Supplementary Table 1 - Demographic and clinical characteristics for individuals in the UK Biobank VTE GWAS analysis

	VTE Cases	VTE Controls
N Individuals	14,222	372,102
Age ± SD, years	60.3 ± 7.1	57.3 ± 7.9
Male, n (%)	6,374 (44.8%)	172,438 (46.3%)
Former Smoker, n (%)	5,517 (38.8%)	130,965 (35.2%)
Current Smoker, n (%)	1,767 (12.4%)	37,878 (10.2%)
Hypertension, n (%)	6,441 (45.3%)	119,426 (32.1%)
Diabetes, n (%)	1,436 (10.1%)	16,519 (4.4%)
Hyperlipidemia, n (%)	3,778 (26.6%)	63,964 (17.2%)
Body-Mass Index ± SD, kg/m^2	29.0 ± 5.5	27.3 ± 4.7
Variants Included in Analysis	13,599,453	

Abbreviations: GWAS, Genome-wide Association Study; SD, Standard Deviation; VTE, Venous Thromboembolism

Supplementary Table 2 - Demographic and clinical characteristics for veterans in the MVP VTE GWAS analysis

	White		Black		Hispanic	
	VTE Cases	VTE Controls	VTE Cases	VTE Controls	VTE Cases	VTE Controls
N Veterans	8,929	181,337	2,261	49,400	654	21,214
Age ± SD, years	71.0 ± 11.3	68.0 ± 12.8	66.1 ± 11.5	61.5 ± 11.6	66.8 ± 13.0	60.5 ± 14.7
Male, n (%)	8,490 (95.0%)	168,161 (92.7%)	2,402 (91.6%)	42,722 (86.5%)	613 (93.7%)	19,258 (90.8%)
Current Smoker, n (%)	1,697 (19.0%)	32,814 (18.1%)	680 (25.9%)	13,505 (27.3%)	77 (11.8%)	3,399 (16.0%)
Former Smoker, n (%)	5,122 (57.4%)	98,706 (54.4%)	1,244 (47.5%)	21,217 (42.9%)	398 (60.9%)	10,317 (48.6%)
Diabetes, n (%)	3,742 (41.9%)	62,283 (34.3%)	1,387 (52.9%)	20,857 (42.2%)	342 (52.3%)	8,431 (39.7%)
Hyperlipidemia, n (%)	3,812 (42.7%)	80,417 (44.3%)	1,057 (40.3%)	19,515 (39.5%)	291 (44.5%)	8,118 (38.3%)
Body-Mass Index ± SD, kg/m^2	31.5 ± 6.7	30.3 ± 5.9	31.3 ± 7.0	30.5 ± 6.2	32.2 ± 7.0	30.8 ± 5.8
Variants Included in Analysis	19,972,400		31,960,759		28,192,968	

Abbreviations: GWAS, Genome-wide Association Study; MVP, Million Veteran Program; SD, Standard Deviation; VTE, Venous Thromboembolism

Supplementary Note

1. Supplementary Results

Of the 22 novel loci, 6 contained at least one gene implicated in the coagulation cascade or platelet function (**Supplementary Table 5)**. Three previously reported suggestive ($5.0x10^{-8} < P < 0.05$) VTE associations at the *GP6*[4], *STXBP5*[5], and *VWF*[5] loci were now observed at genome-wide significance. Across all 33 VTE loci (11 known and 22 novel), 31 were directionally consistent across whites, blacks, and Hispanics in MVP and 22 demonstrated at least nominal significance ($P < 0.05$) in blacks and 7 in Hispanics (**Supplementary Table 13**). 2 known and 4 novel VTE loci demonstrated moderate heterogeneity across the three ethnicities ($50\% < \text{heterogeneity } I^2 < 75\%$), but remained below our pre-specified heterogeneity threshold of 75%. In addition, we found no evidence of association for 3 African specific variants previously reported in an analysis of 393 African ancestry VTE cases that lacked independent replication[6] (**Supplementary Table 14**). In a conditional analysis using combined summary statistics from MVP Europeans and UK Biobank, we identified an additional 15 independent VTE variants across the 33 loci **(Supplementary Table 15)**.

Understanding the full spectrum of phenotypic consequences of a given variant may reveal the mechanism by which a variant or gene leads to disease. Termed a phenome-wide association study (PheWAS), this approach examines the association of a risk variant across a range of phenotypes[7,8]. Using a median of 63 distinct EHR-derived ICD-9/10 diagnosis codes per participant and available clinical laboratory data, we tested each of the 30 autosomal VTE lead risk variants across 1,249 disease phenotypes, symptoms, injuries, and 4 continuous cardiometabolic traits. We found that several of the VTE risk variants demonstrated a range of pleiotropy (**Supplementary Table 16)**. For example, rs2074492 near *HLA-C*, was associated with multiple autoimmune diseases including an increased risk for celiac disease, a disorder previously associated with a greater risk of developing VTE[9]. Interestingly, 4 of the VTE risk loci demonstrated known associations with LDL cholesterol[10-12] (*MYRF*, *HLA-C*, *ABO*, and *SLC44A2*), and 2 with HDL (high-density lipoprotein) cholesterol/triglycerides[10,11] (*MYRF*, *PEPD*). In total, we identified 142 statistically significant ($P < 1.1 \times 10^{-6}$) PheWAS associations across the 30 genetic variants. Results of a PheWAS of the PRS_{VTE} in MVP are also shown in **Supplementary Table 16**.

Next, we sought to better understand how DNA sequence variants might differ in their contribution to vascular disease risk in the arterial and venous territories. Analysis of shared heritability provides a mechanism to better understand the relationship of common variant risk across phenotypes[3,13]. Using linkage disequilibrium score regression[3], we examined the genetic correlation between VTE and i) coronary artery disease (CAD), ii) peripheral artery disease (PAD), and iii) large artery stroke (LAS). We used summary statistics from the European UK Biobank VTE analysis, data from a European MVP release 2.1 PAD analysis[14], summary data of 60,801 CAD cases and 123,504 controls from the CARDIoGRAMplusC4D consortium[15], and 6,688 LAS cases and 454,450 controls from the 2018 MEGASTROKE analysis[16]. We noted a stronger positive correlation between VTE and PAD ($r_g = 0.47$, $P = 2.0x10^{-15}$) than for VTE and CAD ($r_g = 0.27$, $P = 1.2x10^{-7}$) or VTE and LAS ($r_g = 0.35$, $P = 0.02$, **Supplementary**

Fig. 7), suggesting that common variant risk links thrombotic complications across venous and arterial beds, but more greatly with peripheral vasculature. In a sensitivity analysis, the correlation between VTE and myocardial infarction (MI) was similar in direction and magnitude as that for VTE-CAD (VTE-MI r_g = 0.29, P = 2.2x10^{-7}). Association results for the 30 autosomal genome-wide lead VTE risk variants for PAD, CAD, and LAS in the MVP, CARDIoGRAMplusC4D, and MEGASTROKE analyses, respectively, are shown in **Supplementary Table 17.**

We then examined whether the identified VTE risk variants were associated with changes in protein concentrations in circulating plasma and queried recently published protein quantitative trait loci (pQTL) data derived from the plasma samples of 3,301 participants of the INTERVAL study[2,17]. We observed 102 pQTL associations in human plasma at genome-wide significance (P < 5x10^{-8}, **Supplementary Table 18**) including 5 VTE lead variant-protein associations directly related to the coagulation cascade (**Supplementary Table 19)**. VTE risk alleles were associated with decreased concentration of tissue factor pathway inhibitor (TFPI), and increased concentrations of plasminogen activator-inhibitor 1 (PAI-1), Factor VIII (F8), Factor X (F10) and its active form, Factor Xa. In each case, the VTE risk allele was associated with changes in protein concentration resulting in a pro-coagulant effect on the coagulation cascade.

We then attempted to identify causal VTE variants through a fine-mapping analysis leveraging our multi-ethnic summary statistics and the MR-MEGA software[18]. After excluding chromosome X and the major histocompatibility complex because of the complex LD structures across the regions, we constructed credible sets for 29 VTE loci that in aggregate account for ≥99% of the posterior probability of driving the VTE association based on the UK Biobank and trans-ethnic MVP summary statistics. At 12 VTE signals, the credible set included 6 or fewer VTE associated variants (**Supplementary Table 20**). These credible sets included the known causal *F5* Leiden[19] and *F2* G20210A[20] variants, and also included 4 variants at the *ZFPM2* locus - all of which were genome-wide trans-pQTL associations with PAI-1 (**Supplementary Table 9**).

For our VTE PRS analysis, we provide MVP release 3.0 results stratified by sex in **Supplementary Table 21,** and results of the incident event analysis in WHI stratified by WHI sub-study as well as by hormone replacement therapy use are shown in **Supplementary Tables 22-23**.

Lastly, in MVP we performed a manual chart review of 50 VTE cases and 50 controls, which demonstrated that our phenotyping algorithm had a positive predictive value of 96% (95% CI = 85.1-99.3%), and negative predictive value of 100% (95% CI = 91.1-100%).

2. Supplementary Methods

Cohort Descriptions

The design of the Million Veteran Program (MVP) was previously described[21]. In brief, individuals aged 19 to >100 years have been recruited from more than 50 VA Medical Centers nationwide since 2011. Each veteran's electronic health record (EHR) data are

being integrated into the MVP biorepository, including inpatient International Classification of Diseases (ICD-9/10) diagnosis codes, Current Procedural Terminology (CPT) procedure codes, clinical laboratory measurements, and reports of diagnostic imaging modalities. MVP received ethical and study protocol approval by the VA Central Institutional Review Board and informed consent was obtained from all participants.

In UK Biobank, individuals aged 45 to 69 years old were recruited from across the United Kingdom for participation[22]. At enrollment, a trained healthcare provider ascertained participants' medical histories through verbal interview. In addition, participants' EHR including inpatient ICD-9/10 diagnosis codes and Office of Population and Censuses Surveys (OPCS-4) procedure codes, were integrated into UK Biobank. Informed consent was obtained for all participants, and UK Biobank received ethical approval from the Research Ethics Committee (reference number 11/NW/0382). Our study was approved by a local Institutional Review Board at Partners Healthcare (protocol 2013P001840).

We also examined incident VTE data from the Women's Health Initiative (WHI) randomized clinical trial of Hormone Therapy (HT) for our PRS analysis. The overall design of the WHI study has been described previously[23]. In brief, at the inception of the WHI study (1993-1998), 161,808 postmenopausal women between the ages of 50 and 79 years were eligible for inclusion in multiple clinical trials. Exclusion criteria related to the presence of medical conditions predisposing to shortened survival or safety concerns. The protocol and consent forms were approved at each site by institutional review committees and all participants provided written informed consent. The WHI-HT initially comprised 27,347 postmenopausal women who were randomized to receive either estrogen plus progestin or estrogen alone versus placebo until the trials were stopped early in July 2002 and March 2004, respectively. All WHI-HT participants subsequently continued to be followed without intervention until close-out. Of the various components of WHI, VTE was adjudicated by physician adjudicators for participants who were enrolled in the HT trials. The WHI-HT trial was approved by the local institutional review board at the Fred Hutchinson Cancer Research Center.

Quality Control Analysis

In MVP, we excluded: duplicate samples, samples with more heterozygosity than expected, an excess (>2.5%) of missing genotype calls, or discordance between genetically inferred sex and phenotypic gender. In addition, one individual from each pair of related individuals (as measured by the KING[24] software) were removed. Veterans were then divided into three mutually exclusive ethnic groups based on self-identified race/ethnicity and admixture analysis using the ADMIXTURE v1.3 software[25]: 1) non-Hispanic whites (self-identified as "non-Hispanic," "white," and > 80% genetic European ancestry), 2) non-Hispanic blacks (self-identified as "non-Hispanic," "black," and > 50% genetic African ancestry), and 3) Hispanics (self-identified only). In total, 312,571 white, black, and Hispanic MVP participants passed our sample-level quality control. Prior to imputation, variants that were poorly called or that deviated from their expected allele frequency based on reference data from the 1000 Genomes Project[26] were excluded. After pre-phasing using EAGLE[27] v2, genotypes from the 1000 Genomes Project[26] phase 3, version 5 reference panel were imputed into Million Veteran Program (MVP)

participants via Minimac3 software[28]. Ethnicity-specific principal component analysis was performed using the EIGENSOFT software[29].

Following imputation, variant level quality control was performed using the EasyQC R package[30] (www.R-project.org), and exclusion metrics included: ancestry specific Hardy-Weinberg equilibrium[31] $P < 1x10^{-20}$, posterior call probability < 0.9, imputation quality <0.3, minor allele frequency (MAF) < 0.003, call rate < 97.5% for common variants (MAF > 1%), and call rate < 99% for rare variants (MAF < 1%). Variants were also excluded if they deviated > 10% from their expected allele frequency based on reference data from the 1000 Genomes Project[26].

In UK Biobank, approximately 500,000 individuals were genotyped and subsequently imputed to the haplotype reference consortium (HRC) and UK10K reference panels. Details of these procedures are described elsewhere[32]. We performed genome-wide association testing for VTE in the UK Biobank using all variants in the v3 release with minor allele frequency greater than 0.3% and imputation quality INFO > 0.4. To avoid potential population stratification, only European-ancestry samples were included in the analysis. This subset was selected based on self-reported white ethnicity that was subsequently confirmed using genetic principal components analysis. Outliers within the self-reported white samples in the first 6 principal components of ancestry were detected and subsequently removed using the R package *aberrant*[33]. In addition, individuals with sex chromosome aneuploidy (neither XX or XY), discordant self-reported and genetic sex, or excessive heterozygosity or missingness, as defined centrally by the UK Biobank were removed. Finally, one individual from each pair of second-degree or closer relatives (kinship > 0.0884) was removed, selectively retaining VTE cases when possible.

VTE Phenotype and Manual Chart Review

Manual chart review was performed by two blinded trained clinician chart abstractors with a vascular surgeon reviewing discordant cases; the results of chart review for 50 cases and 50 controls otherwise representative of the overall cohort were used for determining the positive and negative predictive values of the phenotyping algorithm, which was standardized to the 50% prevalence of VTE in the validation set. Positive predictive value refers to the ratio of (true positives)/(true positives + false positives) and negative predictive value the ratio of (true negatives)/(true negatives + false negatives). In UK Biobank, individuals were defined as having VTE based on the definition by Klarin and colleagues as previously described[34]. All other individuals were defined as controls.

Discovery and Replication Association Analysis

In MVP, genotyped and imputed DNA sequence variants were tested for association with VTE through logistic regression adjusting for age, sex, and 5 principal components of ancestry assuming an additive model using the SNPTEST (mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html) statistical software program. In our discovery analysis, we performed association analyses separately for each ancestral group (whites, blacks, and Hispanics) and then meta-analyzed using an inverse variance-weighted fixed effects method implemented in the METAL software

program[35]. We excluded variants with a high amount of heterogeneity (I^2 statistic > 75%) across the three ancestries.

In UK Biobank, association testing was performed using a logistic regression model adjusted for age at baseline, sex, genotyping array, and the first 5 principal components of ancestry. All testing was performed in PLINK2 (https://www.cog-genomics.org/plink/2.0/).

We combined results across MVP and UK Biobank cohorts using inverse-variance weighted fixed effects meta-analysis and set a significance threshold of $P < 5 \times 10^{-8}$ (genome-wide significance). In addition, we also required a replication $P < 0.01$ in each of the MVP and UK Biobank analyses (e.g. MVP discovery and subsequent UK Biobank replication, and vice versa), with concordant direction of effect, to minimize false positive findings. Novel loci were defined as being greater than 500,000 base-pairs away from a known VTE genome-wide associated lead variant. Additionally, linkage disequilibrium information from the 1000 Genomes Project[26] was used to determine independent variants where a locus extended beyond 500,000 base-pairs. All logistic regression P values were two-sided.

Conditional Analysis

We used the COJO-GCTA software to perform an approximate, stepwise conditional analysis to identify independent variants within VTE-associated loci. We used summary statistics from the European specific meta-analysis of UK Biobank and MVP datasets (23,151 VTE cases, 553,439 controls) to conduct this analysis combined with an LD-matrix obtained from 20,000 unrelated European individuals randomly sampled from the UK Biobank release v3. We set a threshold $P < 5 \times 10^{-8}$ (genome-wide significance) to declare statistical significance.

PheWAS Disease Definitions, and Association Analysis

Understanding the full spectrum of phenotypic consequences of a given DNA sequence variant may shed light on the mechanism by which a variant/gene leads to disease. For 30 autosomal lead VTE risk variants and the PRS_{VTE} identified in our study, we performed a PheWAS of 1,249 distinct diseases, symptoms, and injuries in MVP leveraging the full catalog of EHR ICD-9/10 diagnosis codes in 227,817 white veterans using the R package PheWAS[36]. We additionally included 4 continuous cardiometabolic traits - LDL cholesterol, HDL cholesterol, triglycerides, and body mass index - given their possible links with VTE causality[34,37]. In total, 1,249 disease phenotypes and 4 continuous traits were available for analysis and we set a statistical threshold of $P < 1.2 \times 10^{-6}$ [0.05/(31 x (1,249 diseases + 4 continuous traits))]. Of 312,571 genotyped veterans passing quality control, we identified 23,172,451 distinct, prevalent ICD-9 diagnosis codes available for analysis. We focused on the largest ethnic group of 227,817 white participants, in which the mean age was 64.3 ± 13.1 years, and 93.3% (212,465) were male.

ICD-9 diagnosis codes were collapsed to clinical disease groups and corresponding controls using the groupings proposed by Denny et al[8]. Diseases were required to have a prevalence of > 0.13% (~300 cases) to be included in the PheWAS analysis. 30 autosomal lead VTE risk DNA sequence variants and the PRS_{VTE} were tested using logistic regression adjusting for age, sex, and five principal components under the assumption of additive effects using the PheWAS R package

(https://github.com/PheWAS/PheWAS) in R v3.2.0 (www.R-project.org). In total, 1,249 disease phenotypes and 4 continuous traits were available for analysis and we set a statistical threshold of $P < 1.1 \times 10^{-6}$ [0.05/(31 x (1,249 diseases + 4 continuous traits))]. For the lipid continuous traits (LDL cholesterol, HDL cholesterol, and triglycerides), maximal LDL cholesterol/triglycerides (after log transformation) and minimal HDL cholesterol were used after inverse normal transformation in MVP as previously described[38]. The body-mass index (BMI) phenotype was formulated in both UK Biobank and MVP and results were combined in an inverse-variance weighted fixed effects meta-analysis. In UK Biobank, BMI was calculated in 374,942 unrelated individuals from the measurement acquired at enrollment. In MVP, BMI was calculated in 218,382 participants from the mean height and mean weight over the 3 years prior to the enrollment date. Outliers were excluded if their mean measurement was < 17 or > 60 kg/m^2. In each case, the BMI phenotype was adjusted for age, age squared, and principal components of ancestry in a linear regression model. The resulting residuals were transformed to approximate normality using inverse normal scores separately by sex as previously described[39]. All logistic and linear regression P values were two-sided.

Shared Heritability within PAD, CAD, and LAS

To better understand the how common genetic variation influences risk for atherosclerosis in multiple vascular beds, we used linkage disequilibrium score regression[3] to calculate the genetic correlation between VTE-PAD, VTE-CAD/VTE-MI and VTE-LAS. Summary statistics from the European UK Biobank VTE GWAS, European MVP PAD GWAS[14], the CARDIoGRAMplusC4D CAD/MI GWAS[15] (predominantly European), and the transancestral LAS MEGASTROKE GWAS meta-analysis (>2/3 European)[16] were used for this analysis. Of note, we used the transancestral meta-analysis statistics from MEGASTROKE because the sample size of the European-ancestry only analysis lacked sufficient power for estimation of genetic correlation. We then queried association results for the 30 autosomal genome-wide lead VTE risk variants for PAD, CAD, and LAS in the MVP PAD[14], CARDIoGRAMplusC4D[15] CAD, and MEGASTROKE[16] LAS GWAS analyses, respectively.

VTE Variant-Plasma Protein Associations

To identify loci that might influence plasma protein concentrations potentially implicated in thromboembolism, we used published protein quantitative trait locus (pQTL) data generated from an aptamer-based multiplex protein assay to quantify 3,622 plasma proteins in 3,301 healthy participants from the INTERVAL study[2,17]. We queried the 30 autosomal VTE risk variants identified in our study for overlap with genome-wide significant (two-sided $P < 5.0 \times 10^{-8}$) variant-protein pairs.

Fine-Mapping of VTE Association Signals

For 29 non-MHC, autosomal, VTE association signals, we used the MR-MEGA[18] software and VTE summary statistics from the UK Biobank (European) and MVP (African, European, and Hispanic) analyses. We first defined a genomic region 1 mega-base on either side of the VTE lead variant restricting to variants with MAF > 1%. Under the assumption of one causal variant at a given locus, we then used multi-dimensional scaling of the Euclidean distance matrix to generate axes of genetic variation to each set

of association statistics between ancestry groups as implemented in MR-MEGA. For each GWAS signal we applied the "meta-regression model," including one axis of genetic variation as a covariate, to each variant passing quality control. From this model, we examined the VTE association for each variant and the heterogeneity in allelic effects that is correlated with ancestry. Subsequently, we derived a posterior probability (using the resultant Bayes' factor) of VTE association and constructed a 99.99% credible set of variants driving each GWAS signal.

Genetic Analysis of Incident VTE Events in WHI

After assessing the associated VTE risk for *F5* p.R506Q and *F2* G20210A carriers as well as the 5% of individuals with the highest PRS_{VTE} relative to the rest of the population in MVP, we sought replication of our findings using incident VTE data from the WHI. The design of the WHI-HT study has been described previously[23]. In brief, at the inception of the WHI study postmenopausal women between the ages of 50 and 79 years were eligible for inclusion in multiple clinical trials. Data used in this analysis included incident VTE events from participants belonging to one of three GWAS sub-studies: 1) the WHI Genomics and Randomized Trials Network (WHI-GARNET, 457 incident VTE events among 4,233 participants), 2) the WHI Memory Study (WHIMS, 180 incident VTE events among 5,637 participants), 3) the WHI Long Life Study (WHI-LLS, 53 incident VTE events among 1,105 participants).

The WHI-GARNET sub-study consisted of individuals selected as a nested case-control sample of coronary heart disease, stroke, venous thrombosis, and incident diabetes events from the parent WHI Hormone Therapy Trial. From 27,347 women who participated in the Hormone Therapy Trial, 4,894 were genotyped on the Illumina Omni-Quad as part of WHI GARNET and imputed using the 1000 Genomes reference panel[26] phase 3, version 5. VTE cases were identified that occurred during the active phase of the Hormone Trial and afterwards. Controls were participants in the Hormone Therapy Trial free of all 4 case conditions. Matching criteria for controls were age, race/ethnicity, hysterectomy status, and enrollment date. GARNET WHI participants were predominantly European (87%), and only European individuals were included in the analysis. In total, 457 VTE incident events were identified among 4,233 individuals after removing 21 observations due to missingness.

The WHIMS sub-study consisted of WHI Hormone Trial women of European ancestry from the following sources: 1) WHI Memory Study (WHIMS) participants[40] who were not in WHI-GARNET, 2) women from the WHI-HT at least 65 years old at enrollment who were neither in WHIMS nor GARNET, and 3) women from the WHI-HT younger than age 65 at enrollment who were neither in WHIMS nor GARNET. In total, 180 incident VTE events were identified among 5,637 individuals after removing 50 observations due to missingness. Participants were genotyped using the Illumina HumanOmniExpress platform and imputed using the 1000 Genomes reference panel[26] phase 3, version 5.

The WHI-LLS (GWAS) sub-study consisted of the phase III cohort of additional eligible women who were added to the LLS study after the decision was made to expand the study population in 2012. In total, 53 VTE incident events were identified among 1,105 individuals after removing 13 observations due to missingness. Participants were

genotyped using the Illumina HumanOmniExpress platform and imputed using the 1000 Genomes reference panel[26] phase 3, version 5.

Cox proportional hazards models were used to estimate hazard ratios (HR) and 95% confidence intervals for the associations of the *F5* p.R506Q and *F2* G20210A mutations with VTE adjusting for age, 10 principal components of ancestry, and hormone therapy intervention status during the active phase of the WHI-HT. We then tested the associated VTE risk for the 5% of individuals with the highest PRS_{VTE} relative to the rest of the population using Cox proportional hazards models adjusting for age, 10 principal components of ancestry, and hormone therapy intervention status during the active phase of the WHI-HT. Results from WHIMS, WHI-LLS, and WHI-GARNET were combined using an inverse-variance weighted fixed effects meta-analysis. Bonferroni-corrected 2-sided P values (P=0.016; 0.05/3) for 3 tests were used to declare statistical significance. Analyses were performed using the R software program (version 3.5.1; Vienna, Austria).

References

1. Bulik-Sullivan, B.K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* **47**, 291-5 (2015).
2. Sun, B.B. *et al.* Genomic atlas of the human plasma proteome. *Nature* **558**, 73-79 (2018).
3. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* **47**, 1236-41 (2015).
4. Bezemer, I.D. *et al.* Gene variants associated with deep vein thrombosis. *Jama* **299**, 1306-14 (2008).
5. Smith, N.L. *et al.* Genetic variation associated with plasma von Willebrand factor levels and the risk of incident venous thrombosis. *Blood* **117**, 6007-11 (2011).
6. Heit, J.A. *et al.* Identification of unique venous thromboembolism-susceptibility variants in African-Americans. *Thromb Haemost* **117**, 758-768 (2017).
7. Denny, J.C. *et al.* PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. *Bioinformatics* **26**, 1205-10 (2010).
8. Denny, J.C. *et al.* Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. *Nat Biotechnol* **31**, 1102-10 (2013).
9. Ungprasert, P., Wijarnpreecha, K. & Tanratana, P. Risk of venous thromboembolism in patients with celiac disease: A systematic review and meta-analysis. *J Gastroenterol Hepatol* **31**, 1240-5 (2016).
10. Teslovich, T.M. *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707-13 (2010).
11. Global Lipids Genetics Consortium *et al.* Discovery and refinement of loci associated with lipid levels. *Nat Genet* **45**, 1274-83 (2013).
12. Liu, D.J. *et al.* Exome-wide association study of plasma lipids in >300,000 individuals. *Nat Genet* (2017).
13. Anttila, V. *et al.* Analysis of shared heritability in common disorders of the brain. *Science* **360**(2018).
14. Klarin, D. *et al.* Genome-wide Association Study of Peripheral Artery Disease in the Million Veteran Program. *Nat Med* **In Press**(2019).
15. CARDIoGRAMplusC4D Consortium. A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* **47**, 1121-30 (2015).
16. Malik, R. *et al.* Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet* (2018).
17. Di Angelantonio, E. *et al.* Efficiency and safety of varying the frequency of whole blood donation (INTERVAL): a randomised trial of 45 000 donors. *Lancet* **390**, 2360-2371 (2017).
18. Magi, R. *et al.* Trans-ethnic meta-regression of genome-wide association studies accounting for ancestry increases power for discovery and improves fine-mapping resolution. *Hum Mol Genet* **26**, 3639-3650 (2017).
19. Bertina, R.M. *et al.* Mutation in blood coagulation factor V associated with resistance to activated protein C. *Nature* **369**, 64-7 (1994).

20. Poort, S.R., Rosendaal, F.R., Reitsma, P.H. & Bertina, R.M. A common genetic variation in the 3'-untranslated region of the prothrombin gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis. *Blood* **88**, 3698-703 (1996).
21. Gaziano, J.M. *et al.* Million Veteran Program: A mega-biobank to study genetic influences on health and disease. *J Clin Epidemiol* **70**, 214-23 (2016).
22. Collins, R. What makes UK Biobank special? *The Lancet* **379**, 1173-1174 (2012).
23. The Women's Health Initiative Study Group. Design of the Women's Health Initiative clinical trial and observational study. *Control Clin Trials* **19**, 61-109 (1998).
24. Manichaikul, A. *et al.* Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867-73 (2010).
25. Alexander, D.H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* **19**, 1655-64 (2009).
26. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* **526**, 68-74 (2015).
27. Loh, P.R., Palamara, P.F. & Price, A.L. Fast and accurate long-range phasing in a UK Biobank cohort. *Nat Genet* **48**, 811-6 (2016).
28. Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. & Abecasis, G.R. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet* **44**, 955-9 (2012).
29. Price, A.L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38**, 904-9 (2006).
30. Winkler, T.W. *et al.* Quality control and conduct of genome-wide association meta-analyses. *Nat Protoc* **9**, 1192-212 (2014).
31. Hyde, C.L. *et al.* Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nat Genet* **48**, 1031-6 (2016).
32. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203-209 (2018).
33. Bellenguez, C., Strange, A., Freeman, C., Donnelly, P. & Spencer, C.C. A robust clustering algorithm for identifying problematic samples in genome-wide association studies. *Bioinformatics* **28**, 134-5 (2012).
34. Klarin, D., Emdin, C.A., Natarajan, P., Conrad, M.F. & Kathiresan, S. Genetic Analysis of Venous Thromboembolism in UK Biobank Identifies the ZFPM2 Locus and Implicates Obesity as a Causal Risk Factor. *Circ Cardiovasc Genet* **10**(2017).
35. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-1 (2010).
36. Carroll, R.J., Bastarache, L. & Denny, J.C. R PheWAS: data analysis and plotting tools for phenome-wide association studies in the R environment. *Bioinformatics* **30**, 2375-6 (2014).
37. Glynn, R.J. *et al.* A randomized trial of rosuvastatin in the prevention of venous thromboembolism. *N Engl J Med* **360**, 1851-61 (2009).
38. Klarin, D. *et al.* Genetics of blood lipids among ~300,000 multi-ethnic participants of the Million Veteran Program. *Nature Genetics* (2018).

39. Locke, A.E. *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518**, 197-206 (2015).
40. Shumaker, S.A. *et al.* The Women's Health Initiative Memory Study (WHIMS): a trial of the effect of estrogen therapy in preventing and slowing the progression of dementia. *Control Clin Trials* **19**, 604-21 (1998).

VA Million Veteran Program

MVP Executive Committee

- Co-Chair: J. Michael Gaziano, M.D., M.P.H.
- Co-Chair: Rachel Ramoni, D.M.D., Sc.D.
- Jim Breeling, M.D. (ex-officio)
- Kyong-Mi Chang, M.D.
- Grant Huang, Ph.D.
- Sumitra Muralidhar, Ph.D.
- Christopher J. O'Donnell, M.D., M.P.H.
- Philip S. Tsao, Ph.D.

MVP Program Office

- Sumitra Muralidhar, Ph.D.
- Jennifer Moser, Ph.D.

MVP Recruitment/Enrollment

- Recruitment/Enrollment Director/Deputy Director, Boston – Stacey B. Whitbourne, Ph.D.; Jessica V. Brewer, M.P.H.
- MVP Coordinating Centers
    - Clinical Epidemiology Research Center (CERC), West Haven – John Concato, M.D., M.P.H.
    - Cooperative Studies Program Clinical Research Pharmacy Coordinating Center, Albuquerque - Stuart Warren, J.D., Pharm D.; Dean P. Argyres, M.S.
    - Genomics Coordinating Center, Palo Alto – Philip S. Tsao, Ph.D.
    - Massachusetts Veterans Epidemiology Research Information Center (MAVERIC), Boston - J. Michael Gaziano, M.D., M.P.H.
    - MVP Information Center, Canandaigua – Brady Stephens, M.S.
- Core Biorepository, Boston – Mary T. Brophy M.D., M.P.H.; Donald E. Humphries, Ph.D.
- MVP Informatics, Boston – Nhan Do, M.D.; Shahpoor Shayan
- Data Operations/Analytics, Boston – Xuan-Mai T. Nguyen, Ph.D.

MVP Science

- Genomics - Christopher J. O'Donnell, M.D., M.P.H.; Saiju Pyarajan Ph.D.; Philip S. Tsao, Ph.D.
- Phenomics - Kelly Cho, M.P.H, Ph.D.
- Data and Computational Sciences – Saiju Pyarajan, Ph.D.
- Statistical Genetics – Elizabeth Hauser, Ph.D.; Yan Sun, Ph.D.; Hongyu Zhao, Ph.D.

MVP Local Site Investigators

- Atlanta VA Medical Center (Peter Wilson)
  1670 Clairmont Rd, Decatur, GA 30033
- Bay Pines VA Healthcare System (Rachel McArdle)

10,000 Bay Pines Blvd Bay Pines FL 33744

- Birmingham VA Medical Center (Louis Dellitalia)
  700 S. 19th Street Birmingham AL 35233
- Cincinnati VA Medical Center (John Harley)
  3200 Vine Street, Cincinnati, OH 45220
- Clement J. Zablocki VA Medical Center (Jeffrey Whittle)
  5000 West National Avenue, Milwaukee, WI 53295
- Durham VA Medical Center (Jean Beckham)
  508 Fulton Street Durham, NC 27705
- Edith Nourse Rogers Memorial Veterans Hospital (John Wells)
  200 Springs Road, Bedford, MA 01730
- Edward Hines, Jr. VA Medical Center (Salvador Gutierrez)
  5000 South 5th Avenue, Hines, IL 60141
- Fayetteville VA Medical Center (Gretchen Gibson)
  1100 N College Ave, Fayetteville, AR 72703
- VA Health Care Upstate New York (Laurence Kaminsky)
  113 Holland Avenue Albany NY 12208
- New Mexico VA Health Care System (Gerardo Villareal)
  1501 San Pedro Drive, S.E.Albuquerque, NM 87108
- VA Boston Healthcare System (Scott Kinlay)
  150 S. Huntington Avenue, Boston, MA 02130
- VA Western New York Healthcare System (Junzhe Xu)
  3495 Bailey Avenue Buffalo, NY 14215-1199
- Ralph H. Johnson VA Medical Center (Mark Hamner)
  109 Bee Street, Mental Health Research, Charleston, SC 29401
- Wm. Jennings Bryan Dorn VA Medical Center (Kathlyn Sue Haddock)
  6439 Garners Ferry Road, Columbia, SC 29209
- VA North Texas Health Care System (Sujata Bhushan)
  4500 S. LANCASTER ROAD, DALLAS, TX 75216
- Hampton VA Medical Center (Pran Iruvanti)
  100 Emancipation Drive, Hampton, VA 23667
- Hunter Holmes McGuire VA Medical Center (Michael Godschalk)
  1201 Broad Rock Blvd., Richmond, VA 23249
- Iowa City VA Health Care System (Zuhair Ballas)
  601 Highway 6 West, Iowa City, IA 52246-2208

- Jack C. Montgomery VA Medical Center (Malcolm Buford)
  1011 Honor Heights Dr., Muskogee, OK 74401
- James A. Haley Veterans' Hospital (Stephen Mastorides)
  13000 Bruce B. Downs Blvd., Tampa, FL 33612
- Louisville VA Medical Center (Jon Klein)
  800 Zorn Avenue, Louisville, KY 40206
- Manchester VA Medical Center (Nora Ratcliffe)
  718 Smyth Road, Manchester, NH 03104
- Miami VA Health Care System (Hermes Florez)

1201 NW 16th Street, 11 GRC, Miami FL 33125

- Michael E. DeBakey VA Medical Center (Alan Swann)
  2002 Holcombe Blvd. Houston TX 77030
- Minneapolis VA Health Care System (Maureen Murdoch)
  One Veterans Drive Minneapolis MN 55417
- N. FL/S. GA Veterans Health System (Peruvemba Sriram)
  1601 SW Archer Road, Gainesville, FL 32608
- Northport VA Medical Center (Shing Shing Yeh)
  79 Middleville Road, Northport, NY 11768
- Overton Brooks VA Medical Center (Ronald Washburn)
  510 East Stoner Ave, Shreveport, LA 71101
- Philadelphia VA Medical Center (Darshana Jhala)
  3900 Woodland Avenue, Philadelphia, PA 19104
- Phoenix VA Health Care System (Samuel Aguayo)
  650 E. Indian School Road, Phoenix, AZ 85012
- Portland VA Medical Center (David Cohen)
  3710 SW U.S. Veterans Hospital Road, Portland, OR 97239
- Providence VA Medical Center (Satish Sharma)
  830 Chalkstone Avenue, Providence, RI 02908
- Richard Roudebush VA Medical Center (John Callaghan)
  1481 West 10th Street, Indianapolis, IN 46202
- Salem VA Medical Center (Kris Ann Oursler)
  1970 Roanoke Blvd.,Salem, VA 24153
- San Francisco VA Health Care System (Mary Whooley)
  4150 Clement Street, San Francisco, CA 94121
- South Texas Veterans Health Care System (Sunil Ahuja)
  7400 Merton Minter Boulevard, San Antonio, TX 78229
- Southeast Louisiana Veterans Health Care System (Amparo Gutierrez)
  2400 Canal Street, New Orleans, LA 70119
- Southern Arizona VA Health Care System (Ronald Schifman)
  3601 S 6th Ave, Tucson, AZ 85723
- Sioux Falls VA Health Care System (Jennifer Greco)
  2501 W 22nd St, Sioux Falls, SD 57105
- St. Louis VA Health Care System (Michael Rauchman)
  915 North Grand Blvd., St. Louis, MO 63106
- Syracuse VA Medical Center (Richard Servatius)
  800 Irving Avenue, Syracuse, NY 13210
- VA Eastern Kansas Health Care System (Mary Oehlert)
  4101 S 4th Street Trafficway, Leavenworth, KS 66048
- VA Greater Los Angeles Health Care System (Agnes Wallbom)
  11301 Wilshire Blvd Los Angeles, CA 90073
- VA Loma Linda Healthcare System (Ronald Fernando)
  11201 Benton Street, Loma Linda, CA 92357
- VA Long Beach Healthcare System (Timothy Morgan)
  5901 East 7th Street Long Beach CA 90822

- VA Maine Healthcare System (Todd Stapley)
  1 VA Center, Augusta, ME 04330
- VA New York Harbor Healthcare System (Scott Sherman)
  423 East 23rd Street New York, NY 10010
- VA Pacific Islands Health Care System (Gwenevere Anderson)
  459 Patterson Rd, Honolulu, HI 96819
- VA Palo Alto Health Care System (Philip Tsao)
  3801 Miranda Avenue Palo Alto, CA 94304-1290
- VA Pittsburgh Health Care System (Elif Sonel)
  University Drive, Pittsburgh, PA 15240
- VA Puget Sound Health Care System (Edward Boyko)
  1660 S. Columbian Way Seattle, WA 98108-1597
- VA Salt Lake City Health Care System (Laurence Meyer)
  500 Foothill Drive Salt Lake City, UT 84148
- VA San Diego Healthcare System (Samir Gupta)
  3350 La Jolla Village Drive, San Diego, CA 92161
- VA Southern Nevada Healthcare System (Joseph Fayad)
  6900 North Pecos Road, North Las Vegas, NV 89086
- VA Tennessee Valley Healthcare System (Adriana Hung)
  1310 24th Ave. South Nashville, TN 37212
- Washington DC VA Medical Center (Jack Lichy)
  50 Irving St, Washington, D. C. 20422
- W.G. (Bill) Hefner VA Medical Center (Robin Hurley)
  1601 Brenner Ave, Salisbury, NC 28144
- White River Junction VA Medical Center (Brooks Robey)
  163 Veterans Drive, White River Junction, VT 05009
- William S. Middleton Memorial Veterans Hospital (Robert Striker)
  2500 Overlook Terrace, Madison, WI 53705

International Network Against Venous Thrombosis (INVENT) Consortium
***Blood* 2019 Collaboration**

Sara Lindstrom, PhD,[1] Lu Wang, PhD,[2] Erin N. Smith, PhD,[3] William Gordon, MS,[4] Astrid van Hylckama Vlieg, PhD,[5] Mariza de Andrade, PhD,[6] Jennifer A. Brody, BA,[7] Jack W. Pattee, BA,[8] Jeffrey Haessler, MS,[9] Ben M. Brumpton, PhD, MPH,[10] Daniel I. Chasman, PhD,[11] Pierre Suchon, MD-PhD,[12] Ming-Huei Chen, PhD,[13] Constance Turman, MS,[14] Marine Germain,[15] Kerri L. Wiggins, MS, RD,[16] James MacDonald, MS,[17] Sigrid K. Braekkan, PhD,[18] Sebastian M. Armasu, MS,[19] Nathan Pankratz, PhD,[20] Rabecca D. Jackson, MD,[21] Jonas B. Nielsen, MD, PhD,[22] Franco Giulianini, PhD,[23] Marja K. Puurunen, MD, PhD,[24] Manal Ibrahim, MD,[25] Susan R. Heckbert, MD, PhD,[26] Theo K. Bammler, PhD,[27] Kelly A. Frazer, PhD,[28] Bryan M. McCauley, MS,[29] Kent Taylor, PhD,[30] James S. Pankow, PhD, MPH,[31] Alexander P. Reiner, MD, MPH,[32] Maiken E. Gabrielsen, PhD,[33] Jean-François Deleuze, PhD,[34] Chris J. O'Donnell, MD,[35] Jihye Kim, PhD, MPH,[36] Barbara McKnight, PhD,[37] Peter Kraft, PhD,[38] John-Bjarne Hansen, MD, PhD,[39] Frits R. Rosendaal, MD, PhD,[40] John A. Heit, MD,[41] Bruce M. Psaty, MD, PhD,[42] Weihong Tang, MD, PhD,[43] Charles Kooperberg, PhD,[44] Kristian Hveem, MD, PhD,[45] Paul M. Ridker, MD, MPH,[46] Pierre-Emmanuel Morange, MD-PhD,[47] Andrew D. Johnson, PhD,[48] Christopher Kabrhel, MD MPH,[49] David-AlexandreTrégouët, PhD,[50] Nicholas L. Smith, PhD[51]

Author Affiliations: [1]Department of Epidemiology, University of Washington, Seattle, Washington, USA; Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA. [2]Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, Washington, USA. [3]Department of Pediatrics and Rady Children's Hospital, University of California San Diego, La Jolla, USA; K.G. Jebsen Thrombosis Research and Expertise Center, Department of Clinical Medicine, UiT – The Arctic University of Norway, Tromsø, Norway. [4]Department of Epidemiology, University of Washington, Seattle, WA, USA. [5]Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands. [6]Department of Health Sciences Research, Mayo Clinic, Rochester, MN USA. [7]Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle WA USA. [8]Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, MN USA. [9]Division of Public Health, Fred Hutchinson Cancer Research Center, Seattle WA, USA. [10]K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Trondheim, Norway; MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK; Clinic of Thoracic and Occupational Medicine, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway. [11]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, USA; Harvard Medical School, Boston, USA. [12]Laboratory of Haematology, La Timone Hospital, Marseille, France; C2VN, Aix Marseille University, INSERM, INRA, C2VN, Marseille, France. [13]Population Sciences Branch, Division of Intramural Research, National Heart, Lung and Blood Institute, Bethesda, MD, USA; NHLBI and Boston University's The Framingham Heart Study, Framingham, MA, USA. [14]Program in Genetic Epidemiology and Statistical Genetics, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [15]INSERM UMR_S

1219, Bordeaux Population Health Research Center, University of Bordeaux, France. [16]Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle WA USA. [17]Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, Washington, USA. [18]K.G. Jebsen Thrombosis Research and Expertise Center, Department of Clinical Medicine, UiT – The Arctic University of Norway, Tromsø, Norway; Division of Internal Medicine, University Hospital of North Norway, Tromsø, Norway. [19]Health Sciences Research, Mayo Clinic, Rochester, MN USA. [20]Department of Laboratory Medicine and Pathology, School of Medicine, University of Minnesota, Minneapolis, MN, USA. [21]Division of Endocrinology, Diabetes and Metabolism, The Ohio State University, Columbus OH, USA. [22]Department of Internal Medicine, Division of Cardiology, University of Michigan Medical School, Ann Arbor, Michigan, USA. [23]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, USA. [24]NHLBI and Boston University's The Framingham Heart Study, Framingham, MA, USA. [25]Laboratory of Haematology, La Timone Hospital, Marseille, France.; C2VN, Aix Marseille University, INSERM, INRA, C2VN, Marseille, France. [26]Department of Epidemiology, University of Washington, Seattle, Washington, USA; Kaiser Permanente Washington Health Research Institute, Kaiser Permanente Washington, Seattle WA USA. [27]Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, Washington, USA. [28]Department of Pediatrics and Rady Children's Hospital, University of California San Diego, La Jolla, USA; K.G. Jebsen Thrombosis Research and Expertise Center, Department of Clinical Medicine, UiT – The Arctic University of Norway, Tromsø, Norway; Institute of Genomic Medicine, University of California San Diego, La Jolla, California, USA. [29]Health Sciences Research, Mayo Clinic, Rochester, MN USA. [30]Los Angeles Biomedical Research Institute and Department of Pediatrics, Harbor-UCLA Medical Center, Torrence CA 90502, USA. [31]Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, Minnesota, USA. [32]Department of Epidemiology, University of Washington, Seattle WA, United States; Division of Public Health, Fred Hutchinson Cancer Research Center, Seattle WA, United States. [33]K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Trondheim, Norway. [34]Centre National de Recherche en Génomique Humaine, Direction de la Recherche Fondamentale, CEA, 91057 Evry, France; CEPH, Fondation Jean Dausset, Paris, France. [35]Million Veteran's Program, Veteran's Administration, Boston, MA; Population Sciences Branch, Division of Intramural Research, National Heart, Lung and Blood Institute, Bethesda, MD, USA; NHLBI and Boston University's The Framingham Heart Study, Framingham, MA, USA. [36]Program in Genetic Epidemiology and Statistical Genetics, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [37]Department of Biostatistics, University of Washington, Seattle WA USA; Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA. [38]Program in Genetic Epidemiology and Statistical Genetics, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [39]K.G. Jebsen Thrombosis Research and Expertise Center, Department of Clinical Medicine, UiT – The Arctic University of Norway, Tromsø, Norway; Division of Internal Medicine, University Hospital of North Norway, Tromsø, Norway. [40]Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands. [41]Health Sciences Research, Mayo Clinic, Rochester,

MN USA. [42]Cardiovascular Health Research Unit, Departments of Medicine, Epidemiology, and Health Services, University of Washington, Seattle WA USA; Kaiser Permanente Washington Health Research Institute, Kaiser Permanente Washington, Seattle WA USA. [43]Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, Minnesota, USA. [44]Division of Public Health, Fred Hutchinson Cancer Research Center, Seattle WA, United States. [45]K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Trondheim, Norway. [46]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, USA; Harvard Medical School, Boston, USA. [47]C2VN, Aix Marseille Univ, INSERM, INRA, C2VN, Marseille, France; Laboratory of Haematology, La Timone Hospital, Marseille, France; CRB Assistance Publique - Hôpitaux de Marseille, HemoVasc (CRB AP-HM HemoVasc), Marseille, France. [48]Population Sciences Branch, Division of Intramural Research, National Heart, Lung and Blood Institute, Bethesda, MD, USA; NHLBI and Boston University's The Framingham Heart Study, Framingham, MA, USA. [49]Center for Vascular Emergencies, Department of Emergency Medicine, Massachusetts General Hospital; Channing Division of Network Medicine, Brigham and Women's Hospital; Harvard Medical School. [50]INSERM UMR_S 1219, Bordeaux Population Health Research Center, University of Bordeaux, France. [51]Department of Epidemiology, University of Washington, Seattle, Washington, USA; Kaiser Permanente Washington Health Research Institute, Kaiser Permanente Washington, Seattle WA USA; Seattle Epidemiologic Research and Information Center, Department of Veterans Affairs Office of Research and Development, Seattle WA USA. On behalf of the INVENT

MEGASTROKE CONSORTIUM

Rainer Malik [1], Ganesh Chauhan [2], Matthew Traylor [3], Muralidharan Sargurupremraj [4,5], Yukinori Okada [6,7,8], Aniket Mishra [4,5], Loes Rutten-Jacobs [3], Anne-Katrin Giese [9], Sander W van der Laan [10], Solveig Gretarsdottir [11], Christopher D Anderson [12,13,14,14], Michael Chong [15], Hieab HH Adams [16,17], Tetsuro Ago [18], Peter Almgren [19], Philippe Amouyel [20,21], Hakan Ay [22,13], Traci M Bartz [23], Oscar R Benavente [24], Steve Bevan [25], Giorgio B Boncoraglio [26], Robert D Brown, Jr. [27], Adam S Butterworth [28,29], Caty Carrera [30,31], Cara L Carty [32,33], Daniel I Chasman [34,35], Wei-Min Chen [36], John W Cole [37], Adolfo Correa [38], Ioana Cotlarciuc [39], Carlos Cruchaga [40,41], John Danesh [28,42,43,44], Paul IW de Bakker [45,46], Anita L DeStefano [47,48], Marcel den Hoed [49], Qing Duan [50], Stefan T Engelter [51,52], Guido J Falcone [53,54], Rebecca F Gottesman [55], Raji P Grewal [56], Vilmundur Gudnason [57,58], Stefan Gustafsson [59], Jeffrey Haessler [60], Tamara B Harris [61], Ahamad Hassan [62], Aki S Havulinna [63,64], Susan R Heckbert [65], Elizabeth G Holliday [66,67], George Howard [68], Fang-Chi Hsu [69], Hyacinth I Hyacinth [70], M Arfan Ikram [16], Erik Ingelsson [71,72], Marguerite R Irvin [73], Xueqiu Jian [74], Jordi Jiménez-Conde [75], Julie A Johnson [76,77], J Wouter Jukema [78], Masahiro Kanai [6,7,79], Keith L Keene [80,81], Brett M Kissela [82], Dawn O Kleindorfer [82], Charles Kooperberg [60], Michiaki Kubo [83], Leslie A Lange [84], Carl D Langefeld [85], Claudia Langenberg [86], Lenore J Launer [87], Jin-Moo Lee [88], Robin Lemmens [89,90], Didier Leys [91], Cathryn M Lewis [92,93], Wei-Yu Lin [28,94], Arne G Lindgren [95,96], Erik Lorentzen [97], Patrik K Magnusson [98], Jane Maguire [99], Ani Manichaikul [36], Patrick F McArdle [100], James F Meschia [101], Braxton D Mitchell [100,102], Thomas H Mosley [103,104], Michael A Nalls [105,106], Toshiharu Ninomiya [107], Martin J O'Donnell [15,108], Bruce M Psaty [109,110,111,112], Sara L Pulit [113,45], Kristiina Rannikmäe [114,115], Alexander P Reiner [65,116], Kathryn M Rexrode [117], Kenneth Rice [118], Stephen S Rich [36], Paul M Ridker [34,35], Natalia S Rost [9,13], Peter M Rothwell [119], Jerome I Rotter [120,121], Tatjana Rundek [122], Ralph L Sacco [122], Saori Sakaue [7,123], Michele M Sale [124], Veikko Salomaa [63], Bishwa R Sapkota [125], Reinhold Schmidt [126], Carsten O Schmidt [127], Ulf Schminke [128], Pankaj Sharma [39], Agnieszka Slowik [129], Cathie LM Sudlow [114,115], Christian Tanislav [130], Turgut Tatlisumak [131,132], Kent D Taylor [120,121], Vincent NS Thijs [133,134], Gudmar Thorleifsson [11], Unnur Thorsteinsdottir [11], Steffen Tiedt [1], Stella Trompet [135], Christophe Tzourio [5,136,137], Cornelia M van Duijn [138,139], Matthew Walters [140], Nicholas J Wareham [86], Sylvia Wassertheil-Smoller [141], James G Wilson [142], Kerri L Wiggins [109], Qiong Yang [47], Salim Yusuf [15], Najaf Amin [16], Hugo S Aparicio [185,48], Donna K Arnett [186], John Attia [187], Alexa S Beiser [47,48], Claudine Berr [188], Julie E Buring [34,35], Mariana Bustamante [189], Valeria Caso [190], Yu-Ching Cheng [191], Seung Hoan Choi [192,48], Ayesha Chowhan [185,48], Natalia Cullell [31], Jean-François Dartigues [193,194], Hossein Delavaran [95,96], Pilar Delgado [195], Marcus Dörr [196,197], Gunnar Engström [19], Ian Ford [198], Wander S Gurpreet [199], Anders Hamsten [200,201], Laura Heitsch [202], Atsushi Hozawa [203], Laura Ibanez [204], Andreea Ilinca [95,96], Martin Ingelsson [205], Motoki Iwasaki [206], Rebecca D Jackson [207], Katarina Jood [208], Pekka Jousilahti [63], Sara Kaffashian [4,5], Lalit Kalra [209], Masahiro Kamouchi [210], Takanari Kitazono [211], Olafur Kjartansson [212], Manja Kloss [213], Peter J Koudstaal [214], Jerzy Krupinski [215], Daniel L Labovitz [216], Cathy C Laurie [118], Christopher R Levi [217], Linxin Li [218], Lars Lind [219], Cecilia M Lindgren [220,221], Vasileios Lioutas [222,48], Yong Mei Liu [223], Oscar L Lopez [224], Hirata Makoto [225], Nicolas Martinez-Majander [172], Koichi Matsuda [225], Naoko Minegishi [203], Joan Montaner [226], Andrew P Morris [227,228], Elena Muiño [31], Martina Müller-Nurasyid [229,230,231], Bo Norrving [95,96], Soichi Ogishima [203], Eugenio A Parati [232], Leema Reddy Peddareddygari [56], Nancy L Pedersen [98,233], Joanna Pera [129], Markus Perola [63,234], Alessandro Pezzini [235], Silvana Pileggi [236], Raquel Rabionet [237], Iolanda Riba-Llena [30], Marta Ribasés [238], Jose R Romero [185,48], Jaume Roquer [239,240], Anthony G Rudd [241,242], Antti-Pekka Sarin [243,244], Ralhan Sarju [199], Chloe Sarnowski [47,48], Makoto Sasaki [245], Claudia L Satizabal [185,48], Mamoru Satoh [245], Naveed Sattar [246], Norie Sawada [206], Gerli Sibolt [172], Ásgeir Sigurdsson [247], Albert Smith [248], Kenji Sobue [245], Carolina Soriano-Tárraga [240], Tara Stanne [249], O Colin Stine [250], David J Stott [251], Konstantin Strauch [229,252], Takako Takai [203], Hideo Tanaka [253,254], Kozo Tanno [245], Alexander

Teumer [255], Liisa Tomppo [172], Nuria P Torres-Aguila [31], Emmanuel Touze [256,257], Shoichiro Tsugane [206], Andre G Uitterlinden [258], Einar M Valdimarsson [259], Sven J van der Lee [16], Henry Völzke [255], Kenji Wakai [253], David Weir [260], Stephen R Williams [261], Charles DA Wolfe [241,242], Quenna Wong [118], Huichun Xu [191], Taiki Yamaji [206], Dharambir K Sanghera [125,169,170], Olle Melander [19], Christina Jern [171], Daniel Strbian [172,173], Israel Fernandez-Cadenas [31,30], W T Longstreth, Jr [174,65], Arndt Rolfs [175], Jun Hata [107], Daniel Woo [82], Jonathan Rosand [12,13,14], Guillaume Pare [15], Jemma C Hopewell [176], Danish Saleheen [177], Kari Stefansson [11,178], Bradford B Worrall [179], Steven J Kittner [37], Sudha Seshadri [180,48], Myriam Fornage [74,181], Hugh S Markus [3], Joanna MM Howson [28], Yoichiro Kamatani [6,182], Stephanie Debette [4,5], Martin Dichgans [1,183,184]

1 Institute for Stroke and Dementia Research (ISD), University Hospital, LMU Munich, Munich, Germany
2 Centre for Brain Research, Indian Institute of Science, Bangalore, India
3 Stroke Research Group, Division of Clinical Neurosciences, University of Cambridge, UK
4 INSERM U1219 Bordeaux Population Health Research Center, Bordeaux, France
5 University of Bordeaux, Bordeaux, France
6 Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan
7 Department of Statistical Genetics, Osaka University Graduate School of Medicine, Osaka, Japan
8 Laboratory of Statistical Immunology, Immunology Frontier Research Center (WPI-IFReC), Osaka University, Suita, Japan.
9 Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA
10 Laboratory of Experimental Cardiology, Division of Heart and Lungs, University Medical Center Utrecht, University of Utrecht, Utrecht,Netherlands
11 deCODE genetics/AMGEN inc, Reykjavik, Iceland
12 Center for Genomic Medicine, Massachusetts General Hospital (MGH), Boston, MA, USA
13 J. Philip Kistler Stroke Research Center, Department of Neurology, MGH, Boston, MA, USA
14 Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA
15 Population Health Research Institute, McMaster University, Hamilton, Canada
16 Department of Epidemiology, Erasmus University Medical Center, Rotterdam, Netherlands
17 Department of Radiology and Nuclear Medicine, Erasmus University Medical Center, Rotterdam, Netherlands
18 Department of Medicine and Clinical Science, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan
19 Department of Clinical Sciences, Lund University, Malmö, Sweden
20 Univ. Lille, Inserm, Institut Pasteur de Lille, LabEx DISTALZ-UMR1167, Risk factors and molecular determinants of aging-related diseases, F-59000 Lille, France
21 Centre Hosp. Univ Lille, Epidemiology and Public Health Department, F-59000 Lille, France
22 AA Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA
23 Cardiovascular Health Research Unit, Departments of Biostatistics and Medicine, University of Washington, Seattle, WA, USA
24 Division of Neurology, Faculty of Medicine, Brain Research Center, University of British Columbia, Vancouver, Canada
25 School of Life Science, University of Lincoln, Lincoln, UK
26 Department of Cerebrovascular Diseases, Fondazione IRCCS Istituto Neurologico "Carlo Besta", Milano, Italy
27 Department of Neurology, Mayo Clinic Rochester, Rochester, MN, USA

28 MRC/BHF Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK
29 The National Institute for Health Research Blood and Transplant Research Unit in Donor Health and Genomics, University of Cambridge, UK
30 Neurovascular Research Laboratory, Vall d'Hebron Institut of Research, Neurology and Medicine Departments-Universitat Autònoma de Barcelona, Vall d'Hebrón Hospital, Barcelona, Spain
31 Stroke Pharmacogenomics and Genetics, Fundacio Docència i Recerca MutuaTerrassa, Terrassa, Spain
32 Children's Research Institute, Children's National Medical Center, Washington, DC, USA
33 Center for Translational Science, George Washington University, Washington, DC, USA
34 Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA, USA
35 Harvard Medical School, Boston, MA, USA
36 Center for Public Health Genomics, Department of Public Health Sciences, University of Virginia, Charlottesville, VA, USA
37 Department of Neurology, University of Maryland School of Medicine and Baltimore VAMC, Baltimore, MD, USA
38 Departments of Medicine, Pediatrics and Population Health Science, University of Mississippi Medical Center, Jackson, MS, USA
39 Institute of Cardiovascular Research, Royal Holloway University of London, UK & Ashford and St Peters Hospital, Surrey UK
40 Department of Psychiatry,The Hope Center Program on Protein Aggregation and Neurodegeneration (HPAN),Washington University, School of Medicine, St. Louis, MO, USA
41 Department of Developmental Biology, Washington University School of Medicine, St. Louis, MO, USA
42 NIHR Blood and Transplant Research Unit in Donor Health and Genomics, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK
43 Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK
44 British Heart Foundation, Cambridge Centre of Excellence, Department of Medicine, University of Cambridge, Cambridge, UK
45 Department of Medical Genetics, University Medical Center Utrecht, Utrecht, Netherlands
46 Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, Netherlands
47 Boston University School of Public Health, Boston, MA, USA
48 Framingham Heart Study, Framingham, MA, USA
49 Department of Immunology, Genetics and Pathology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden
50 Department of Genetics, University of North Carolina, Chapel Hill, NC, USA
51 Department of Neurology and Stroke Center, Basel University Hospital, Switzerland
52 Neurorehabilitation Unit, University and University Center for Medicine of Aging and Rehabilitation Basel, Felix Platter Hospital, Basel, Switzerland
53 Department of Neurology, Yale University School of Medicine, New Haven, CT, USA
54 Program in Medical and Population Genetics, The Broad Institute of Harvard and MIT, Cambridge, MA, USA
55 Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD, USA
56 Neuroscience Institute, SF Medical Center, Trenton, NJ, USA
57 Icelandic Heart Association Research Institute, Kopavogur, Iceland
58 University of Iceland, Faculty of Medicine, Reykjavik, Iceland
59 Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden
60 Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

61 Laboratory of Epidemiology and Population Science, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA
62 Department of Neurology, Leeds General Infirmary, Leeds Teaching Hospitals NHS Trust, Leeds, UK
63 National Institute for Health and Welfare, Helsinki, Finland
64 FIMM - Institute for Molecular Medicine Finland, Helsinki, Finland
65 Department of Epidemiology, University of Washington, Seattle, WA, USA
66 Public Health Stream, Hunter Medical Research Institute, New Lambton, Australia
67 Faculty of Health and Medicine, University of Newcastle, Newcastle, Australia
68 School of Public Health, University of Alabama at Birmingham, Birmingham, AL, USA
69 Department of Biostatistical Sciences, Wake Forest School of Medicine, Winston-Salem, NC, USA
70 Aflac Cancer and Blood Disorder Center, Department of Pediatrics, Emory University School of Medicine, Atlanta, GA, USA
71 Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, CA, USA
72 Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden
73 Epidemiology, School of Public Health, University of Alabama at Birmingham, USA
74 Brown Foundation Institute of Molecular Medicine, University of Texas Health Science Center at Houston, Houston, TX, USA
75 Neurovascular Research Group (NEUVAS), Neurology Department, Institut Hospital del Mar d'Investigació Mèdica, Universitat Autònoma de Barcelona, Barcelona, Spain
76 Department of Pharmacotherapy and Translational Research and Center for Pharmacogenomics, University of Florida, College of Pharmacy, Gainesville, FL, USA
77 Division of Cardiovascular Medicine, College of Medicine, University of Florida, Gainesville, FL, USA
78 Department of Cardiology, Leiden University Medical Center, Leiden, the Netherlands
79 Program in Bioinformatics and Integrative Genomics, Harvard Medical School, Boston, MA, USA
80 Department of Biology, East Carolina University, Greenville, NC, USA
81 Center for Health Disparities, East Carolina University, Greenville, NC, USA
82 University of Cincinnati College of Medicine, Cincinnati, OH, USA
83 RIKEN Center for Integrative Medical Sciences, Yokohama, Japan
84 Department of Medicine, University of Colorado Denver, Anschutz Medical Campus, Aurora, CO, USA
85 Center for Public Health Genomics and Department of Biostatistical Sciences, Wake Forest School of Medicine, Winston-Salem, NC, USA
86 MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Institute of Metabolic Science, Cambridge Biomedical Campus, Cambridge, UK
87 Intramural Research Program, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA
88 Department of Neurology, Radiology, and Biomedical Engineering, Washington University School of Medicine, St. Louis, MO, USA
89 KU Leuven – University of Leuven, Department of Neurosciences, Experimental Neurology, Leuven, Belgium
90 VIB Center for Brain & Disease Research, University Hospitals Leuven, Department of Neurology, Leuven, Belgium
91 Univ.-Lille, INSERM U 1171. CHU Lille. Lille, France
92 Department of Medical and Molecular Genetics, King's College London, London, UK
93 SGDP Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK
94 Northern Institute for Cancer Research, Paul O'Gorman Building, Newcastle University, Newcastle, UK

95 Department of Clinical Sciences Lund, Neurology, Lund University, Lund, Sweden
96 Department of Neurology and Rehabilitation Medicine, Skåne University Hospital, Lund, Sweden
97 Bioinformatics Core Facility, University of Gothenburg, Gothenburg, Sweden
98 Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
99 University of Technology Sydney, Faculty of Health, Ultimo, Australia
100 Department of Medicine, University of Maryland School of Medicine, MD, USA
101 Department of Neurology, Mayo Clinic, Jacksonville, FL, USA
102 Geriatrics Research and Education Clinical Center, Baltimore Veterans Administration Medical Center, Baltimore, MD, USA
103 Division of Geriatrics, School of Medicine, University of Mississippi Medical Center, Jackson, MS, USA
104 Memory Impairment and Neurodegenerative Dementia Center, University of Mississippi Medical Center, Jackson, MS, USA
105 Laboratory of Neurogenetics, National Institute on Aging, National institutes of Health, Bethesda, MD, USA
106 Data Tecnica International, Glen Echo MD, USA
107 Department of Epidemiology and Public Health, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan
108 Clinical Research Facility, Department of Medicine, NUI Galway, Galway, Ireland
109 Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA, USA
110 Department of Epidemiology, University of Washington, Seattle, WA
111 Department of Health Services, University of Washington, Seattle, WA, USA
112 Kaiser Permanente Washington Health Research Institute, Seattle, WA, USA
113 Brain Center Rudolf Magnus, Department of Neurology, University Medical Center Utrecht, Utrecht, The Netherlands
114 Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK
115 Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK
116 Fred Hutchinson Cancer Research Center, University of Washington, Seattle, WA, USA
117 Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA
118 Department of Biostatistics, University of Washington, Seattle, WA, USA
119 Nuffield Department of Clinical Neurosciences, University of Oxford, UK
120 Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA, USA
121 Division of Genomic Outcomes, Department of Pediatrics, Harbor-UCLA Medical Center, Torrance, CA, USA
122 Department of Neurology, Miller School of Medicine, University of Miami, Miami, FL, USA
123 Department of Allergy and Rheumatology, Graduate School of Medicine, the University of Tokyo, Tokyo, Japan
124 Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA
125 Department of Pediatrics, College of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA
126 Department of Neurology, Medical University of Graz, Graz, Austria
127 University Medicine Greifswald, Institute for Community Medicine, SHIP-KEF, Greifswald, Germany
128 University Medicine Greifswald, Department of Neurology, Greifswald, Germany
129 Department of Neurology, Jagiellonian University, Krakow, Poland
130 Department of Neurology, Justus Liebig University, Giessen, Germany
131 Department of Clinical Neurosciences/Neurology, Institute of Neuroscience and Physiology, Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden

132 Sahlgrenska University Hospital, Gothenburg, Sweden
133 Stroke Division, Florey Institute of Neuroscience and Mental Health, University of Melbourne, Heidelberg, Australia
134 Austin Health, Department of Neurology, Heidelberg, Australia
135 Department of Internal Medicine, Section Gerontology and Geriatrics, Leiden University Medical Center, Leiden, the Netherlands
136 INSERM U1219, Bordeaux, France
137 Department of Public Health, Bordeaux University Hospital, Bordeaux, France
138 Genetic Epidemiology Unit, Department of Epidemiology, Erasmus University Medical Center Rotterdam, Netherlands
139 Center for Medical Systems Biology, Leiden, Netherlands
140 School of Medicine, Dentistry and Nursing at the University of Glasgow, Glasgow, UK
141 Department of Epidemiology and Population Health, Albert Einstein College of Medicine, NY, USA
142 Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, MS, USA
143 A full list of members and affiliations appears in the Supplementary Note
144 Department of Human Genetics, McGill University, Montreal, Canada
145 Department of Pathophysiology, Institute of Biomedicine and Translation Medicine, University of Tartu, Tartu, Estonia
146 Department of Cardiac Surgery, Tartu University Hospital, Tartu, Estonia
147 Clinical Gene Networks AB,Stockholm, Sweden
148 Department of Genetics and Genomic Sciences, The Icahn Institute for Genomics and Multiscale Biology Icahn School of Medicine at Mount Sinai, New York, NY , USA
149 Department of Pathophysiology, Institute of Biomedicine and Translation Medicine, University of Tartu, Biomeedikum, Tartu, Estonia
150 Integrated Cardio Metabolic Centre, Department of Medicine, Karolinska Institutet, Karolinska Universitetssjukhuset, Huddinge, Sweden.
151 Clinical Gene Networks AB, Stockholm, Sweden
152 Sorbonne Universités, UPMC Univ. Paris 06, INSERM, UMR_S 1166, Team Genomics & Pathophysiology of Cardiovascular Diseases, Paris, France
153 ICAN Institute for Cardiometabolism and Nutrition, Paris, France
154 Department of Biomedical Engineering, University of Virginia, Charlottesville, VA, USA
155 Group Health Research Institute, Group Health Cooperative, Seattle, WA, USA
156 Seattle Epidemiologic Research and Information Center, VA Office of Research and Development, Seattle, WA, USA
157 Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA, USA
158 Department of Medical Research, Bærum Hospital, Vestre Viken Hospital Trust, Gjettum, Norway
159 Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, Singapore
160 National Heart and Lung Institute, Imperial College London, London, UK
161 Department of Gene Diagnostics and Therapeutics, Research Institute, National Center for Global Health and Medicine, Tokyo, Japan
162 Department of Epidemiology, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA, USA
163 Department of Cardiology,University Medical Center Groningen, University of Groningen, Netherlands
164 MRC-PHE Centre for Environment and Health, School of Public Health, Department of Epidemiology and Biostatistics, Imperial College London, London, UK
165 Department of Epidemiology and Biostatistics, Imperial College London, London, UK
166 Department of Cardiology, Ealing Hospital NHS Trust, Southall, UK
167 National Heart, Lung and Blood Research Institute, Division of Intramural Research, Population Sciences Branch, Framingham, MA, USA

168 A full list of members and affiliations appears at the end of the manuscript
169 Department of Phamaceutical Sciences, Collge of Pharmacy, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA
170 Oklahoma Center for Neuroscience, Oklahoma City, OK, USA
171 Department of Pathology and Genetics, Institute of Biomedicine, The Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden
172 Department of Neurology, Helsinki University Hospital, Helsinki, Finland
173 Clinical Neurosciences, Neurology, University of Helsinki, Helsinki, Finland
174 Department of Neurology, University of Washington, Seattle, WA, USA
175 Albrecht Kossel Institute, University Clinic of Rostock, Rostock, Germany
176 Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK
177 Department of Genetics, Perelman School of Medicine, University of Pennsylvania, PA, USA
178 Faculty of Medicine, University of Iceland, Reykjavik, Iceland
179 Departments of Neurology and Public Health Sciences, University of Virginia School of Medicine, Charlottesville, VA, USA
180 Department of Neurology, Boston University School of Medicine, Boston, MA, USA
181 Human Genetics Center, University of Texas Health Science Center at Houston, Houston, TX, USA
182 Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan
183 Munich Cluster for Systems Neurology (SyNergy), Munich, Germany
184 German Center for Neurodegenerative Diseases (DZNE), Munich, Germany
185 Boston University School of Medicine, Boston, MA, USA
186 University of Kentucky College of Public Health, Lexington, KY, USA
187 University of Newcastle and Hunter Medical Research Institute, New Lambton, Australia
188 Univ. Montpellier, Inserm, U1061, Montpellier, France
189 Centre for Research in Environmental Epidemiology, Barcelona, Spain
190 Department of Neurology, Università degli Studi di Perugia, Umbria, Italy
191 Department of Medicine, University of Maryland School of Medicine, Baltimore, MD, USA
192 Broad Institute, Cambridge, MA, USA
193 Univ. Bordeaux, Inserm, Bordeaux Population Health Research Center, UMR 1219, Bordeaux, France
194 Bordeaux University Hospital, Department of Neurology, Memory Clinic, Bordeaux, France
195 Neurovascular Research Laboratory. Vall d'Hebron Institut of Research, Neurology and Medicine Departments-Universitat Autònoma de Barcelona. Vall d'Hebrón Hospital, Barcelona, Spain
196 University Medicine Greifswald, Department of Internal Medicine B, Greifswald, Germany
197 DZHK, Greifswald, Germany
198 Robertson Center for Biostatistics, University of Glasgow, Glasgow, UK
199 Hero DMC Heart Institute, Dayanand Medical College & Hospital, Ludhiana, India
200 Atherosclerosis Research Unit, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden
201 Karolinska Institutet, Stockholm, Sweden
202 Division of Emergency Medicine, and Department of Neurology, Washington University School of Medicine, St. Louis, MO, USA
203 Tohoku Medical Megabank Organization, Sendai, Japan
204 Department of Psychiatry, Washington University School of Medicine, St. Louis, MO, USA

205 Department of Public Health and Caring Sciences / Geriatrics, Uppsala University, Uppsala, Sweden
206 Epidemiology and Prevention Group, Center for Public Health Sciences, National Cancer Center, Tokyo, Japan
207 Department of Internal Medicine and the Center for Clinical and Translational Science, The Ohio State University, Columbus, OH, USA
208 Institute of Neuroscience and Physiology, the Sahlgrenska Academy at University of Gothenburg, Goteborg, Sweden
209 Department of Basic and Clinical Neurosciences, King's College London, London, UK
210 Department of Health Care Administration and Management, Graduate School of Medical Sciences, Kyushu University, Japan
211 Department of Medicine and Clinical Science, Graduate School of Medical Sciences, Kyushu University, Japan
212 Landspitali National University Hospital, Departments of Neurology & Radiology, Reykjavik, Iceland
213 Department of Neurology, Heidelberg University Hospital, Germany
214 Department of Neurology, Erasmus University Medical Center
215 Hospital Universitari Mutua Terrassa, Terrassa (Barcelona), Spain
216 Albert Einstein College of Medicine, Montefiore Medical Center, New York, NY, USA
217 John Hunter Hospital, Hunter Medical Research Institute and University of Newcastle, Newcastle, NSW, Australia
218 Centre for Prevention of Stroke and Dementia, Nuffield Department of Clinical Neurosciences, University of Oxford, UK
219 Department of Medical Sciences, Uppsala University, Uppsala, Sweden
220 Genetic and Genomic Epidemiology Unit, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK
221 The Wellcome Trust Centre for Human Genetics, Oxford, UK
222 Beth Israel Deaconess Medical Center, Boston, MA, USA
223 Wake Forest School of Medicine, Wake Forest, NC, USA
224 Department of Neurology, University of Pittsburgh, Pittsburgh, PA, USA
225 BioBank Japan, Laboratory of Clinical Sequencing, Department of Computational biology and medical Sciences, Graduate school of Frontier Sciences, The University of Tokyo, Tokyo, Japan
226 Neurovascular Research Laboratory, Vall d'Hebron Institut of Research, Neurology and Medicine Departments-Universitat Autònoma de Barcelona. Vall d'Hebrón Hospital, Barcelona, Spain
227 Department of Biostatistics, University of Liverpool, Liverpool, UK
228 Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK
229 Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany
230 Department of Medicine I, Ludwig-Maximilians-Universität, Munich, Germany
231 DZHK (German Centre for Cardiovascular Research), partner site Munich Heart Alliance, Munich, Germany
232 Department of Cerebrovascular Diseases, Fondazione IRCCS Istituto Neurologico "Carlo Besta", Milano, Italy
233 Karolinska Institutet, MEB, Stockholm, Sweden
234 University of Tartu, Estonian Genome Center, Tartu, Estonia, Tartu, Estonia
235 Department of Clinical and Experimental Sciences, Neurology Clinic, University of Brescia, Italy
236 Translational Genomics Unit, Department of Oncology, IRCCS Istituto di Ricerche Farmacologiche Mario Negri, Milano, Italy
237 Department of Genetics, Microbiology and Statistics, University of Barcelona, Barcelona, Spain

238 Psychiatric Genetics Unit, Group of Psychiatry, Mental Health and Addictions, Vall d'Hebron Research Institute (VHIR), Universitat Autònoma de Barcelona, Biomedical Network Research Centre on Mental Health (CIBERSAM), Barcelona, Spain
239 Department of Neurology, IMIM-Hospital del Mar, and Universitat Autònoma de Barcelona, Spain
240 IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain
241 National Institute for Health Research Comprehensive Biomedical Research Centre, Guy's & St. Thomas' NHS Foundation Trust and King's College London, London, UK
242 Division of Health and Social Care Research, King's College London, London, UK
243 FIMM-Institute for Molecular Medicine Finland, Helsinki, Finland
244 THL-National Institute for Health and Welfare, Helsinki, Finland
245 Iwate Tohoku Medical Megabank Organization, Iwate Medical University, Iwate, Japan
246 BHF Glasgow Cardiovascular Research Centre, Faculty of Medicine, Glasgow, UK
247 deCODE Genetics/Amgen, Inc., Reykjavik, Iceland
248 Icelandic Heart Association, Reykjavik, Iceland
249 Institute of Biomedicine, the Sahlgrenska Academy at University of Gothenburg, Goteborg, Sweden
250 Department of Epidemiology, University of Maryland School of Medicine, Baltimore, MD, USA
251 Institute of Cardiovascular and Medical Sciences, Faculty of Medicine, University of Glasgow, Glasgow, UK
252 Chair of Genetic Epidemiology, IBE, Faculty of Medicine, LMU Munich, Germany
253 Division of Epidemiology and Prevention, Aichi Cancer Center Research Institute, Nagoya, Japan
254 Department of Epidemiology, Nagoya University Graduate School of Medicine, Nagoya, Japan
255 University Medicine Greifswald, Institute for Community Medicine, SHIP-KEF, Greifswald, Germany
256 Department of Neurology, Caen University Hospital, Caen, France
257 University of Caen Normandy, Caen, France
258 Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, Netherlands
259 Landspitali University Hospital, Reykjavik, Iceland
260 Survey Research Center, University of Michigan, Ann Arbor, MI, USA
261 University of Virginia Department of Neurology, Charlottesville, VA, USA