



# Genetic Architecture of Abdominal Aortic Aneurysm in the Million Veteran Program

**BACKGROUND:** Abdominal aortic aneurysm (AAA) is an important cause of cardiovascular mortality; however, its genetic determinants remain incompletely defined. In total, 10 previously identified risk loci explain a small fraction of AAA heritability.

**METHODS:** We performed a genome-wide association study in the Million Veteran Program testing  $\approx 18$  million DNA sequence variants with AAA (7642 cases and 172 172 controls) in veterans of European ancestry with independent replication in up to 4972 cases and 99 858 controls. We then used mendelian randomization to examine the causal effects of blood pressure on AAA. We examined the association of AAA risk variants with aneurysms in the lower extremity, cerebral, and iliac arterial beds, and derived a genome-wide polygenic risk score (PRS) to identify a subset of the population at greater risk for disease.

**RESULTS:** Through a genome-wide association study, we identified 14 novel loci, bringing the total number of known significant AAA loci to 24. In our mendelian randomization analysis, we demonstrate that a genetic increase of 10 mm Hg in diastolic blood pressure (odds ratio, 1.43 [95% CI, 1.24–1.66];  $P=1.6\times 10^{-6}$ ), as opposed to systolic blood pressure (odds ratio, 1.06 [95% CI, 0.97–1.15];  $P=0.2$ ), likely has a causal relationship with AAA development. We observed that 19 of 24 AAA risk variants associate with aneurysms in at least 1 other vascular territory. A 29-variant PRS was strongly associated with AAA (odds ratio<sub>PRS</sub>, 1.26 [95% CI, 1.18–1.36];  $P_{PRS}=2.7\times 10^{-11}$  per SD increase in PRS), independent of family history and smoking risk factors (odds ratio<sub>PRS+family history+smoking</sub>, 1.24 [95% CI, 1.14–1.35];  $P_{PRS}=1.27\times 10^{-6}$ ). Using this PRS, we identified a subset of the population with AAA prevalence greater than that observed in screening trials informing current guidelines.

**CONCLUSIONS:** We identify novel AAA genetic associations with therapeutic implications and identify a subset of the population at significantly increased genetic risk of AAA independent of family history. Our data suggest that extending current screening guidelines to include testing to identify those with high polygenic AAA risk, once the cost of genotyping becomes comparable with that of screening ultrasound, would significantly increase the yield of current screening at reasonable cost.

Derek Klarin, MD

:

Scott M. Damrauer<sup>ID</sup>, MD\*

Philip S. Tsao<sup>ID</sup>, PhD\*

On behalf of the Veterans Affairs Million Veteran Program

\*Drs Damrauer and Tsao jointly supervised this work.

†A complete list of the investigators in the Veterans Affairs Million Veteran Program is provided in the [Data Supplement](#).

The full author list is available on page 1643.

**Key Words:** aneurysm ■ aortic diseases ■ genome-wide association study ■ humans

Sources of Funding, see page 1644

© 2020 The Authors. *Circulation* is published on behalf of the American Heart Association, Inc., by Wolters Kluwer Health, Inc. This is an open access article under the terms of the [Creative Commons Attribution License](#), which permits use, distribution, and reproduction in any medium, provided that the original work is properly cited.

<https://www.ahajournals.org/journal/circ>

## Clinical Perspective

### What Is New?

- A genome-wide association study of abdominal aortic aneurysm (AAA) further defines the role of common genetic variation in disease susceptibility.
- Mendelian randomization analysis supports a causal role for smoking liability and diastolic blood pressure in the pathogenesis of AAA.
- A polygenic risk score based on 29 genome-wide AAA risk variants strongly associates with AAA independent of family history and 6 additional clinical risk factors.

### What Are the Clinical Implications?

- These results highlight novel AAA genetic associations with therapeutic implications, including *LPA* and *PCSK9*.
- Diastolic blood pressure, as opposed to systolic blood pressure, is likely of greater significance in the pathogenesis of AAA.
- Polygenic risk score analysis can identify a subset of the population at significantly increased genetic risk of AAA independent of family history; extending current screening guidelines to include testing for those with high polygenic AAA risk deserves further investigation, particularly once genotyping becomes standard of care in healthcare systems and can be performed at nominal cost.

**A**bdominal aortic aneurysm (AAA) is a complex disease affected by both environmental<sup>1</sup> and genetic factors,<sup>2</sup> and heritability has been estimated to be as high as 70%.<sup>3</sup> Genome-wide association studies (GWAS) have revealed only 10 loci reaching genome-wide significance,<sup>4–7</sup> leaving a significant portion of AAA heritability unknown. Previously published discovery efforts have been limited by small sample sizes, with the largest previous GWAS analyzing ≈5000 AAA cases.<sup>7</sup>

Large-scale biobanks combining genetic data with electronic health record (EHR)-derived phenotypes have continued to evolve over the past decade.<sup>8,9</sup> The Million Veteran Program (MVP) was established in 2011 to study how genes affect health in the Veterans Health Administration. Recent efforts have demonstrated that a Veterans Health Administration-based biobank provides the ability to aid genetic discovery for understudied vascular diseases with a higher prevalence among US veterans.<sup>10,11</sup> Leveraging the MVP resource, we sought to perform a genetic discovery analysis for AAA; explore the spectrum of phenotypic consequences associated with AAA risk variants; examine the genetic relationship between smoking and blood pressure and AAA; test the effects of AAA risk variants on aneurysms

in other territories; and develop and evaluate a genome-wide polygenic risk score (PRS) for AAA to identify a subset of the population at higher risk for disease.

## METHODS

The full summary-level association data from the MVP AAA discovery analysis from this article are available through dbGaP (the database of Genotypes and Phenotypes; accession code: phs001672.v2.p1).

### Study Populations

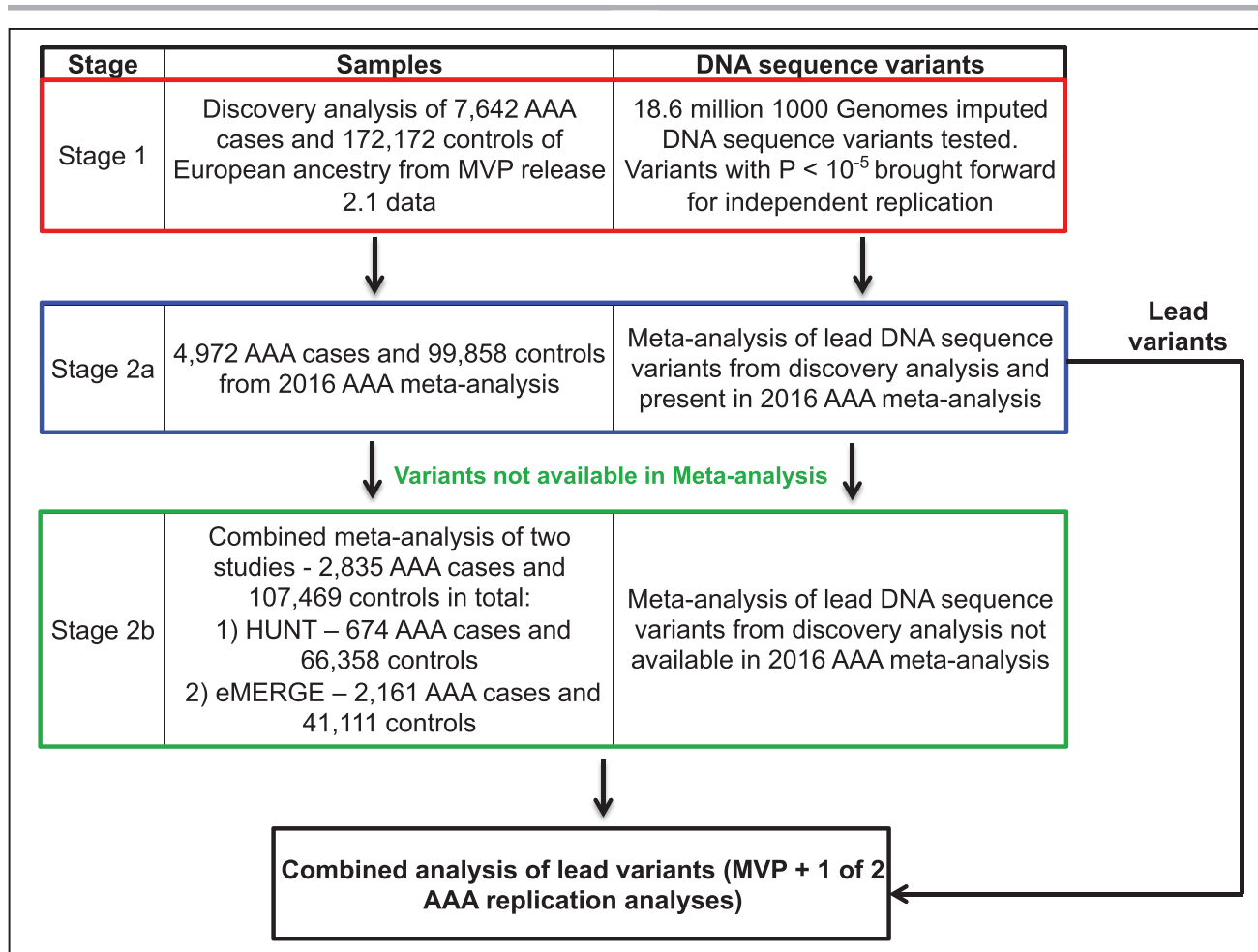
We conducted a discovery genetic association analysis using DNA samples and phenotypic data from the MVP (Figure 1 and Figure 1 in the Data Supplement). In the MVP, individuals ages 18 to >100 years have been recruited from 63 Veterans Affairs medical centers across the United States. After performing quality control, we identified 227 817 European participants for AAA discovery analysis from MVP release 2.1. After applying our phenotypic algorithm, we evaluated 7642 AAA cases and 172 172 controls free from clinical evidence of disease. For variants with suggestive associations ( $P < 10^{-5}$ ), we sought replication of our findings with data from either of 2 independent datasets. We first attempted replication of lead AAA variants using summary statistics from the 2016 AAA meta-analysis consisting of 4972 AAA cases and 99 858 controls<sup>7</sup> (stage 2a).

If association statistics for promising DNA sequence variants from stage 1 were not available for replication, we sought replication of these variants from a combined meta-analysis of 2 studies: HUNT (the Trøndelag Health Study; 674 AAA cases and 66 358 controls) and the eMERGE Network (electronic Medical Records and Genomics; 2161 AAA cases and 41 111 controls). Network datasets consisted of up to 2835 AAA cases and 107 469 controls in total (stage 2b). HUNT is a population-based health survey conducted in the county of Nord-Trøndelag, Norway, since 1984. Individuals were included at 3 different time points over ≈20 years (HUNT1 [1984 to 1986], HUNT2 [1995 to 1997], and HUNT3 [2006 to 2008]). The eMERGE Network is a collaboration of genetic biobanks with phenotypic data linked to EHRs. The study design and sample collection of eMERGE Network cohorts has been described previously.<sup>12</sup> Deidentified samples linked to EHRs were supplied from multiple investigators from different institutions with approval across the consortium for analysis. Additional details of genetic data and quality control are available in the Data Supplement.

The MVP received ethical and study protocol approval by the Veterans Affairs Central Institutional Review Board and informed consent was obtained for all participants. Each additional study received approval from its local institutional review board.

### AAA Phenotype Definitions

From the participants passing quality control in MVP, individuals were defined as having AAA or being a disease-free control using a previously adjudicated<sup>13</sup> definition initially proposed by Denny et al.<sup>14</sup> AAA cases were defined as the presence of 2 instances of any of the following International Classification of Diseases (ICD)-9 or ICD-10 codes in a



**Figure 1. Abdominal aortic aneurysm (AAA) genome-wide association study design.**

DNA sequence variants with suggestive association ( $P < 10^{-5}$ ) in discovery (stage 1) were brought forward for independent replication and tested using summary statistics from the 2016 AAA meta-analysis consisting of 4972 AAA cases and 99 858 controls<sup>7</sup> (stage 2a) or from a combined meta-analysis of HUNT (the Trøndelag Health Study) and the eMERGE Network (electronic Medical Records and Genomics) datasets consisting of a total of 2835 AAA cases and 107 469 controls (stage 2b) if variants were not available in stage 2a. MVP indicates Million Veteran Program.

participant's EHR: 441.3, 441.4, I71.3, or I71.4. Controls were defined as possessing no occurrences of the aforementioned ICD codes, as well as no occurrences of the ICD-9 codes 440 through 448 or ICD-10 codes I71 through I75, I77 through I79, or K55. This identified 7642 AAA cases and 172 172 controls of European ancestry for genetic discovery, and an independent set of 1656 AAA cases and 44 908 controls of European ancestry and 718 AAA cases and 46 380 controls of African ancestry for our PRS analysis. In HUNT, all patients diagnosed with AAA from 1995 through 2014 in Nord-Trøndelag were identified from hospital and outpatient registries from the 3 local hospitals: St Olav's Hospital, Levanger Hospital, and Namsos Hospital. All diagnoses were verified manually using patient chart data. The unique personal identification number was used to link data on AAA diagnoses and death with exposure data from HUNT. This identified 674 AAA cases and 66 358 controls for downstream analysis. In the eMERGE Network, BioMe, and the Penn Medicine Biobank, AAA was defined in a manner identical to the definition proposed by Denny et al.<sup>14</sup> In the Mayo Clinic Vascular Disease Biorepository, which consists of patients referred for noninvasive vascular evaluation

including abdominal ultrasound, AAA cases were defined as having an infrarenal abdominal aortic diameter  $\geq 3$  cm or a history of open or endovascular AAA repair. Controls were not known to have AAA and had no ICD-9 diagnosis codes for AAA.

### Phenome-Wide Association Study of AAA Risk Variants

Of the genotyped veterans passing quality control, we identified 23 172 451 distinct, prevalent ICD diagnosis codes available for analysis. We focused on the entire group of 227 817 White participants passing quality control, in which the mean age was  $64.3 \pm 13.1$  years, and 93.3% (212 465) were male. ICD-9 or ICD-10 diagnosis codes were collapsed to clinical disease groups and corresponding controls using the groupings proposed by Denny et al.<sup>14</sup> and tested using logistic regression. Diseases were required to have a prevalence of  $>0.13\%$  ( $\approx 300$  cases) to be included in the Phenome-Wide Association Study (PheWAS) analysis. We tested a total of 1249 distinct diseases, symptoms, and injuries in MVP using the R package PheWAS.<sup>15</sup>

## Shared Heritability With AAA Risk Factors

To better understand how common genetic variation influences risk for AAA and its risk factors, we used linkage disequilibrium score regression<sup>16</sup> to calculate the genetic correlation between AAA and 10 known risk factors for disease and aneurysm rupture. MVP association summary statistics for AAA were uploaded to LD Hub (<http://ldsc.broadinstitute.org/ldhub/>) and genetic correlation analyses performed between AAA, 4 smoking categories (current smoker, ever smoker, cigarettes/day, and chronic obstructive pulmonary disease), and 2 blood pressure traits (systolic blood pressure [SBP] and diastolic blood pressure [DBP]) from UK Biobank and previously published GLGC plasma lipid (low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, and triglycerides)<sup>17</sup> and coronary artery disease statistics from the CARDIoGRAMplusC4D (Coronary Artery Disease Genome-Wide Replication and Meta-Analysis Plus Coronary Artery Disease Genetics) GWAS.<sup>18</sup>

## Smoking and Blood Pressure Mendelian Randomization Analyses

Mendelian randomization (MR) analyses for 3 smoking phenotypes (smoking initiation, smoking heaviness, and smoking cessation) and 2 blood pressure phenotypes (SBP and DBP) with AAA development were performed. Genetic instruments were selected as DNA sequence variants that associated with the exposure at genome-wide significance ( $P < 5 \times 10^{-8}$ ) with an  $R^2 < 0.001$ . All clumping was performed using the TwoSampleMR package of R.<sup>19</sup> Genetic instruments for the smoking initiation, smoking heaviness, and smoking cessation exposures were derived from publicly available GSCAN (GWAS & Sequencing Consortium of Alcohol and Nicotine Use) data from up to 1 232 091 participants<sup>20</sup>; instruments of the SBP and DBP traits were derived from the International Consortium of Blood Pressure and UK Biobank Discovery GWAS<sup>21</sup> analysis in up to 757 601 individuals. Inverse variance-weighted MR was used for the primary analysis, with weighted median<sup>22</sup> MR performed as sensitivity analysis, allowing for up to 50% of the weight of each instrument to be drawn from invalid instruments while controlling type I error. Diagnostic leave-one-out, single-variant, funnel-plot, and MR-Egger<sup>23</sup> analyses were performed to evaluate for evidence of heterogeneity and horizontal pleiotropy. In addition, we performed the MR-PRESSO test (Mendelian Randomization Pleiotropy Residual Sum and Outlier)<sup>24</sup> to identify evidence of horizontal pleiotropy, which consists of 3 parts: (1) the global test for horizontal pleiotropy; (2) the outlier corrected causal estimate, which corrects for the detected horizontal pleiotropy; and (3) the distortion test, which tests whether the causal estimate is significantly different after outlier adjustment.

## Aneurysm Associations in Other Vascular Territories

We sought to better understand how DNA sequence variants might differ in contribution to risk for aneurysmal disease in the iliac, lower extremity, and cerebral vascular beds. We first

examined the strength of the relationship between AAA and aneurysms in other vascular beds. Next, for lead AAA risk variants identified in our primary analysis, we tested their effect on iliac, lower extremity, and cerebral aneurysms in European MVP participants using the definition proposed by Denny et al.<sup>14</sup> We performed a sensitivity analysis to examine the association with isolated aneurysms in these territories by excluding individuals with a diagnosis of AAA from the analysis.

## AAA PRS Generation

A weighted PRS represents an individual's risk of a given disease conferred by the sum of the effects of many common DNA sequence variants. A weight is assigned to each genetic variant based on its strength of association with disease risk ( $\beta$ ). Individuals are then additively scored in a weighted fashion based on the number of risk alleles they carry for each variant in the PRS.

To generate our scores, we used summary statistics from the MVP release 2.1 European ancestry GWAS analysis (7642 AAA cases, 172 172 controls) and a linkage disequilibrium panel of 20 000 randomly selected European samples from UK Biobank. To increase the number of independent variants included in our score, we performed a pruning and thresholding analysis using the linkage disequilibrium-driven clumping procedure in PLINK version 1.90b (clump). In brief, this algorithm formed "clumps" around variants with AAA association (in this case,  $P < 0.01$ ,  $P < 1 \times 10^{-4}$ ,  $P < 1 \times 10^{-6}$ ) and with an  $R^2 > 0.0001$  based on the linkage disequilibrium reference. From our initial set of summary statistics, the algorithm selected only 1 associated variant from each clump below our prespecified  $P$  value threshold. The final output from this procedure generated 3 scores of 29, 301, and 3699 independent ( $R^2 < 0.0001$ ) AAA-associated variants, representing the strongest disease-associated variant for each linkage disequilibrium-based clump across the genome.

## AAA PRS Analysis

We first tested the associated AAA risk for each of the PRS<sub>AAA</sub> using ascertained AAA case-control cohort data from the Mayo Clinic Vascular Disease Biorepository consisting of 1022 AAA cases and 7750 controls. Additional adjustments in the association were made to account for the presence of smoking and a family history of AAA. Once the best performing PRS<sub>AAA</sub> was identified, we validated our results using AAA data from 3 additional population-based datasets: (1) prevalent AAA data from MVP release 3.0 using 1656 AAA cases and 44 908 controls of European ancestry and 718 AAA cases and 46 380 controls of African ancestry independent from the individuals in the MVP discovery GWAS and specifically designated for PRS analysis; (2) a set of 195 AAA cases and 9348 controls from the BioMe Biobank Program; and (3) a set of 396 AAA cases and 9835 controls from the Penn Medicine Biobank. We tested the association of the continuous PRS<sub>AAA</sub> with prevalent AAA in each cohort and identified the prevalence of AAA in those individuals with the highest 5% of polygenic risk. We calculated the prevalence of AAA in each PRS quintile and in those individuals with the highest 5% of polygenic risk among men aged 50 or older, because individuals in their 6th decade of life or older are significantly

more likely to have AAA.<sup>25</sup> Last, we tested the association of our continuous PRS<sub>AAA</sub> while including 6 traditional AAA risk factors in the association model: smoking, hypertension, low-density lipoprotein cholesterol with statin adjustment, high-density lipoprotein cholesterol, triglycerides, and coronary artery disease as a marker of atherosclerosis burden.

## Statistical Analysis

In our primary discovery analysis, genotyped and imputed DNA sequence variants in individuals of European ancestry were tested for association with AAA using logistic regression adjusting for age, sex, and 5 principal components of ancestry assuming an additive model using the SNPTESTv2.5.3 statistical software program.

For variants with suggestive AAA associations ( $P < 10^{-5}$ ), we sought replication of our findings from either an in silico query from the 2016 AAA meta-analysis<sup>7</sup> (stage 2a) or the HUNT/eMERGE Network datasets (stage 2b). In HUNT, the GWAS was performed using a logistic mixed model implemented in SAIGE (Scalable and Accurate Implementation of Generalized Mixed Model).<sup>26</sup> Covariates in the model were sex, history of cardiovascular event as defined from HUNT questionnaires, history of hypertension, history of diabetes mellitus, myocardial infarction, angina pectoris, and smoking status (ever/never). In the eMERGE Network, association testing was performed using the PLATO software<sup>27</sup> adjusted for year of birth, sex, 6 principal components, and sample collection site. Analyses were restricted to eMERGE Network locations with adult participants and those contributing cases to the analysis. Discovery and replication results were combined using an inverse variance-weighted fixed-effects method as implemented in the METAL software program.<sup>28</sup> We did not attempt replication of any variant in either stage and required  $P < 0.05$  with concordant direction of effect for successful replication. We then combined statistical evidence across MVP and the replication studies and set a significance threshold of  $P < 5 \times 10^{-8}$  (genome-wide significance). Novel loci were defined as being  $> 500\,000$  base pairs away from a known AAA genome-wide associated lead variant. Linkage disequilibrium information from the 1000 Genomes Project<sup>29</sup> was used to determine independent variants when a locus extended beyond 500 000 base pairs. All logistic regression values of  $P$  were 2-sided.

In the PheWAS analysis, DNA sequence variants were tested using logistic regression adjusting for age, sex, and 5 principal components. Variants were declared to be significantly associated with the disease if they met  $P < 1.7 \times 10^{-6}$  ( $P < 0.05 / [1249 \text{ diseases} \times 24 \text{ variants}]$ ). In our genetic correlation analyses, traits were declared to be significant if they demonstrated a correlation  $P < 0.005$  ( $P < 0.05 / 10$  AAA risk factors).

In our MR analyses, a random-effects inverse variance-weighted method was used as the main analysis, with sensitivity analyses performed for the statistically significant associations as described previously. We set a 2-sided  $P < 0.01$  (0.05/5 traits) for statistical significance.

In addition,  $\chi^2$  tests of independence were used to examine whether the occurrence of AAA with aneurysms in other vascular territories was more frequent than would be expected by chance. In the association analysis of AAA risk variants with

aneurysms in other vascular territories, we performed logistic regression adjusting for age, sex, and 5 principal components of ancestry assuming an additive model using the R statistical software program<sup>30</sup> (R version 3.3). In our sensitivity analysis, risk variants were tested using the same logistic model with individuals with AAA excluded. Given the known prior association with AAA, we used a nominally significant  $P < 0.05$  to declare significance. All  $P$  values were 2-sided.

In our PRS analysis, logistic regression models were used to estimate odds ratios (ORs) and 95% CIs for the associations of each of the continuous PRSs (1 SD unit), first adjusting only for age, sex, and 5 principal components. Family history and smoking status were then included in the model as appropriate. We then performed logistic regression adjusting for age, sex, 5 principal components, statins, and 6 clinical risk factors for AAA in the association model. We calculated prevalence of AAA for the 5% of individuals with the highest PRS<sub>AAA</sub> relative to the rest of the population and generated CIs using R (version 3.3). All  $P$  values were 2-sided.

## RESULTS

The MVP discovery analysis comprised 7642 patients with AAA and 172 172 controls of European ancestry. Their baseline characteristics are presented in Table 1. Participants with AAA were more likely to be older, male, and prescribed statin therapy; to have a history of smoking; and to have type 2 diabetes mellitus. Through genotype imputation, we obtained 18.6 million DNA sequence variants for analysis. After discovery, a total of 549 variants at 16 loci met a genome-wide significance threshold (Figures II and III in the Data Supplement). We replicated all 10 previously described genome-wide AAA loci with association  $P < 10^{-5}$  (Table 2 and Table I in the Data Supplement). The *9p21* variant rs4007642 was the top association result (49.5% frequency for the T allele; OR, 1.21 [95% CI, 1.17–1.25];  $P = 6.9 \times 10^{-29}$ ).

A total of 70 lead variants at candidate novel loci demonstrated association  $P < 10^{-5}$  in the MVP discovery analysis and were taken forward for replication. Of those, 28 were available for independent testing in the 2016 AAA meta-analysis, and the remaining were tested in the combined HUNT and eMERGE Network datasets. After independent replication, 14 novel loci exceeded genome-wide significance ( $P < 5 \times 10^{-8}$ ; Table 2 and Table II in the Data Supplement). The remaining variants either did not meet the prespecified  $P < 0.05$  with consistent direction of effect required for independent replication or did not surpass the genome-wide significance threshold after meta-analysis (Table III in the Data Supplement).

Understanding the full spectrum of phenotypic consequences of a given DNA sequence variant can help identify the mechanism by which a variant or gene leads to disease. Using a median of 65 distinct ICD-9 or ICD-10 EHR-derived diagnosis codes per participant, we tested each of the 24 AAA lead risk variants

**Table 1. Demographic and Clinical Characteristics of Veterans in the Million Veteran Program Abdominal Aortic Aneurysm Genome-Wide Association Studies Analysis**

Characteristics	Abdominal Aortic Aneurysm	
	No. of Cases	No. of Controls
Veterans	7642	172 172
Age, y	76.2±8.4	67.1±13.8
Male	7586 (99.3)	159 065 (92.4)
Statin therapy prescription	4102 (53.7)	63 120 (36.7)
Diabetes mellitus	2986 (39.1)	49 581 (28.8)
Maximum low-density lipoprotein cholesterol, mg/dL	143.1±38.9	138.3±38.1
Former smoker	4022 (52.6)	77 322 (44.9)
Current smoker	2781 (36.4)	38 483 (22.4)
Variants passing quality control	18 689 289	

Values are mean ± SD or n (%).

(14 newly identified in our study, 10 known) across 1249 disease phenotypes. In total, we identified 232 statistically significant ( $P < 1.7 \times 10^{-6}$ ) PheWAS associations across the 24 genetic variants. Thirteen of the loci replicated known associations with atherosclerosis in at least 1 arterial territory (peripheral, coronary, or cerebral). We also found that several of the DNA sequence variants associated with a range of known risk factors for AAA (Table IV in the Data Supplement). For example, we found 7 PheWAS associations with hyperlipidemia and hypertriglyceridemia. These loci have previously associated with either low-density lipoprotein cholesterol (*LDLR*, *PCSK9*, *SORT1*, *LPA*, *APOE*)<sup>17</sup> or triglycerides (*APOA5*, *TRIB1*),<sup>17</sup> both of which have been causally implicated in AAA development.<sup>31</sup> The AAA risk allele for rs7255 near the lipid droplet-associated hydrogenase (*LDAH*) gene was associated with a decreased risk of hypercholesterolemia, suggesting that a mechanism other than increased serum cholesterol likely links the locus to aneurysmal degeneration. In *CHRNA3* (encoding cholinergic receptor nicotinic  $\alpha$ -3), rs55958997 demonstrated associations with chronic obstructive pulmonary disease and ventilator dependence. This DNA sequence variant is strongly correlated ( $R^2=0.83$ ) with variants shown to predict nicotine dependence and a reduced likelihood for cigarette smoking cessation<sup>32</sup> and appears to confer AAA risk entirely through its effect on smokers (Figure IV in the Data Supplement).

We next sought to more broadly understand the relationship of traditional AAA risk factors with disease at the genome-wide level. Analysis of shared heritability provides a mechanism to better understand the relationship of common variant risk across phenotypes.<sup>16,33</sup> Using linkage disequilibrium score regression,<sup>16</sup> we examined the genetic correlation between AAA and 10 known risk factors for disease and aneurysm rupture (Figure V in the Data Supplement). We found that 9

of the 10 risk factor traits were significantly correlated with AAA after Bonferroni correction ( $P < 0.005$ ), and all risk factors showed a consistent direction of effects between the epidemiologically posited association and the genetic correlation. Current smoking and coronary artery disease (a surrogate for atherosclerosis) demonstrated the strongest positive correlation (Table V in the Data Supplement). These results suggest a common genetic architecture for AAA and its individual risk factors.

In observational studies, lipids, smoking, and hypertension are modifiable, independent risk factors for AAA. We have previously examined the causal relationship between lipids and AAA,<sup>31</sup> and here we performed MR analyses for 3 smoking phenotypes—smoking initiation, smoking heaviness, and smoking cessation—with AAA development. Genetic instruments of 84, 20, and 8 variants for the smoking initiation, smoking heaviness, and smoking cessation exposures, respectively, were derived from publicly available GSCAN data<sup>20</sup> (Table VI in the Data Supplement). Using the inverse variance-weighted method, we observed that a genetically increased risk of ever being a regular smoker (OR 2.71 per 1 log-odds unit increase in risk of smoking initiation [95% CI, 1.83–4.01];  $P=6.2 \times 10^{-7}$ ) and a genetic increase in smoking heaviness (OR 2.53 per 1 U increase in daily smoking or  $\approx 5$  cigarettes/d [95% CI, 1.78–3.61];  $P=3.3 \times 10^{-7}$ ) were both associated with increased risk of AAA (Figure 2A). Our genetic instrument for likelihood of smoking cessation was found to be protective (OR 0.21 per 1 log-odds unit increase in likelihood of smoking cessation [95% CI, 0.05–0.89];  $P=0.03$ ), although this result did not survive Bonferroni correction. Our results remained robust to multiple sensitivity analyses, including the weighted median<sup>22</sup> and leave-one-out analyses, as well as MR-PRESSO<sup>24</sup> and MR-Egger<sup>23</sup> tests for evidence of horizontal pleiotropy (Table VII and Figures VI and VII in the Data Supplement).

We then examined the effects of 2 blood pressure traits, SBP and DBP, with AAA through MR. We generated genetic instruments of 390 and 396 variants for the SBP and DBP traits, respectively, from the combined International Consortium for Blood Pressure and UK Biobank discovery GWAS analysis<sup>21</sup> (Table VIII in the Data Supplement). We observed that a genetic 10 mmHg increase in DBP was associated with an increased risk of AAA (OR, 1.43 per 10 mmHg increase in DBP [95% CI, 1.24–1.66];  $P=1.6 \times 10^{-6}$ ; Figure 2B). In sensitivity analyses, the MR-PRESSO global test was significant for evidence of horizontal pleiotropy for DBP; however, the outlier-corrected results demonstrated nearly identical findings (Table IX and Figures VIII and IX in the Data Supplement). We did not detect a significant association between a genetic 10 mmHg increase in SBP and AAA risk (OR, 1.06 per 10 mmHg increase in SBP [95% CI, 0.97–1.15];  $P=0.2$ ).

**Table 2.** Risk Loci Associated With Abdominal Aortic Aneurysm in the MVP Genome-Wide Association Studies Discovery Analysis

Chr:Pos	rsid	EA	NEA	EAF	Overall OR	Overall 95% CI	Overall P	Annotation	Gene/Locus‡	Stage§
Known loci*										
1:109818530	rs646776	T	C	0.7796	1.15	1.11–1.2	6.45E-12	downstream_gene_variant	<i>CELSR2/SORT1</i>	MVP only
1:154419892	rs12730935	G	A	0.6028	1.14	1.1–1.18	3.01E-14	intron_variant	<i>IL6R</i>	MVP only
1:214442481	rs12125521	T	C	0.7115	1.10	1.06–1.14	1.13E-06	intergenic_variant	<i>(SMYD2)</i>	MVP only
9:124422403	rs7025486	A	G	0.2571	1.12	1.08–1.16	5.75E-09	intron_variant	<i>DAB2IP</i>	MVP only
9:22093299	rs4007642	T	A	0.4951	1.21	1.17–1.25	6.94E-29	intron_variant	<i>CDKN2B-AS1</i>	MVP only
12:57527283	rs11172113	T	C	0.5981	1.10	1.06–1.14	5.99E-08	intron_variant	<i>LRP1</i>	MVP only
13:22871446	rs7994761	A	G	0.7772	1.14	1.09–1.19	9.49E-10	intergenic_variant	<i>(LINC00540)</i>	MVP only
19:11191300	rs73015016	G	A	0.8765	1.15	1.09–1.21	1.28E-07	intergenic_variant	<i>(LDLR)</i>	MVP only
20:44608901	rs8124182	A	G	0.1654	1.15	1.1–1.2	8.00E-10	upstream_gene_variant	<i>FTLP1/MMP9</i>	MVP only
21:39819830	rs2836411	T	C	0.3345	1.10	1.06–1.14	8.05E-07	intron_variant	<i>ERG</i>	MVP only
Novel loci†										
2:20878820	rs7255	T	C	0.4631	1.10	1.07–1.13	8.58E-13	non_coding_transcript_exon_variant	<i>AC012065.7/LDAH</i>	2a
4:88772535	rs10023907	T	C	0.6659	1.09	1.06–1.12	1.90E-08	downstream_gene_variant	<i>MEPE</i>	2a
6:36648816	rs3176336	T	A	0.4028	1.10	1.07–1.13	9.50E-11	intron_variant	<i>CDKN1A</i>	2a
8:126495818	rs10808546	C	T	0.5646	1.10	1.07–1.13	1.05E-10	intron_variant	<i>RP11-136O12.2/TRIB1</i>	2a
10:91002804	rs1412445	T	C	0.3392	1.10	1.07–1.13	1.46E-10	intron_variant	<i>LIPA</i>	2a
11:116648917	rs964184	G	C	0.1396	1.18	1.14–1.23	4.59E-19	3_prime_UTR_variant	<i>ZNF259/APOA5</i>	2a
11:130280667	rs4936098	G	A	0.6306	1.13	1.10–1.16	7.00E-16	intron_variant	<i>ADAMTS8</i>	2a
16:84919525	rs35254673	G	A	0.2545	1.09	1.06–1.13	3.11E-08	intron_variant	<i>CRISPLD2</i>	2a
18:20193668	rs4401144	T	C	0.4831	1.11	1.08–1.14	3.63E-14	regulatory_region_variant	<i>(CTAGE1)</i>	2a
19:45411941	rs429358	C	T	0.1401	1.17	1.12–1.21	1.16E-15	missense_variant	<i>APOE</i>	2a
1:55505647	rs11591147	G	T	0.9839	1.58	1.38–1.82	6.43E-11	missense_variant	<i>PCSK9</i>	2b
6:160985526	rs118039278	A	G	0.0663	1.28	1.21–1.35	3.64E-18	intron_variant	<i>LPA</i>	2b
15:78915872	rs55958997	A	C	0.3679	1.12	1.09–1.16	9.06E-14	upstream_gene_variant	<i>CHRNA3</i>	2b
20:62481351	rs73149487	G	T	0.9572	1.26	1.16–1.36	8.33E-09	intergenic_variant	<i>(ABHD16B)</i>	2b

Chr indicates chromosome; EA, effect allele; EAF, effect allele frequency; MVP, Million Veteran Program; NEA, non-effect allele; OR, odds ratio; Pos, position; and rsid, RefSNP identification number.

\*Overall OR, 95% CI, and *P* (2-sided) represent logistic regression statistics after MVP discovery analysis; *n* = 7642 cases and 172 172 controls.

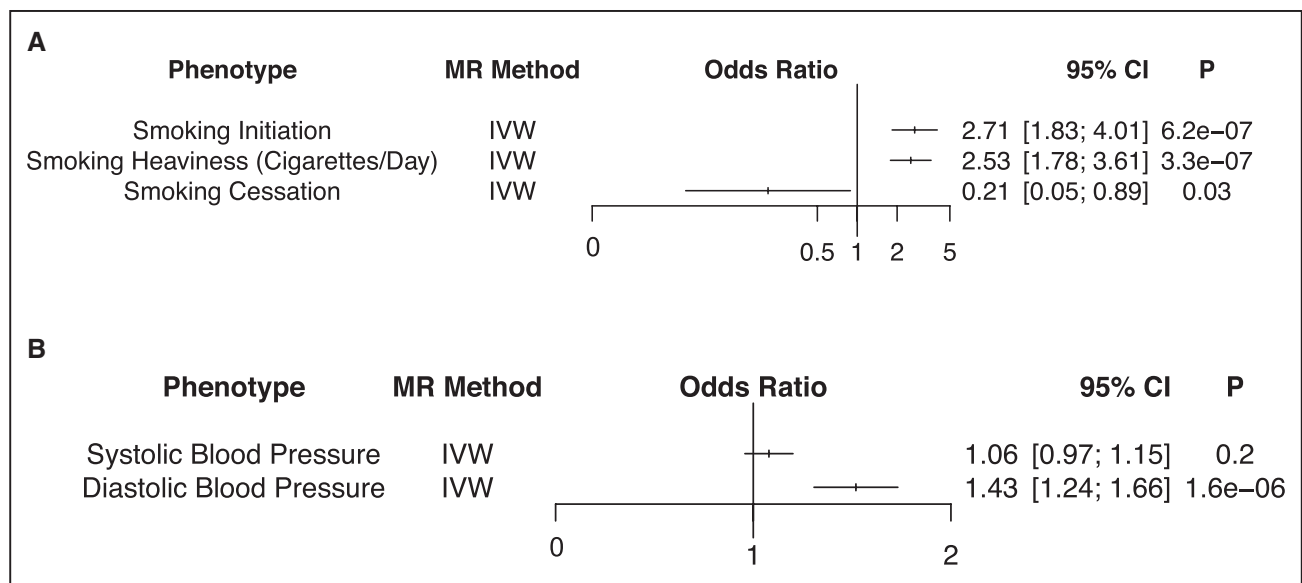
†Overall OR, 95% CI, and *P* (2-sided) represent logistic regression statistics after meta-analysis of MVP and replication studies (total *n* depends on replication source; see Table III in the Data Supplement).

‡Genes for variants that are outside the transcript boundary of a protein-coding gene are shown with nearest candidate gene in parentheses (eg, [ABHD16B]).

§Refers to stage of analysis or replication for variant.

Next, we examined the strength of the relationship between AAA and aneurysms in other vascular territories. In total, we identified 659 cerebral, 520 lower extremity, and 844 iliac aneurysm cases among MVP participants of European ancestry. Whereas a significant relationship was observed for AAA and all 3 aneurysm types, we observed a stronger association among AAA and iliac aneurysms ( $\chi^2 P < 1 \times 10^{-300}$ ) as well as AAA and lower extremity aneurysms ( $\chi^2 P < 1 \times 10^{-300}$ ) than AAA

and cerebral aneurysms ( $\chi^2 P = 6 \times 10^{-12}$ ). A total of 66% of patients with AAA were also noted to have an iliac aneurysm, 44% a lower extremity aneurysm, and 9.7% a cerebral aneurysm (Table X in the Data Supplement). We then extended our analysis to examine the underlying genetics of these diseases. For each of the 24 lead AAA risk variants, we tested their association with cerebral, iliac artery, and lower extremity aneurysms. We observed that 4 AAA risk variants demonstrated at least



**Figure 2. Smoking and blood pressure mendelian randomization (MR) analyses.**

Logistic regression association results of the (A) smoking and (B) blood pressure genetic instruments with the abdominal aortic aneurysm (AAA) outcome in 2-sample MR analyses. The smoking initiation and cessation odds ratios correspond to the change in AAA risk per 1 U log-odds increase in likelihood of the exposure, whereas the smoking heaviness odds ratio corresponds to the change in AAA risk per  $\approx 5$  cigarettes/d increase in smoking. The systolic and diastolic blood pressure odds ratios correspond to the change in AAA risk per 10-mmHg increase in the blood pressure trait. Two-sided values of  $P$  are displayed. IVW indicates inverse variance-weighted.

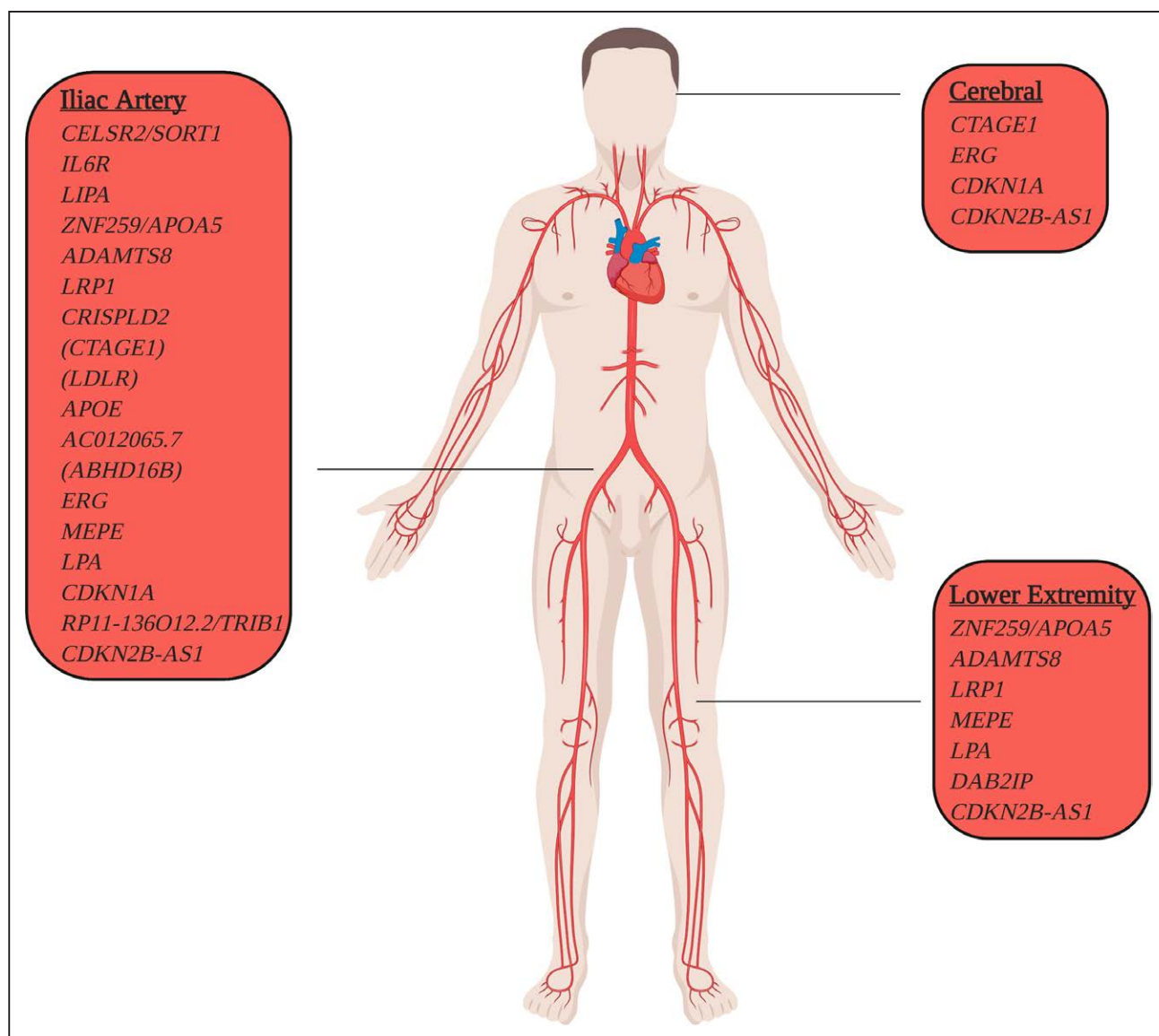
nominal association ( $P < 0.05$ ) with cerebral aneurysms, 7 with lower extremity aneurysms, and 18 with iliac aneurysms (Figure 3 and Table XI in the Data Supplement). In a sensitivity analysis, we examined the association of the AAA risk variants with isolated iliac, lower extremity, and cerebral aneurysms by excluding individuals with a diagnosis of AAA from the analysis. We noted a greater attenuation of association for the AAA risk variants with iliac artery and lower extremity aneurysms than for cerebral aneurysms, likely driven by the greater overlap of AAA diagnosis for aneurysms of the iliac and lower extremity vessels than in the cerebral territory (Table XII in the Data Supplement).

We aimed to examine the contribution of polygenic inheritance on AAA risk. We hypothesized that those at the right tail of the normally distributed AAA PRS (highest 5%) would identify a population at increased risk of AAA who may be candidates for screening. We generated 3 distinct AAA PRSs using a pruning and thresholding method from European MVP release 2.1 AAA summary statistics ( $R^2 < 0.0001$  for each score, association  $P < 0.01$ ,  $P < 1 \times 10^{-4}$ , or  $P < 1 \times 10^{-6}$ ; Table XIII in the Data Supplement). We first tested the associated AAA risk for each of the PRS<sub>AAA</sub> using ascertained AAA case-control cohort data from the Mayo Clinic Vascular Disease Biorepository<sup>34</sup> consisting of 1022 AAA cases and 7750 controls. Demographic and clinical characteristics of this European ancestry cohort are depicted in Table XIV in the Data Supplement. We observed that the 29 variant score ( $R^2 < 0.0001$ ,  $P < 1 \times 10^{-6}$ ; score 1) performed best based on its effect estimate and  $P$  value of association, with a 1-SD increase in

PRS<sub>AAA</sub> associated with a 26% increased risk of AAA (OR<sub>PRS</sub> 1.26 [95% CI, 1.18–1.36];  $P_{PRS} = 2.7 \times 10^{-11}$ ; Figure 4A). Including family history of AAA or smoking in the model only slightly attenuated the association effect estimate for AAA (OR<sub>PRS+FH+smoking</sub> 1.24 [95% CI, 1.14–1.35];  $P_{PRS} = 1.27 \times 10^{-6}$ ; Figure 4A).

We sought validation of our PRS findings using AAA data from 3 additional population-based datasets: (1) prevalent AAA data from MVP release 3.0, which comprised 1656 AAA cases and 44 908 controls of European ancestry and 718 AAA cases and 46 380 controls of African ancestry that were entirely independent from the individuals in the MVP discovery GWAS; (2) a set of 195 AAA cases and 9348 controls from the BioMe Biobank Program; and (3) a set of 396 AAA cases and 9835 controls from the Penn Medicine Biobank (Table XV in the Data Supplement). We observed that the 29-variant PRS<sub>AAA</sub> was associated with AAA in all of the cohorts, though the effect estimate was diminished most in individuals of African ancestry (Figure 4B). In addition, among European individuals with high polygenic risk in each cohort (highest 5%), we noted an AAA prevalence of 5.9% (95% CI, 5.1–6.7%) for all individuals and 8.6% (95% CI, 7.3–9.8%) among men aged 50 or older (Table XVI in the Data Supplement). The prevalence estimates observed in individuals of European ancestry were greater than those found in individuals of African ancestry (1.7% among all individuals in the PRS top 5% and 2.5% among men aged 50 or older). Prevalence results stratified by PRS quintile are depicted in Table XVII in the Data Supplement.





**Figure 3. Abdominal aortic aneurysm (AAA) risk loci with aneurysmal pleiotropy.**

AAA risk loci with evidence of association with aneurysms in other vascular territories. Within our Million Veteran Program European cohort, we identified 659 patients with cerebral, 520 with lower extremity, and 844 with iliac aneurysm. We tested the 24 sentinel AAA risk variants for association with cerebral, lower extremity, and iliac aneurysms using logistic regression after adjusting for age, sex, and principal components of ancestry. AAA risk loci are displayed based on their pleiotropic aneurysmal association ( $P < 0.05$ ).

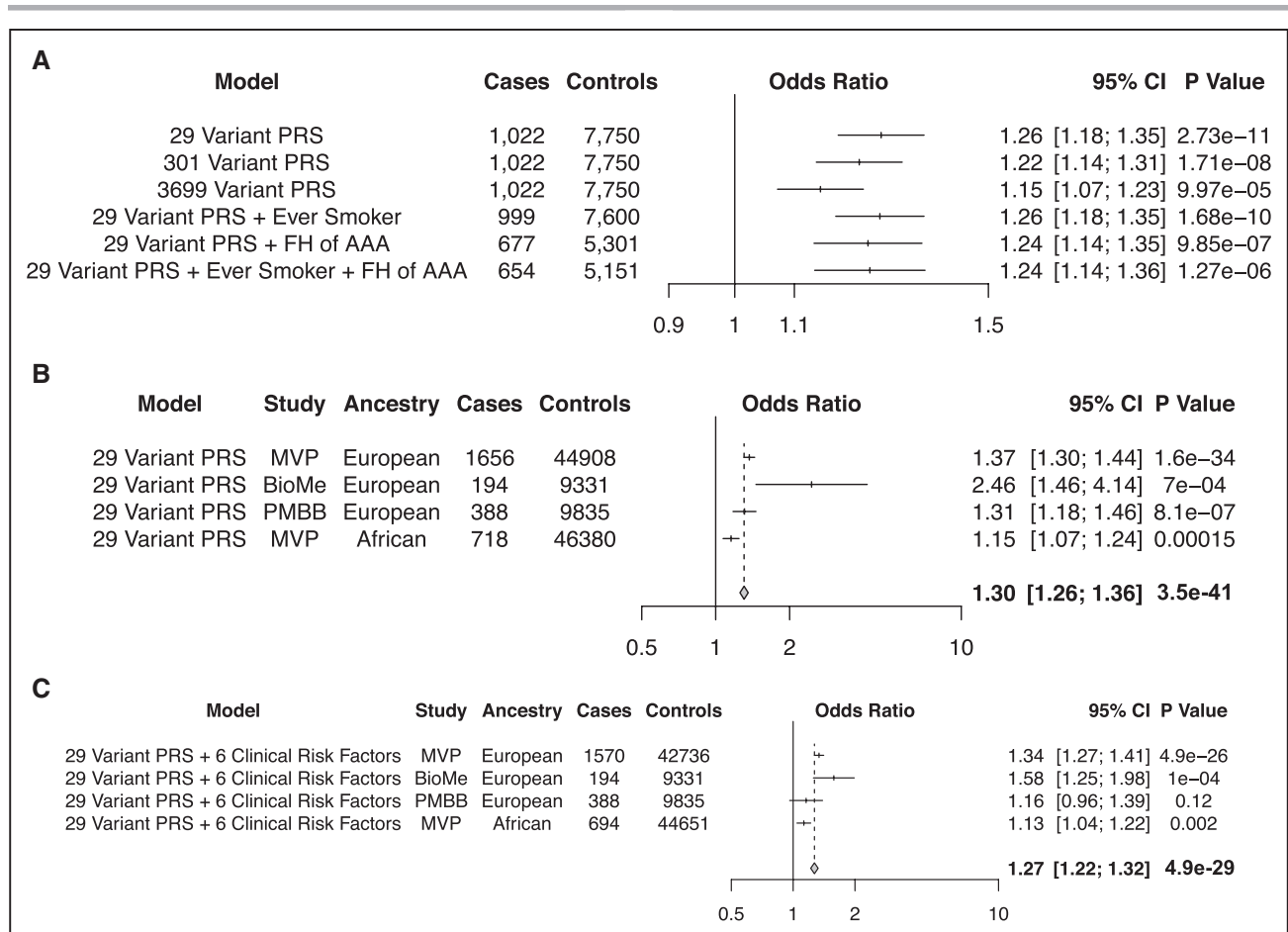
We then included 6 clinical AAA risk factors in the association model to assess whether the PRS represented an orthogonal risk factor beyond what could be used to predict AAA risk on clinical grounds alone. We observed an attenuation in both effect estimate and association  $P$  value but the results remained statistically significant (Figure 4C).

## DISCUSSION

We identified 14 novel loci and examined the overlap of AAA genetic variation with its risk factors at the genome-wide level. We tested AAA lead risk variants with risk of aneurysms in other vascular territories, and through MR, we demonstrate that DBP, as opposed to

SBP, likely has a causal relationship with AAA development. Last, we developed a genome-wide PRS for AAA that identifies a subset of the population at greater risk for AAA independent of known smoking and family history risk factors.

These findings permit several conclusions. First, our findings lend human genetic support to DBP, as opposed to SBP, being of greater significance in the pathogenesis of AAA. Observational epidemiologic data from the CALIBER study (Cardiovascular Research Using Linked Bespoke Studies and Electronic Health Records) demonstrated that increased DBP was associated with incident AAA<sup>35</sup> and that unlike other atherosclerotic syndromes, elevated SBP was not predictive of AAA risk. Our MR results support these findings and



**Figure 4. Genome-wide polygenic risk score (PRS) for abdominal aortic aneurysm (AAA).**

**A**, AAA odds ratios and 2-sided  $P$  values in an ascertained case-control cohort from the Mayo Clinic Vascular Disease Biorepository per 1 SD increase in  $PRS_{AAA}$  for 3 separate PRSs. The addition of smoking or family history (FH) to the association model only slightly attenuated the effect estimate observed. Case/control counts in the models including smoking and FH are reduced after accounting for missingness for these variables across the cohort. **B**, AAA odds ratios and 2-sided  $P$  values in 4 different populations (Europeans in the Million Veteran Program [MVP], BioMe, and Penn Medicine Biobank [PMBB]; Africans in the MVP) per 1 SD increase in  $PRS_{AAA}$ . Results were combined in an inverse variance-weighted fixed effects meta-analysis. **C**, AAA odds ratios and 2-sided  $P$  values in 4 different populations (Europeans in the MVP, BioMe, and PMBB; Africans in the MVP) per 1 SD increase in  $PRS_{AAA}$ , accounting for 6 additional clinical risk factors in the association model. Results were combined in an inverse variance-weighted fixed effects meta-analysis.

suggest a causal role for elevated DBP in aneurysm development. Given that diastole represents  $\approx 2/3$  of the cardiac cycle, our results suggest that abdominal aortic exposure to a hypertensive environment over a long period of time may be more important than exposure to brief, elevated peaks in blood pressure in promoting aneurysmal degeneration. Whereas many of the same causal pathways for atherosclerosis have been implicated in AAA, including dyslipidemia and tobacco exposure, a key difference appears to be that genetically increased SBP—previously shown to predispose to an increased risk for atherosclerosis<sup>36</sup>—is not as important in the pathogenesis of AAA.

Second, our results highlight biological similarities and differences among aneurysms of the aortic, cerebral, lower extremity, and iliac systems. Previous analyses demonstrated that cerebral aneurysms may be observed in up to 10% of patients with AAA,<sup>37</sup> and estimates for lower extremity and iliac aneurysms among those with AAA approach 14%<sup>38</sup> and 40%,<sup>39</sup>

respectively. In our population-based genetic analysis, we observed a much greater amount of disease overlap between AAA and iliac or lower extremity aneurysms than previously reported, with the shared genetic risk associations reflecting this distribution. Our findings support the concept that the biology underlying AAA, iliac, and lower extremity aneurysms are likely closely related, whereas cerebral aneurysms appear to be a physiologic entity divergent in genetic origin. Whereas the final common pathway of arterial wall degeneration is common to all 4 diseases, these results suggest that therapies targeting AAA are more likely to be efficacious for preventing aneurysmal disease in the lower extremity and iliac vascular beds.

Last, our data identify a role for polygenic risk prediction in the targeting of asymptomatic individuals for AAA screening. AAA remains underdiagnosed in the general population. A 2013 report examining Medicare data demonstrated that up to 40% of individuals with AAA were diagnosed late in their disease course, and

late diagnosis was associated with an increased likelihood of rupture at the time of repair.<sup>40</sup> The US Preventive Service Task Force recently released updated guidelines<sup>41</sup> regarding the utility of screening asymptomatic patients on the basis of pooled data from 4 randomized control trials.<sup>42–45</sup> Across these studies, a net reduction in aneurysm-related mortality and incidence of rupture was observed by screening men >65 years of age, leading to the recommendation to screen individuals aged 65 to 75 with a history of smoking with a single visceral ultrasound. This recommendation was based on the assumption that 6% to 7% of individuals in this cohort will have an identified AAA.<sup>25</sup> Here, through the use of a 29-variant PRS<sub>AAA</sub>, we identified a subset of the European ancestry population who had high genetic risk and an AAA prevalence similar to (5.9% across the top 5% PRS<sub>AAA</sub>) or greater than (8.6% among men over 50 in the top 5% PRS<sub>AAA</sub>) that observed in screening trials informing current guidelines. Furthermore, a family history of AAA, often touted as an equivalent means of assessing genetic risk, only slightly attenuated the PRS association. Results from our study suggest that PRSs represent an orthogonal source of disease risk not captured by standard AAA risk factors, and testing individuals with a high PRS<sub>AAA</sub> through a noninvasive abdominal duplex study may increase the yield of AAA screening, particularly once genotyping becomes standard of care in healthcare systems and can be performed at nominal cost.

These results also underscore an important issue with respect to applicability of polygenic scoring across ancestry groups. To date, the overwhelming majority of PRSs have been developed using summary statistics from patients with European ancestry. Whereas the 29-variant score constructed here associated with AAA in individuals with African ancestry, the attenuated effect estimate could limit its predictive accuracy. As the field of genetics moves forward, it is critical that large-scale association studies include individuals of diverse genetic backgrounds to limit the potential for ethnic disparities in precision medicine.

Our study should be interpreted within the context of its limitations. First, our AAA phenotype is based on EHR data and may result in misclassification of case status. Such misclassification should, however, reduce statistical power for discovery and on average bias results toward the null. Second, although those with the highest PRS<sub>AAA</sub> are at increased risk for AAA, the PRS mechanism of action represents a combination of many causal risk factors, rather than a single pathway that leads to disease. However, assessment of individual risk can aid in identifying those at highest risk for AAA and more likely to obtain benefit from screening, regardless of mechanism. Third, the Veterans Health Administration healthcare system contains a much higher proportion of men than the general population, and our ability to detect female-specific associations was more limited.

However, AAA is a disease process that overwhelmingly affects men,<sup>25</sup> and our results should therefore still be generalizable to patients with AAA.

We identified 14 novel genomic loci associated with AAA risk, explored the phenotypic consequences of blood pressure on AAA, identified AAA risk variants associated with aneurysms in other vascular beds, and generated a PRS for AAA that may improve the current screening paradigm once genotyping costs are further reduced in the future. Our data provide mechanistic insights into the genetic architecture of AAA that can inform clinical care.

## ARTICLE INFORMATION

Received April 11, 2020; accepted July 23, 2020.

This article does not represent the views of the Department of Veterans Affairs or the US government.

The Data Supplement is available with this article at <https://www.ahajournals.org/doi/suppl/10.1161/CIRCULATIONAHA.120.047544>.

## Authors

Derek Klarin, MD; Shefali Setia Verma, PhD; Renae Judy, MS; Ozan Dikilitas<sup>1</sup>, MD; Brooke N. Wolford<sup>2</sup>, BS; Ishan Paranjpe, BS; Michael G. Levin<sup>3</sup>, MD; Cuiping Pan<sup>4</sup>, PhD; Catherine Tcheandjieu<sup>5</sup>, PhD; Joshua M. Spin, MD, PhD; Julie Lynch, PhD, RN, MBA; Themistocles L. Assimes<sup>6</sup>, MD, PhD; Linn Åldstedt Nyronning, MD; Erney Mattsson, MD; Todd L. Edwards, PhD; Josh Denny, MD, MS; Eric Larson, MD; Ming Ta Michael Lee, PhD; David Carrell, PhD; Yanfei Zhang, MD, PhD; Gail P. Jarvik, MD, PhD; Ali G. Gharavi<sup>7</sup>, MD; John Harley<sup>8</sup>, MD, PhD; Frank Mentch, PhD; Jennifer A. Pacheco, MS; Hakon Hakonarson, MD; Anne Heidi Skogholt, MS; Laurent Thomas<sup>9</sup>, PhD; Maiken Elvestad Gabrielsen, PhD; Kristian Hveem, MD, PhD; Jonas Bille Nielsen<sup>10</sup>, MD, PhD; Wei Zhou, PhD; Lars Fritsche<sup>11</sup>, PhD; Jie Huang, MD, PhD; Pradeep Natarajan<sup>12</sup>, MD, MMSc; Yan V. Sun, PhD; Scott L. DuVall, PhD; Daniel J. Rader<sup>13</sup>, MD; Kelly Cho, PhD; Kyong-Mi Chang, MD; Peter W.F. Wilson<sup>14</sup>, MD; Christopher J. O'Donnell, MD, MPH; Sekar Kathiresan<sup>15</sup>, MD; Salvatore T. Scali, MD; Scott A. Berceci, MD, PhD; Cristen Willer, PhD; Gregory T. Jones<sup>16</sup>, PhD; Matthew J. Bown, MD; Girish Nadkarni, MD; Iftikhar J. Kullo<sup>17</sup>, MD; Marylyn Ritchie, PhD; Scott M. Damrauer<sup>18</sup>, MD\*; Philip S. Tsao<sup>19</sup>, PhD\*; On behalf of the Veterans Affairs Million Veteran Program

\*Drs Damrauer and Tsao jointly supervised this work.

## Correspondence

Philip S. Tsao, PhD, VA Palo Alto Health Care System, 3801 Miranda Avenue, Palo Alto, CA 94304. Email [philip.tsao@va.gov](mailto:philip.tsao@va.gov)

## Affiliations

Malcolm Randall VA Medical Center, Gainesville, FL (D.K., S.T.S., S.A.B.). Division of Vascular Surgery and Endovascular Therapy, University of Florida College of Medicine, Gainesville (D.K., S.T.S., S.A.B.). Center for Genomic Medicine (D.K., W.Z., P.N.) and Department of Medicine (P.N.), Massachusetts General Hospital, Harvard Medical School, Boston. Program in Medical and Population Genetics (D.K.) and Stanley Center for Psychiatric Research (W.Z.), Broad Institute of MIT and Harvard, Cambridge, MA. Department of Genetics (S.S.V., M.R.), Department of Surgery (R.J., S.M.D.), Division of Cardiovascular Medicine (M.G.L.), Department of Medicine (M.G.L., D.J.R., K.-M.C.), and Department of Pediatrics (H.H.), Perelman School of Medicine, University of Pennsylvania, Philadelphia. Corporal Michael J. Crescenz VA Medical Center, Philadelphia, PA (R.J., M.G.L., K.-M.C., S.M.D.). Department of Cardiovascular Medicine, Mayo Clinic, Rochester, MN (O.D., I.J.K.). Department of Computational Medicine and Bioinformatics (B.N.W., C.W.), Department of Biostatistics (L.F.), Department of Internal Medicine, Division of Cardiology (C.W.), and Department of Human Genetics (C.W.), University of Michigan Medical School, Ann Arbor. Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY (I.P., G.N.). Palo Alto Epidemiology Research and Information Center for Genomics (C.P.), VA Palo Alto Health Care System (C.T., J.M.S., T.L.A., P.S.T.), CA. Division of Cardiovascular Medicine, Department of Medicine (C.T., J.M.S., T.L.A., P.S.T.), and Department of Pediatric Cardiology (C.T.), Stanford University

School of Medicine, CA. Edith Nourse VA Medical Center, Bedford, MA (J.L.). VA Informatics and Computing Infrastructure, VA Salt Lake City Health Care System, UT (J.L., S.L.D.). Department of Vascular Surgery, St. Olavs Hospital, Trondheim, Norway (L.Å.N., E.M.). Department of Circulation and Medical Imaging (L.Å.N., E.M.), Faculty of Medicine and Health Sciences (A.H.S., L.T., M.E.G., K.H., J.B.N.), and Department of Clinical and Molecular Medicine (L.T.), Norwegian University of Science and Technology, Trondheim, Norway. Division of Epidemiology, Department of Medicine, Vanderbilt-Ingram Cancer Center (T.L.E.), Vanderbilt Genetics Institute (T.L.E., J.D.), and Department of Biomedical Informatics (J.D., E.L., D.C.), Vanderbilt University Medical Center, Nashville, TN. Kaiser Permanente Washington Health Research Institute, Seattle (J.D., E.L., D.C.). Departments of Medicine and Health Services (E.L.) and Division of Medical Genetics, Departments of Medicine and Genome Sciences (G.P.J.), University of Washington, Seattle. Genomic Medicine Institute, Geisinger Health System, Danville, PA (M.T.M.L., Y.Z.). Division of Nephrology and Center for Precision Medicine and Genomics, Columbia University, New York, NY (A.G.G.). Center for Autoimmune Genomics and Etiology (CAGE), Cincinnati Children's Hospital Medical Center, OH (J.H.). Department of Pediatrics, University of Cincinnati College of Medicine, OH (J.H.). US Department of Veterans Affairs, Cincinnati, OH (J.H.). Center for Applied Genomics, The Children's Hospital of Philadelphia, PA (F.M., H.H.). Center for Genetic Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL (J.A.P.). K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, Department of Epidemiology Research, Statens Serum Institute, Copenhagen, Denmark (J.B.N.). Analytic and Translational Genetics Unit (W.Z.) and Cardiovascular Research Center (P.N.), Massachusetts General Hospital, Boston. Boston VA Healthcare System, MA (J.H., P.N., K.C., C.J.O.). Department of Epidemiology, Emory University Rollins School of Public Health, Atlanta, GA (Y.V.S.). Atlanta VA Health Care System, Decatur, GA (Y.V.S., P.W.F.W.). Division of Epidemiology, Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City (S.L.D.). Emory Clinical Cardiovascular Research Institute, Atlanta, GA (P.W.F.W.). Cardiovascular Medicine Division, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA (C.J.O.). Verve Therapeutics, Cambridge, MA (S.K.). Department of Surgical Sciences, Dunedin School of Medicine, University of Otago, New Zealand (G.T.J.). Department of Cardiovascular Sciences and NIHR Leicester Biomedical Research Centre, University of Leicester, United Kingdom (M.J.B.).

### Sources of Funding

This work was supported by funding from the Department of Veterans Affairs Office of Research and Development, Million Veteran Program Grant MVP000; Department of Veterans Affairs awards I01-01BX03340 (Drs Cho and Wilson), I01-BX003362 (Drs Chang and Tsao), and IK2-CX001780 (Dr Damrauer); Veterans Affairs Informatics and Computing Infrastructure (VINCI) grant VA HSR RES 130457 (Dr DuVall); and National Institutes of Health grant K08HL140203 (Dr Natarajan). Phase III of the eMERGE Network (electronic Medical Records and Genomics) was initiated and funded by the National Human Genome Research Institute through the following grants: U01HG8657 (Kaiser Washington/University of Washington), U01HG8685 (Brigham and Women's Hospital), U01HG8672 (Vanderbilt University Medical Center), U01HG8666 (Cincinnati Children's Hospital Medical Center), U01HG6379 (Mayo Clinic), U01HG8679 (Geisinger Clinic), U01HG8680 (Columbia University Health Sciences), U01HG8684 (Children's Hospital of Philadelphia), U01HG8673 (Northwestern University), U01HG8701 (Vanderbilt University Medical Center serving as the Coordinating Center), U01HG8676 (Partners Healthcare/Broad Institute), and U01HG8664 (Baylor College of Medicine). Phase I and II of the eMERGE Network was initiated and funded by the National Human Genome Research Institute through the following grants: U01HG006828 (Cincinnati Children's Hospital Medical Center/Boston Children's Hospital), U01HG006830 (Children's Hospital of Philadelphia), U01HG006389 (Essentia Institute of Rural Health, Marshfield Clinic Research Foundation, and Pennsylvania State University), U01HG006382 (Geisinger Clinic), U01HG006375 (Group Health Cooperative/University of Washington), U01HG006379 (Mayo Clinic), U01HG006380 (Icahn School of Medicine at Mount Sinai), U01HG006388 (Northwestern University), U01HG006378 (Vanderbilt University Medical Center), U01HG006385 (Vanderbilt University Medical Center, serving as the Coordinating Center), U01HG004438 (Center for Inherited Disease Research, serving as a Genotyping Center), and U01HG004424 (the Broad Institute, serving as a Genotyping Center).

### Disclosures

Dr Klarin has received consulting fees from Regeneron Pharmaceuticals unrelated to the work in this article. Dr DuVall reports grants and nonfinancial support

from the Department of Veterans Affairs during the conduct of the study; and grants from AbbVie Inc, Amgen Inc, Anolinx LLC, Astellas Pharma Inc, AstraZeneca Pharmaceuticals LP, Boehringer Ingelheim International GmbH, Celgene Corporation, Eli Lilly and Company, Genentech Inc, Genomic Health Inc, Gil-ead Sciences Inc, GlaxoSmithKline PLC, Innocrin Pharmaceuticals Inc, Janssen Pharmaceuticals Inc, Kantar Health, Myriad Genetics Laboratories Inc, Novartis International AG, and Parexel International Corporation outside the submitted work. Dr Damrauer receives research support to his institution from RenalytixAI and is a paid consultant for Calico Labs. Dr Kathiresan is a founder of Maze Therapeutics, Verve Therapeutics, and San Therapeutics; holds equity in Cata-basis and San Therapeutics; is a member of the scientific advisory boards for Regeneron Genetics Center and Corvidia Therapeutics; served as a consultant for Acceleron, Eli Lilly, Novartis, Merck, NovoNordisk, Novo Ventures, Ionis, Alnylam, Aegerion, Huag Partners, Noble Insights, Leerink Partners, Bayer Healthcare, Illumina, Color Genomics, MedGenome, Quest, and Medscape; and reports patents related to a method of identifying and treating a person having a predisposition to or afflicted with cardiometabolic disease (20180010185) and a genetics risk predictor (20190017119).

### Supplemental Materials

Supplemental Methods  
Data Supplement Figures I–IX  
Data Supplement Tables I–XVII  
Data Supplement Excel File I  
References 46–54

### REFERENCES

- Sidloff D, Stather P, Dattani N, Bown M, Thompson J, Sayers R, Choke E. Aneurysm global epidemiology study: public health measures can further reduce abdominal aortic aneurysm mortality. *Circulation*. 2014;129:747–753. doi: 10.1161/CIRCULATIONAHA.113.005457
- Pinard A, Jones GT, Milewicz DM. Genetics of thoracic and abdominal aortic diseases. *Circ Res*. 2019;124:588–606. doi: 10.1161/CIRCRESAHA.118.312436
- Wahlgren CM, Larsson E, Magnusson PK, Hultgren R, Swedenborg J. Genetic and environmental contributions to abdominal aortic aneurysm development in a twin population. *J Vasc Surg*. 2010;51:3–7; discussion 7. doi: 10.1016/j.jvs.2009.08.036
- Gretarsdóttir S, Baas AF, Thorleifsson G, Holm H, den Heijer M, de Vries JP, Kranendonk SE, Zeebregts CJ, van Sterkenburg SM, Geelkerken RH, et al. Genome-wide association study identifies a sequence variant within the DAB2IP gene conferring susceptibility to abdominal aortic aneurysm. *Nat Genet*. 2010;42:692–697. doi: 10.1038/ng.622
- Bown MJ, Jones GT, Harrison SC, Wright BJ, Bumpstead S, Baas AF, Gretarsdóttir S, Badger SA, Bradley DT, Burnand K, et al. Cardiogram Consortium; Global BPgen Consortium; DIAGRAM Consortium; VRCNZ Consortium. Abdominal aortic aneurysm is associated with a variant in low-density lipoprotein receptor-related protein 1. *Am J Hum Genet*. 2011;89:619–627. doi: 10.1016/j.ajhg.2011.10.002
- Bradley DT, Hughes AE, Badger SA, Jones GT, Harrison SC, Wright BJ, Bumpstead S, Baas AF, Grétarsdóttir S, Burnand K, et al. A variant in LDLR is associated with abdominal aortic aneurysm. *Circ Cardiovasc Genet*. 2013;6:498–504. doi: 10.1161/CIRCGENETICS.113.000165
- Jones GT, Tromp G, Kuivaniemi H, Gretarsdóttir S, Baas AF, Giusti B, Strauss E, Van't Hof FN, Webb TR, Erdman R, et al. Meta-analysis of genome-wide association studies for abdominal aortic aneurysm identifies four new disease-specific risk loci. *Circ Res*. 2017;120:341–353. doi: 10.1161/CIRCRESAHA.116.308765
- Collins R. What makes UK Biobank special? *Lancet*. 2012;379:1173–1174. doi: 10.1016/S0140-6736(12)60404-8
- Gaziano JM, Concato J, Brophy M, Fiore L, Pyarajan S, Breeling J, Whitbourne S, Deen J, Shannon C, Humphries D, et al. Million Veteran Program: a mega-biobank to study genetic influences on health and disease. *J Clin Epidemiol*. 2016;70:214–223. doi: 10.1016/j.jclinepi.2015.09.016
- Klarin D, Lynch J, Aragam K, Chaffin M, Assimes TL, Huang J, Lee KM, Shao Q, Huffman JE, Natarajan P, et al; VA Million Veteran Program. Genome-wide association study of peripheral artery disease in the Million Veteran Program. *Nat Med*. 2019;25:1274–1279. doi: 10.1038/s41591-019-0492-5
- Klarin D, Busenkell E, Judy R, Lynch J, Levin M, Haessler J, Aragam K, Chaffin M, Haas M, Lindström S, et al; INVENT Consortium; Veterans Affairs' Million Veteran Program. Genome-wide association analysis of

- venous thromboembolism identifies new risk loci and genetic overlap with arterial vascular disease. *Nat Genet.* 2019;51:1574–1579. doi: 10.1038/s41588-019-0519-3
12. McCarty CA, Chisholm RL, Chute CG, Kullo IJ, Jarvik GP, Larson EB, Li R, Masys DR, Ritchie MD, Roden DM, et al; eMERGE Team. The eMERGE Network: a consortium of biorepositories linked to electronic medical records data for conducting genomic studies. *BMC Med Genomics.* 2011;4:13. doi: 10.1186/1755-8794-4-13
  13. Cai T, Zhang Y, Ho YL, Link N, Sun J, Huang J, Cai TA, Damrauer S, Ahuja Y, Honerlaw J, et al; VA Million Veteran Program. Association of interleukin 6 receptor variant with cardiovascular disease effects of interleukin 6 receptor blocking therapy: a phenome-wide association study. *JAMA Cardiol.* 2018;3:849–857. doi: 10.1001/jamacardio.2018.2287
  14. Denny JC, Bastarache L, Ritchie MD, Carroll RJ, Zink R, Mosley JD, Field JR, Pulley JM, Ramirez AH, Bowton E, et al. Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. *Nat Biotechnol.* 2013;31:1102–1110. doi: 10.1038/nbt.2749
  15. Carroll RJ, Bastarache L, Denny JC. R PheWAS: data analysis and plotting tools for phenome-wide association studies in the R environment. *Bioinformatics.* 2014;30:2375–2376. doi: 10.1093/bioinformatics/btu197
  16. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, Duncan L, Perry JR, Patterson N, Robinson EB, et al; ReproGen Consortium; Psychiatric Genomics Consortium; Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control Consortium 3. An atlas of genetic correlations across human diseases and traits. *Nat Genet.* 2015;47:1236–1241. doi: 10.1038/ng.3406
  17. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, Pirruccello JP, Ripatti S, Chasman DI, Willer CJ, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature.* 2010;466:707–713. doi: 10.1038/nature09270
  18. Cardiogram plus C4D Consortium. A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet.* 2015;47:1121–1130. doi: 10.1038/ng.3396
  19. Zheng J, Erzurumluoglu AM, Elsworth BL, Kemp JP, Howe L, Haycock PC, Hemani G, Tansey K, Laurin C, Pourcain BS, et al; Early Genetics and Lifecourse Epidemiology (EAGLE) Eczema Consortium. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics.* 2017;33:272–279. doi: 10.1093/bioinformatics/btw613
  20. Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, Chen F, Datta G, Davila-Velderrain J, McGuire D, Tian C, et al; 23andMe Research Team; HUNT All-In Psychiatry. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat Genet.* 2019;51:237–244. doi: 10.1038/s41588-018-0307-5
  21. Evangelou E, Warren HR, Mosen-Ansorena D, Mifsud B, Pazoki R, Gao H, Ntritsos G, Dimou N, Cabrera CP, Karaman I, et al; Million Veteran Program. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nat Genet.* 2018;50:1412–1425. doi: 10.1038/s41588-018-0205-x
  22. Burgess S, Bowden J, Fall T, Ingelsson E, Thompson SG. Sensitivity analyses for robust causal inference from mendelian randomization analyses with multiple genetic variants. *Epidemiology.* 2017;28:30–42. doi: 10.1097/EDE.0000000000000559
  23. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol.* 2015;44:512–525. doi: 10.1093/ije/dyv080
  24. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from mendelian randomization between complex traits and diseases. *Nat Genet.* 2018;50:693–698. doi: 10.1038/s41588-018-0099-7
  25. Chaikof EL, Dalman RL, Eskandari MK, Jackson BM, Lee WA, Mansour MA, Mastracci TM, Mell M, Murad MH, Nguyen LL, et al. The Society for Vascular Surgery practice guidelines on the care of patients with an abdominal aortic aneurysm. *J Vasc Surg.* 2018;67:2–77.e2. doi: 10.1016/j.jvs.2017.10.044
  26. Zhou W, Nielsen JB, Fritsche LG, Dey R, Gabrielsen ME, Wolford BN, LeFaive J, VandeHaar P, Gagliano SA, Gifford A, et al. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat Genet.* 2018;50:1335–1341. doi: 10.1038/s41588-018-0184-y
  27. Hall MA, Wallace J, Lucas A, Kim D, Basile AO, Verma SS, McCarty CA, Brilliant MH, Peissig PL, Kitchner TE, et al. PLATO software provides analytic framework for investigating complexity beyond genome-wide association studies. *Nat Commun.* 2017;8:1167. doi: 10.1038/s41467-017-00802-2
  28. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics.* 2010;26:2190–2191. doi: 10.1093/bioinformatics/btq340
  29. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature.* 2015;526:68–74. doi: 10.1038/nature15393
  30. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2008.
  31. Klarin D, Damrauer SM, Cho K, Sun YV, Teslovich TM, Honerlaw J, Gagnon DR, DuVall SL, Li J, Peloso GM, et al; Global Lipids Genetics Consortium; Myocardial Infarction Genetics (MIGen) Consortium; Geisinger-Regeneron DiscovEHR Collaboration; VA Million Veteran Program. Genetics of blood lipids among ~300,000 multi-ethnic participants of the Million Veteran Program. *Nat Genet.* 2018;50:1514–1523. doi: 10.1038/s41588-018-0222-9
  32. Thorgerirsson TE, Geller F, Sulem P, Rafnar T, Wiste A, Magnusson KP, Manolescu A, Thorleifsson G, Stefansson H, Ingason A, et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature.* 2008;452:638–642. doi: 10.1038/nature06846
  33. Anttila V, Bulik-Sullivan B, Finucane HK, Walters RK, Bras J, Duncan L, Escott-Price V, Falcone GJ, Gormley P, Malik R, et al. Analysis of shared heritability in common disorders of the brain. *Science.* 2018;360:eaap8757. doi: 10.1126/science.aap8757
  34. Ye Z, Kallou FS, Dalenberg AK, Kullo IJ. An electronic medical record-linked biorepository to identify novel biomarkers for atherosclerotic cardiovascular disease. *Glob Cardiol Sci Pract.* 2013;2013:82–90. doi: 10.5339/gcsp.2013.10
  35. Rapsomaniki E, Timmis A, George J, Pujades-Rodriguez M, Shah AD, Denaxas S, White IR, Caulfield MJ, Deanfield JE, Smeeth L, et al. Blood pressure and incidence of twelve cardiovascular diseases: lifetime risks, healthy life-years lost, and age-specific associations in 1-25 million people. *Lancet.* 2014;383:1899–1911. doi: 10.1016/S0140-6736(14)60685-1
  36. Lieb W, Jansen H, Loley C, Pencina MJ, Nelson CP, Newton-Cheh C, Kathiresan S, Reilly MP, Assimes TL, Boerwinkle E, et al; CARDIoGRAM. Genetic predisposition to higher blood pressure increases coronary artery disease risk. *Hypertension.* 2013;61:995–1001. doi: 10.1161/HYPERTENSIONAHA.111.00275
  37. Rouchaud A, Brandt MD, Rydberg AM, Kadivel R, Flemming K, Kallmes DF, Brinjikji W. Prevalence of intracranial aneurysms in patients with aortic aneurysms. *AJNR Am J Neuroradiol.* 2016;37:1664–1668. doi: 10.3174/ajnr.A4827
  38. Diwan A, Sarkar R, Stanley JC, Zelenock GB, Wakefield TW. Incidence of femoral and popliteal artery aneurysms in patients with abdominal aortic aneurysms. *J Vasc Surg.* 2000;31:863–869. doi: 10.1067/mva.2000.105955
  39. Richards T, Dharmadasa A, Davies R, Murphy M, Perera R, Walton J. Natural history of the common iliac artery in the presence of an abdominal aortic aneurysm. *J Vasc Surg.* 2009;49:881–885. doi: 10.1016/j.jvs.2008.11.025
  40. Mell MW, Hlatky MA, Shreibati JB, Dalman RL, Baker LC. Late diagnosis of abdominal aortic aneurysms substantiates underutilization of abdominal aortic aneurysm screening for Medicare beneficiaries. *J Vasc Surg.* 2013;57:1519–1523, 1523.e1. doi: 10.1016/j.jvs.2012.12.034
  41. Owens DK, Davidson KW, Krist AH, Barry MJ, Cabana M, Caughey AB, Doubeni CA, Epling JW Jr, Kubik M, Landefeld CS, et al. Screening for abdominal aortic aneurysm: US Preventive Services Task Force recommendation statement. *JAMA.* 2019;322:2211–2218. doi: 10.1001/jama.2019.18928
  42. Scott RA, Wilson NM, Ashton HA, Kay DN. Influence of screening on the incidence of ruptured abdominal aortic aneurysm: 5-year results of a randomized controlled study. *Br J Surg.* 1995;82:1066–1070. doi: 10.1002/bjs.1800820821
  43. Ashton HA, Buxton MJ, Day NE, Kim LG, Marteau TM, Scott RA, Thompson SG, Walker NM; Multicentre Aneurysm Screening Study Group. The Multicentre Aneurysm Screening Study (MASS) into the effect of abdominal aortic aneurysm screening on mortality in men: a randomised controlled trial. *Lancet.* 2002;360:1531–1539. doi: 10.1016/S0140-6736(02)11522-4
  44. Norman PE, Jamrozik K, Lawrence-Brown MM, Le MT, Spencer CA, Tuohy RJ, Parsons RW, Dickinson JA. Population based randomised controlled trial on impact of screening on mortality from abdominal aortic aneurysm. *BMJ.* 2004;329:1259. doi: 10.1136/bmj.38272.478438.55

45. Lindholt JS, Juul S, Fasting H, Henneberg EW. Screening for abdominal aortic aneurysms: single centre randomised controlled trial. *BMJ*. 2005;330:750. doi: 10.1136/bmj.38369.620162.82
46. Hunter-Zinck H, Shi Y, Li M, Gorman BR, Ji SG, Sun N, Webster T, Liem A, Hsieh P, Devineni P, et al; VA Million Veteran Program. Genotyping array design and data quality control in the Million Veteran Program. *Am J Hum Genet*. 2020;106:535–548. doi: 10.1016/j.ajhg.2020.03.004
47. Loh PR, Palamara PF, Price AL. Fast and accurate long-range phasing in a UK Biobank cohort. *Nat Genet*. 2016;48:811–816. doi: 10.1038/ng.3571
48. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet*. 2012;44:955–959. doi: 10.1038/ng.2354
49. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006;38:904–909. doi: 10.1038/ng1847
50. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*. 2009;19:1655–1664. doi: 10.1101/gr.094052.109
51. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. *Bioinformatics*. 2010;26:2867–2873. doi: 10.1093/bioinformatics/btq559
52. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559–575. doi: 10.1086/519795
53. McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A, Kang HM, Fuchsberger C, Danecek P, Sharp K, et al; Haplotype Reference Consortium. A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet*. 2016;48:1279–1283. doi: 10.1038/ng.3643
54. Stanaway IB, Hall TO, Rosenthal EA, Palmer M, Naranbhai V, Knevel R, Namjou-Khales B, Carroll RJ, Kiryluk K, Gordon AS, et al; eMERGE Network. The eMERGE genotype set of 83,717 subjects imputed to ~40 million variants genome wide and association with the herpes zoster medical record phenotype. *Genet Epidemiol*. 2019;43:63–81. doi: 10.1002/gepi.22167