>0.70</td>
<td>0.79</td>
<td>5.86E-06</td>
<td>-------</td>
<td>interonic</td>
<td>AS3MT</td>
<td>0</td>
</tr>
<tr>
<td>rs2161277</td>
<td>5q12.1</td>
<td>62902018</td>
<td>t</td>
<td>c</td>
<td>0.12</td>
<td>1.49</td>
<td>6.11E-06</td>
<td>++++++++</td>
<td>intergenic</td>
<td>HTR1A</td>
<td>389613</td>
</tr>
<tr>
<td>rs1926030</td>
<td>10q24.32</td>
<td>104845646</td>
<td>t</td>
<td>c</td>
<td>0.60</td>
<td>0.80</td>
<td>6.58E-06</td>
<td>-------</td>
<td>intergenic</td>
<td>NT5C2</td>
<td>0</td>
</tr>
<tr>
<td>rs1553632</td>
<td>12q13.13</td>
<td>50210347</td>
<td>a</td>
<td>g</td>
<td>0.06</td>
<td>2.17</td>
<td>7.89E-06</td>
<td>++++++++</td>
<td>intergenic</td>
<td>SLC4A8</td>
<td>14533</td>
</tr>
<tr>
<td>rs7961653</td>
<td>12q13.13</td>
<td>50161967</td>
<td>t</td>
<td>c</td>
<td>0.94</td>
<td>0.58</td>
<td>8.17E-06</td>
<td>---→??→</td>
<td>intronic</td>
<td>SLC4A8</td>
<td>0</td>
</tr>
<tr>
<td>rs11249878</td>
<td>8p23.1</td>
<td>8587665</td>
<td>a</td>
<td>c</td>
<td>0.70</td>
<td>0.79</td>
<td>8.25E-06</td>
<td>--→-----</td>
<td>intergenic</td>
<td>CLDN23</td>
<td>9411</td>
</tr>
<tr>
<td>rs7913682</td>
<td>10q24.32</td>
<td>104674382</td>
<td>a</td>
<td>t</td>
<td>0.42</td>
<td>1.27</td>
<td>9.05E-06</td>
<td>++++++++?</td>
<td>intronic</td>
<td>CNNM2</td>
<td>0</td>
</tr>
<tr>
<td>rs746293</td>
<td>10q24.32</td>
<td>104887244</td>
<td>t</td>
<td>g</td>
<td>0.40</td>
<td>1.24</td>
<td>9.05E-06</td>
<td>++++++++</td>
<td>intronic</td>
<td>NT5C2</td>
<td>0</td>
</tr>
<tr>
<td>rs716316</td>
<td>20p12.1</td>
<td>14856741</td>
<td>t</td>
<td>c</td>
<td>0.36</td>
<td>1.24</td>
<td>9.74E-06</td>
<td>++++++++</td>
<td>intronic</td>
<td>MACROD2</td>
<td>0</td>
</tr>
</tbody>
</table>

Freq=frequency, bp=base pairs, Direction=direction of effect in each study, studies in alphabetic order: AGES-Reykjavik, ARIC, ASPS, CHS, FHS, PROSPER, RS II, RSIII, TASCOG, 3C-Dijon.
### Supplemental Table VI. GWAS meta-analysis result with adjustment for WMH load at baseline

(filtered MAF5%, RSQ 0.3)

<table>
<thead>
<tr>
<th>SNP</th>
<th>Locus</th>
<th>Position</th>
<th>Coded allele</th>
<th>Non-coded allele</th>
<th>Freq coded allele</th>
<th>Odds ratio</th>
<th>P value</th>
<th>Direction</th>
<th>Function</th>
<th>Closest Gene</th>
<th>Distance to Gene (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2391776</td>
<td>13q34</td>
<td>109041426</td>
<td>t</td>
<td>c</td>
<td>0.86</td>
<td>0.54</td>
<td>1.01E-06</td>
<td>--?------?</td>
<td>intergenic</td>
<td>IRS2</td>
<td>162759</td>
</tr>
<tr>
<td>rs7664442</td>
<td>4p15.31</td>
<td>22627802</td>
<td>a</td>
<td>g</td>
<td>0.92</td>
<td>0.65</td>
<td>2.26E-06</td>
<td>--------++</td>
<td>intergenic</td>
<td>GBA3</td>
<td>197509</td>
</tr>
<tr>
<td>rs10017649</td>
<td>4p15.31</td>
<td>22610312</td>
<td>a</td>
<td>g</td>
<td>0.08</td>
<td>1.53</td>
<td>2.32E-06</td>
<td>++++++++</td>
<td>intergenic</td>
<td>GBA3</td>
<td>180019</td>
</tr>
<tr>
<td>rs4761974</td>
<td>12q13.13</td>
<td>50166712</td>
<td>a</td>
<td>t</td>
<td>0.94</td>
<td>0.50</td>
<td>2.94E-06</td>
<td>---?----?</td>
<td>intronic</td>
<td>SLC4A8</td>
<td>0</td>
</tr>
<tr>
<td>rs2240255</td>
<td>16q23.1</td>
<td>73330299</td>
<td>t</td>
<td>c</td>
<td>0.25</td>
<td>1.34</td>
<td>5.79E-06</td>
<td>+??????</td>
<td>intronic</td>
<td>FA2H</td>
<td>0</td>
</tr>
<tr>
<td>rs11118844</td>
<td>1q41</td>
<td>219976911</td>
<td>a</td>
<td>g</td>
<td>0.86</td>
<td>1.49</td>
<td>6.95E-06</td>
<td>++++++++</td>
<td>UTR5</td>
<td>DUSP10</td>
<td>0</td>
</tr>
<tr>
<td>rs2161277</td>
<td>5q12.1</td>
<td>62902018</td>
<td>t</td>
<td>c</td>
<td>0.12</td>
<td>1.54</td>
<td>7.78E-06</td>
<td>+++???????</td>
<td>intergenic</td>
<td>HTR1A</td>
<td>389613</td>
</tr>
<tr>
<td>rs2174190</td>
<td>4p15.31</td>
<td>22561798</td>
<td>a</td>
<td>c</td>
<td>0.09</td>
<td>1.48</td>
<td>8.39E-06</td>
<td>++++++++</td>
<td>intergenic</td>
<td>GBA3</td>
<td>131505</td>
</tr>
</tbody>
</table>

Freq=frequency, bp=base pairs, Direction=direction of effect in each study, studies in alphabetic order: AGES-Reykjavik, ARIC, ASPS, CHS, FHS, PROSPER, RS II, RS III, TASCOG, 3C-Dijon.
Supplemental Figure I.

Quantile-quantile (QQ) plots showing the observed (y-axis) versus the expected (x-axis) $-\log_{10}(p$-values) in individual cohort analyses of WMH progression without adjustment for WMH burden at baseline (filtered MAF5%, RSQ 0.3). The red line shows the distribution under the null-hypothesis.

$\lambda$=genomic control inflation factor
Supplemental Figure II.

Quantile-quantile (QQ) plots showing the observed (y-axis) versus the expected (x-axis) −log10(p-values) in individual cohort analyses of WMH progression adjusted for WMH burden at baseline (filtered MAF5%, RSQ 0.3). The red line shows the distribution under the null-hypothesis. λ=genomic control inflation factor.
Supplemental Figure III.

Scatter plot showing the mean WML volume at baseline (in cm$^3$) on the x-axis and the mean annual WML volume increase (in cm$^3$) on the y-axis for each cohort. CHS is not included as the protocol of this cohort did not include volumetric WML measurement.
Supplemental References


22. 3C Study Group. Vascular factors and risk of dementia: Design of the three-city study and baseline characteristics of the study population. Neuroepidemiology. 2003;22:316-325
White Matter Lesion Progression: Genome-Wide Search for Genetic Influences

Stroke. 2015;46:3048-3057; originally published online October 8, 2015;
doi: 10.1161/STROKEAHA.115.009252

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/46/11/3048

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Stroke can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Stroke is online at:
http://stroke.ahajournals.org/subscriptions/