Three-dimensional aortic geometry: clinical correlates, prognostic value and genetic architecture

Cameron Beeche†, BS, PhD student[1,2]; Marie-Joe Dib†, PhD[2,4]; Bingxin Zhao[3], PhD; Joe David Azzo, MD[4]; Hannah Maynard, MPH, BS[2]; Jeffrey Duda, PhD[5]; James Gee[5], PhD; Oday Salman, MD[4]; Penn Medicine BioBank; Walter R. Witschey PhD[5]; Julio A. Chirinos, MD, PhD[2,4*]

[1]Department of Bioengineering, University of Pennsylvania, 3400 Spruce Street, Philadelphia, PA, 19104
[2]Division of Cardiovascular Medicine, Hospital of the University of Pennsylvania, 3400 Spruce Street, Philadelphia, PA, 19104
[3]Department of Statistics and Data Science, University of Pennsylvania, 3400 Spruce Street, Philadelphia, PA, 19104
[3]University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, 3400 Spruce Street, Philadelphia, PA, 19104
[5]Department of Radiology, Perelman School of Medicine, University of Pennsylvania, 3400 Spruce Street, Philadelphia, PA, 19104

† **Contributed equally to this work.**

Funding:
J.A.C. is supported by NIH grants R01-HL 121510, R33-HL-146390, R01HL153646, R01-AG058969, 1R01-HL104106, P01-HL094307, R03-HL146874, K24-AG070459 and R56-HL136730. W.R.W. is supported by NIH grants P41-EB029460, R01 HL169378, R01 HL137984, UL1 TR001878. J.G is supported by R01EB031722, R01HL133889.

***Address for correspondence:**
Julio A. Chirinos, MD, PhD
South Tower, Rm. 11-138.
Perelman Center for Advanced Medicine.
3400 Civic Center Blvd. Philadelphia, PA. 19104.
Tel: 215-573-6606; Fax: 215-746-7415
Email: julio.chirinos@uphs.upenn.edu
Twitter: @JulioChirinosMd

Abstract

Aortic structure and function impact cardiovascular health through multiple mechanisms. Aortic structural degeneration increases left ventricular afterload, pulse pressure and promotes target organ damage. Despite the impact of aortic structure on cardiovascular health, aortic 3D-geometry has yet to be comprehensively assessed. Using a convolutional neural network (U-Net) combined with morphological operations, we quantified aortic 3D-geometric phenotypes (AGPs) from 53,612 participants in the UK Biobank and 8,066 participants in the Penn Medicine Biobank. AGPs reflective of structural aortic degeneration, characterized by arch unfolding, descending aortic lengthening and luminal dilation exhibited cross-sectional associations with hypertension and cardiac diseases, and were predictive for new-onset hypertension, heart failure, cardiomyopathy, and atrial fibrillation. We identified 237 novel genetic loci associated with 3D-AGPs. Fibrillin-2 gene polymorphisms were identified as key determinants of aortic arch-3D structure. Mendelian randomization identified putative causal effects of aortic geometry on the risk of chronic kidney disease and stroke.

Keywords: Aorta, genetics, geometry, 3-dimensional, elongation, unfolding

1. Introduction

The aorta is the largest conduit artery in the human body[1]. In addition to its conduit function, the aorta provides a cushioning function via its compliance, allowing for the near steady flow in the systemic microvasculature despite the intermittent left ventricular ejection. Aortic structural parameters (including geometry and wall stiffness) have been shown to be key determinants of pulsatile arterial hemodynamics[1,2]. A stiff aorta loses its cushioning function, resulting in excess arterial pulsatility and damage to the systemic microvasculature of target organs.[1] In addition, a stiff aorta increases pulsatile left ventricular afterload, contributing to the pathogenesis of left ventricular hypertrophy and heart failure. Multiple studies have shown that aortic stiffness is a key determinant of pulse pressure, is involved in the pathogenesis of isolated systolic hypertension, and independently predicts the risk of incident cardiovascular events[1,2].

Although there is a joint interplay between aortic stiffness and geometry as determinants of pulsatile hemodynamics, only the former has been widely studied as a mediator of target organ damage and cardiovascular disease (CVD). In contrast, available studies regarding genetic and biologic determinants of aortic geometry have been limited to cross-sectional diameter assessments at a few discrete aortic locations[3-5]. Importantly, 3-D aortic geometry is complex and has been assessed in only a few relatively small studies. Previous cross-sectional studies among 210-250 patients undergoing aortic imaging demonstrated that the aorta elongates with age and that thoracic aortic length is greater among patients with acute aortic dissection[6,7].

Despite its prominent age-associated changes and its key hemodynamic role, studies related to the population variability, clinical correlates, prognostic significance, and the genetic architecture of 3-D aortic geometric phenotypes (AGPs) are lacking. Narrow geometric phenotyping of the aorta neglects important three-dimensional aspects of its geometry, including

elongation, tortuosity/unfolding, and curvature, all of which influence aortic hemodynamic function. Analyses of image-derived cardiac phenotypes acquired through the segmentation of cardiovascular magnetic resonance imaging have been used to uncover the genetic basis for cardiac structure and function[8,9]. Analyses of aortic 3D-structure in large community-based samples are required to uncover its demographic and clinical correlates. Moreover, genetic analysis of such phenotypes could identify specific loci affecting aortic geometry and uncover causal effects of aortic geometry on relevant cardiovascular diseases through Mendelian randomization (MR).

In this study, we investigated the patterns, distribution, epidemiologic and clinical correlates, prognostic significance, and the genetic basis of 3D aortic geometry among 53,612 participants enrolled in the UK Biobank[10] (UKB) and 8,066 individuals enrolled in the Penn Medicine Biobank[11] (PMBB). We leveraged a deep convolutional neural network to segment the thoracic aorta, then used morphological operations to identify six sub-segments of the aorta to generate 51 unique 3D AGPs that represent key 3D aortic geometric features, including overall and segmental centerline length, diameter, tortuosity, and curvature. We then assessed the relationships between 3D-AGPs with: **(1)** age, sex, various cardiovascular risk factors and prevalent CVD, and **(2)** the risk of incident CVD among participants without CVD at baseline. We further performed GWAS of multiple novel AGPs to investigate the genetic architecture of aortic 3D-geometry, followed by genetic correlation and MR analyses to assess the putative causal effects of AGPs on outcomes of interest. **Figure 1** provides a summary of our study.

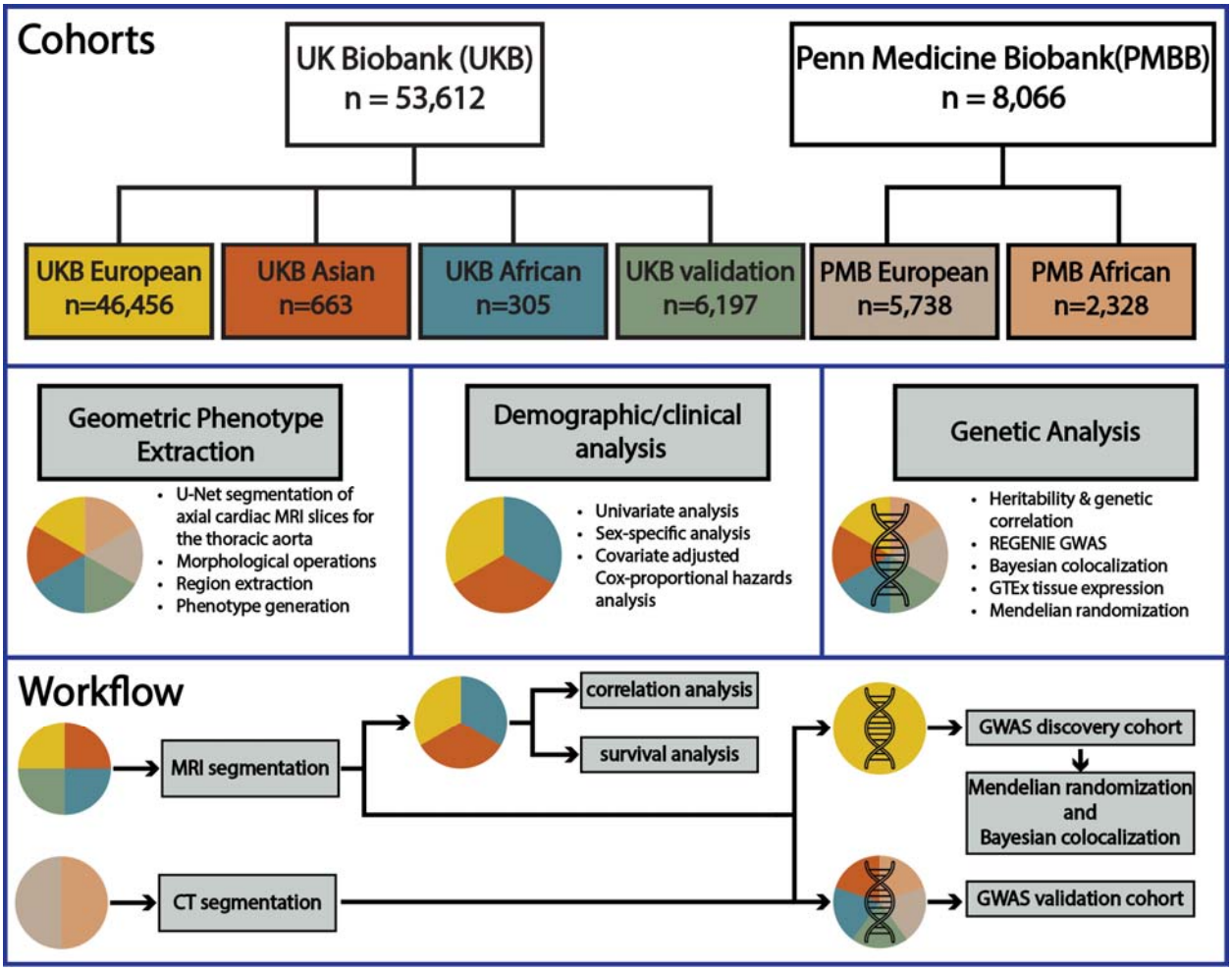


Figure 1. Overview of **(1)** study cohorts, **(2)** image segmentation, demographic and clinical correlation analysis, and genetic analysis, and **(3)** analysis workflow.

2. Results

As detailed in Methods, we automatically extracted 51 unique 3-D AGPS (**Fig. 2A; Supplemental Fig. 1**) from 49,282 UKB participants with CMR studies via U-Net segmentation, and various morphological operations to identify anatomical segments and compute overall and segment-specific AGPs (**Figs. 2A and 2B**). **Figure 2C** shows the correlation matrix of various 3D-AGPs. There were four well-defined AGP clusters, which related to: aortic arch phenotypes (cluster 1), aortic curvature (cluster 2), aortic diameter (cluster 3) and descending aortic centerline length (cluster 4). Metrics of descending aortic length (cluster 4) segregated separately from overall thoracic aortic length, which clustered with metrics of aortic arch 3D-geometry and overall thoracic aorta tortuosity (cluster 1).

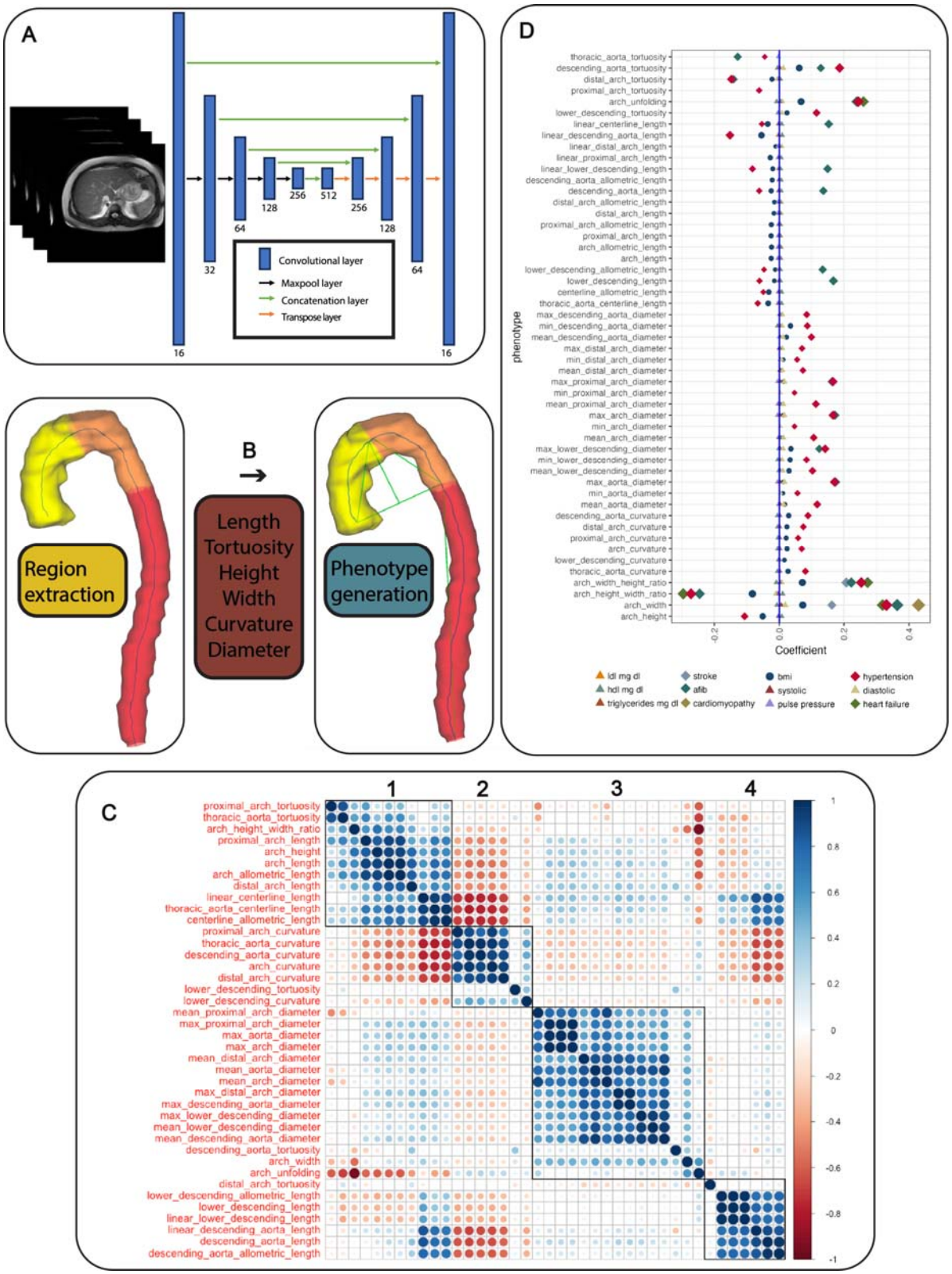

A
16
32
64
128
256
512
256
128
64
16
Convolutional layer
Maxpool layer
Concatenation layer
Transpose layer
B
Region extraction
Length
Tortuosity
Height
Width
Curvature
Diameter
Phenotype generation
D
phenotype
Coefficient
-0.2
0.0
0.2
0.4
ldl mg dl
hdl mg dl
triglycerides mg dl
stroke
afib
cardiomyopathy
bmi
systolic
pulse pressure
hypertension
diastolic
heart failure
C
1
2
3
4
proximal_arch_tortuosity
thoracic_aorta_tortuosity
arch_height_width_ratio
proximal_arch_length
arch_height
arch_length
arch_allometric_length
distal_arch_length
linear_centerline_length
thoracic_aorta_centerline_length
centerline_allometric_length
proximal_arch_curvature
thoracic_aorta_curvature
descending_aorta_curvature
arch_curvature
distal_arch_curvature
lower_descending_tortuosity
lower_descending_curvature
mean_proximal_arch_diameter
max_proximal_arch_diameter
max_aorta_diameter
max_arch_diameter
mean_distal_arch_diameter
mean_aorta_diameter
mean_arch_diameter
max_distal_arch_diameter
max_descending_aorta_diameter
max_lower_descending_diameter
mean_lower_descending_diameter
mean_descending_aorta_diameter
descending_aorta_tortuosity
arch_width
arch_unfolding
distal_arch_tortuosity
lower_descending_allometric_length
lower_descending_length
linear_lower_descending_length
linear_descending_aorta_length
descending_aorta_length
descending_aorta_allometric_length
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1

Figure 2. (A) Overview of the U-Net segmentation architecture. (B) Thoracic aorta region extraction and phenotype generation. (C) Correlation matrix of select aortic geometric phenotypes (AGPS), showing the 4 variable clusters (see text for details). (D) Age and sex-adjusted relationship between AGPs and various demographic parameters, laboratory parameters and prevalent clinical cardiovascular diagnoses.

Demographic and clinical correlates of aortic geometry

We examined the associations of AGPs with demographic factors, traditional risk factors for CVD, and prevalent cardiovascular clinical diagnoses using the UKB cohort. Univariable linear regression models identified significant associations at the Bonferroni-corrected significance level ($p<1.05e-5$). **Supplemental Figure 3** shows the standardized regression coefficient for each parameter, modeled as a predictor of the respective AGP. We found 38 AGPs to be age-related, with arch length and overall length having the strongest positive association, whereas arch curvature and overall curvature had strong negative associations (**Supplemental Fig. 3**, **Supplemental Table 1**). Fifteen AGPs were associated with BMI, with arch width and arch unfolding having the strongest positive associations. Sex was found to be an important determinant of multiple AGPs. Male sex was associated with greater aortic diameter as well as allometrically-indexed aortic length and was associated with lower curvature.

Given the impact of sex and age on aortic geometry, we re-analyzed associations of AGPs in sex and age-adjusted models, as well as sex-stratified analyses (**Fig. 2D and Supplemental Table 2**). Sex-stratified analyses are shown in **Supplemental Figures 4-5** and **Supplemental Tables 3-4**. In age and sex-adjusted analyses, multiple AGPs were associated with hypertension, heart failure and cardiomyopathy. Aortic arch width exhibited the strongest association with all 3

diagnoses. These 3 conditions were also associated with the arch unfolding and inversely associated with the arch height-to-width ratio. Overall, these results indicated an association between a wide, unfolded arch and the prevalence of these 3 conditions.

There was also a direct association between aortic diameter, and hypertension, which was seen for the overall thoracic aorta and various subsegments. There were positive associations between diameter measures and the prevalence of heart failure. BMI was associated with greater arch unfolding and descending aortic tortuosity, and less prominently associated with a greater diameter and a shorter overall aortic centerline length. Interestingly, a longer descending aorta was positively associated with a heart failure but negatively associated with hypertension.

Relationship to the risk of incident outcomes

We investigated the relationship between AGPs and incident cardiovascular events / *de novo* diagnoses among individuals free of the disease of interest at the time of imaging. Outcomes of interest included hypertension, cardiomyopathy, heart failure and atrial fibrillation. Standardized hazard ratios (HR) and 95% confidence intervals (CI) derived from Cox proportional-hazards models assessing the relationship between AGPs at baseline and the risk of incident outcomes are shown in **Supplemental Figure 6-9**, and **Supplemental Table 5**. Selected AGPs, along with representative Kaplan-Meier plots for each outcome are shown in **Figure 3**. All models are adjusted for classic cardiovascular risk factors (age, sex, systolic blood pressure, diastolic blood pressure, LDL-cholesterol, HDL-cholesterol, triglycerides, and BMI).

The risk of incident atrial fibrillation was associated with a greater arch unfolding index (HR=1.21; CI=1.11-1.33; $p=3.28x10^{-5}$), aortic arch width (HR=1.398; CI =1.29-1.50; $p=7.85x10^{-19}$), larger aortic diameter (HR=1.32; CI=1.21-1.44; $p=1.71x10^{-11}$)as well as greater thoracic

aortic centerline length (HR=1.20; CI=1.11-1.30; $p=2.47x10^{-6}$) and descending aortic length (HR=1.18; CI=1.09-1.27; $p=1.52x10^{-5}$). Similarly, the risk of incident heart failure and cardiomyopathy were associated with a greater arch unfolding index (Heart failure HR=1.22; CI=1.06-1.40; p=0.003), aortic arch width (Heart failure HR=1.27; CI=1.13-1.42; $p=2.20x10^{-5}$), larger aortic diameter (HR=1.25; CI=1.10-1.42; p=0.0003), greater descending aortic centerline length (HR=1.11; CI=1.00-1.24; p=0.04). Finally, the risk of hypertension was directly related to a greater arch width (HR=1.13; CI=1.06-1.20; p=0.0001), greater arch unfolding (HR=1.09; CI=1.01-1.17; p=0.01), and greater aortic diameter (HR=1.13; CI=1.05-1.21; p=0.0002).

These results were highly replicable in the sex specific models summarized in **Supplemental Tables 6-7** respectively (Pearson correlation between male and female analyses=0.77), with modest differences occurring for the outcome of cardiomyopathy (absolute mean difference=0.18).

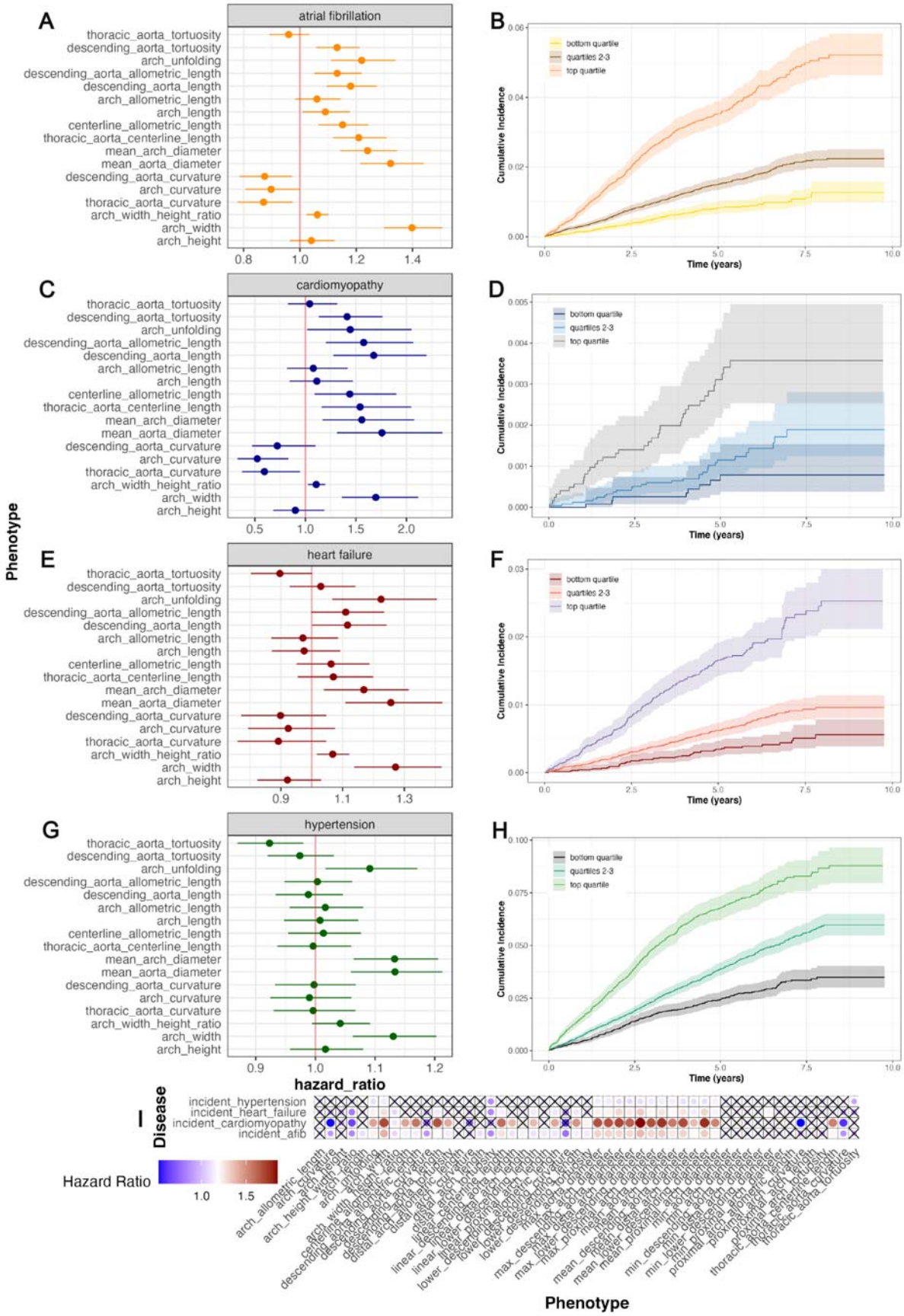

A
atrial fibrillation
B
C
cardiomyopathy
D
E
heart failure
F
G
hypertension
H
thoracic_aorta_tortuosity
descending_aorta_tortuosity
arch_unfolding
descending_aorta_allometric_length
descending_aorta_length
arch_allometric_length
arch_length
centerline_allometric_length
thoracic_aorta_centerline_length
mean_arch_diameter
mean_aorta_diameter
descending_aorta_curvature
arch_curvature
thoracic_aorta_curvature
arch_width_height_ratio
arch_width
arch_height
Phenotype
hazard_ratio
bottom quartile
quartiles 2-3
top quartile
Cumulative Incidence
Time (years)
I
Disease
incident_hypertension
incident_heart_failure
incident_cardiomyopathy
incident_afib
Hazard Ratio
1.0
1.5
Phenotype

Figure 3. Standardized hazard ratios with 95% confidence intervals for aortic geometric phenotypes and representative stratified cumulative incidence functions, for (A-B) atrial fibrillation, (C-D) cardiomyopathy, (E-F) heart failure, and (G-H) hypertension. Cumulative incidence function graphs (B, D, F, H) were stratified using quartiles of arch width. (I) Visualization of hazard ratios for all AGPs.

SNP-based heritability estimates of AGPs

Figure 4A shows estimates of the SNP-based heritability (h^2) for 51 AGPs in the UKB European ancestry sub cohort (n=46,456). The h^2 estimates, standard errors, and p-values are provided in **Supplemental Table 8**. After Bonferroni correction ($p<1x10^{-4}$), 48 AGPs were significantly heritable, with an average heritability of 0.19 (s.e.=0.10). The AGP exhibiting the highest levels of heritability was notably the diameter of the lower descending aorta showed the largest heritability (h^2=0.41). Thoracic aortic length (allometrically indexed and non-indexed) showed h^2 of 0.23 and 0.27, respectively (**Supplemental Table 8**). Aortic arch width heritability was 0.19 (s.e.=0.01).

Genetic architecture of aortic geometry

We performed GWAS on all 51 AGPs in the UKB European ancestry sub cohort (n=46,456). **Figure 4B** shows a summary of the findings. All GWAS are available in **Supplemental Datasets** (will be made available upon publication). LD and genomic inflation (λ) were tested to verify the robustness of each GWAS (LD mean=1.01±0.01, λ mean=1.10±0.06). LD scores, genomic inflation and GWAS-based heritability is shown in **Supplemental Table 9**. After adjusting for genome-wide significance ($P<5x10^{-8}$) we identified 237 regions (LD $r^2<0.1$) containing 2.6% of

the entire genome. More specifically, among aortic length GWAS results, we identified 41 regions associated with aortic length (137 loci), 27 aortic tortuosity regions (28 loci), 31 aortic curvature regions (41 loci), 163 aortic diameter regions (748 loci), and 16 aortic arch regions (18 loci).

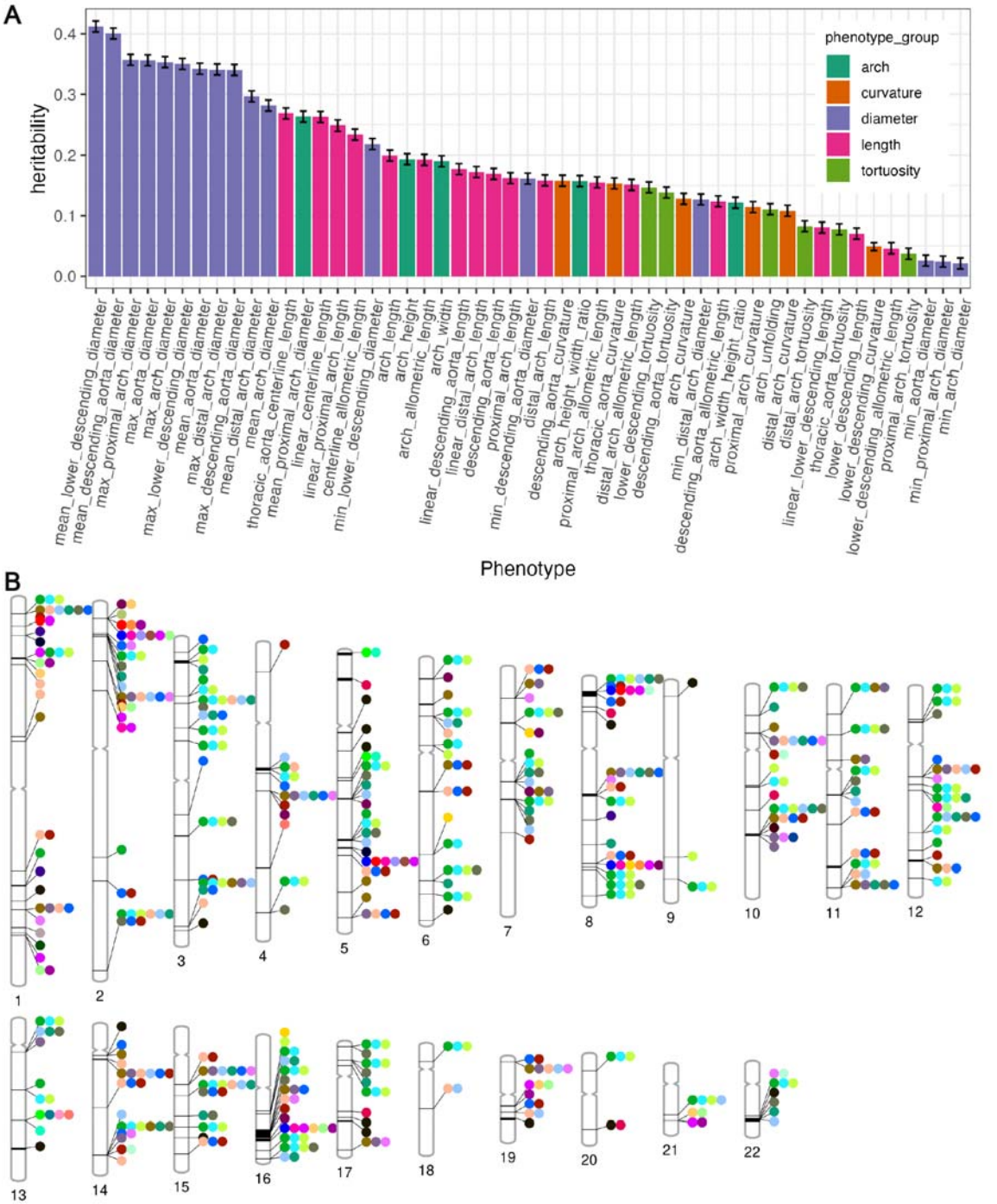

A
heritability
0.0
0.1
0.2
0.3
0.4
phenotype_group
arch
curvature
diameter
length
tortuosity
mean_lower_descending_diameter
mean_descending_aorta_diameter
max_proximal_arch_diameter
max_aorta_diameter
max_arch_diameter
max_lower_descending_diameter
mean_aorta_diameter
max_distal_arch_diameter
max_descending_aorta_diameter
mean_distal_arch_diameter
mean_arch_diameter
thoracic_aorta_centerline_length
mean_proximal_arch_diameter
linear_centerline_length
linear_proximal_arch_diameter
centerline_allometric_length
min_lower_descending_diameter
arch_length
arch_height
arch_allometric_length
arch_width
linear_descending_aorta_length
linear_distal_arch_length
descending_aorta_length
proximal_arch_length
min_descending_aorta_length
distal_arch_length
descending_aorta_diameter
arch_height_width_ratio
proximal_arch_allometric_length
thoracic_aorta_curvature
distal_arch_allometric_length
lower_descending_length
descending_aorta_tortuosity
min_distal_arch_diameter
arch_curvature
descending_aorta_allometric_length
arch_width_height_ratio
proximal_arch_curvature
arch_unfolding
distal_arch_curvature
linear_lower_descending_length
distal_arch_tortuosity
thoracic_aorta_tortuosity
lower_descending_length
lower_descending_curvature
lower_descending_allometric_length
proximal_arch_length
min_aorta_tortuosity
min_proximal_arch_diameter
min_arch_diameter
Phenotype
B
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
arch_allometric_length
arch_curvature
arch_height
arch_height_width_ratio
arch_length
arch_unfolding
arch_width
proximal_arch_allometric_length
proximal_arch_curvature
proximal_arch_length
desc_aorta_curvature
desc_aorta_tortuosity
distal_arch_allometric_length
distal_arch_curvature
distal_arch_length
linear_proximal_arch_length
linear_distal_arch_length
lower_desc_curvature
lower_desc_tortuosity
max_aorta_diameter
max_arch_diameter
max_proximal_arch_diameter
max_desc_aorta_diameter
max_distal_arch_diameter
max_lower_desc_diameter
mean_aorta_diameter
mean_arch_diameter
mean_proximal_arch_diameter
mean_desc_aorta_diameter
mean_distal_arch_diameter
mean_lower_desc_diameter
min_desc_aorta_diameter
min_distal_arch_diameter
min_lower_desc_diameter
centerline_allometric_length
thoracic_aorta_curvature
thoracic_aorta_centerline_length
linear_centerline_length
annulus_thorax_tortuosity

Figure 4. (A) SNP-based heritability of all significant (p<1e-4) aortic geometric phenotypes. (B) visualization of significant genetic loci.

Validation of GWAS results

In the UKB validation cohort (**Supplemental Table S11**), we replicated 9 aortic length loci at Bonferroni significance (p<0.05/41); 2 tortuosity loci (p<0.05/27); 3 aortic curvature loci (p<0.05/31); 32 aortic diameter loci (p<0.05/163); 5 arch loci (p<0.05<16). After overlap reduction we verified 36 loci (p<0.05/237).

In our PMBB external validation cohort (**Supplemental Table 12**) among participants with European ancestry we replicated 15 aortic length loci in the unadjusted length GWAS (p<0.05); 1 aortic tortuosity loci; 5 aortic curvature associations at nominal significance; 21 aortic diameter associations at Bonferroni significance; 7 arch loci at nominal significance. After overlap reduction we replicated 22 loci at Bonferroni significance.

Assessment in participants of different ancestries

We assessed whether the genetic associations demonstrated among participants of European ancestry in the UKB were present among participants with Asian (n=663) and African (n=305) genetic ancestry in the UKB, as well as participants in the PMBB with African ancestry (n=2,328).

In our PMBB African cohort (**Supplemental Table 13**), we replicated 2 aortic length loci at Bonferroni significance; 3 tortuosity loci at nominal significance; 3 curvature loci at Bonferroni significance; 9 diameter loci at Bonferroni significance; 1 arch loci at Bonferroni significance. After overlap reduction we replicated 21 loci at Bonferroni significance. In the

UKB African cohort (**Supplemental Table 14**), we replicated 21 aortic length loci at nominal significance; 6 tortuosity loci at nominal significance; 1 curvature loci at Bonferroni significance; 3 diameter loci at Bonferroni significance; 2 arch loci at Bonferroni significance. After overlap reduction we verified 3 loci at Bonferroni significance.

In the UKB Asian cohort (**Supplemental Table 15**), we replicated 1 aortic length loci at the Bonferroni significance; 3 tortuosity loci at nominal significance ($P<0.05$); 1 curvature loci at Bonferroni significance; 5 diameter loci at Bonferroni significance; 3 arch loci at nominal significance. After overlap reduction we replicated 39 loci at Bonferroni significance.

Concordance of results with previous GWAS findings

To assess the concordance of our results with previous findings from GWAS, we queried independent SNPs among AGPs to previously reported GWAS summary data (NHGRI-EBI GWAS Catalog). We identified shared variants between our measurements for ascending and descending thoracic aortic diameter with previous in-plane measurements of the ascending and descending thoracic aorta[3]. We highlight genetic concordance between thoracic aortic length (arch length, and aortic arch height) and aortic root diameter[12]. Furthermore, we identified shared variants between descending aortic diameter and left and right ventricular end-systolic and end-diastolic volume, whereas the diameter of the aortic arch shared variants with left ventricular wall thickness[8,13]. Specifically, the genes, *SH2B3*, *ATXN2*, *BAG3*, and *CKDN1A* were identified in GWAS of the aorta and the left ventricle. Refer to **Supplemental Table 16** for a complete list of concordant results.

Genetic correlation between AGPs and disease traits

We used our GWAS summary statistics and disease publicly available GWAS to assess the genetic correlation across AGPs and disease traits (**Supplemental Table 17, Supplemental Fig. 10**). Notably, genetically determined arch width was significantly positively correlated with genetic risk of stroke (p=0.0006), heart failure ($p=1.4x10^{-8}$), atrial fibrillation ($p=3.24x10^{-11)}$, hypertension ($p=1.16x10^{-11}$), and non-ischemic cardiomyopathy (p=0.0002). Genetically determined aortic length was positively associated with atrial fibrillation (p=0.007) and negatively associated with CKD (p=0.006). Genetically determined arch unfolding was negatively associated with risk of stroke (p=0.005), atrial fibrillation ($p=3.64x10^{-11}$), heart failure ($p = 6.33x10^{-11}$) and hypertension ($p=4.63x10^{-11}$). Particularly, genetically determined distal arch tortuosity was negatively correlated with cardioembolic stroke (p=0.005). Lastly, several genetically predicted aortic curvature phenotypes were negatively correlated with genetically determined risk for atrial fibrillation.

Two-sample Mendelian randomization

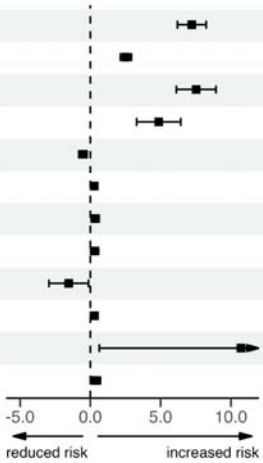
We estimated the potentially causal relationships between AGPs of interest and 14 cardioembolic disease outcomes using a two-sample MR approach. Using AGP GWAS' from the UKB as exposures, MR identified significant associations between genetically determined lower descending curvature and an increased risk of small vessel stroke ($p=2.59x10^{-43}$), chronic kidney disease ($p=1.97x10^{-37}$), large artery stroke ($p=1.20x10^{-25}$) and cardioembolic stroke ($p=1.20x10^{-9}$) (**Supplemental Table 18**). MR also highlighted a significant association between longer linear ascending aortic length and a reduced risk of large artery stroke (p=0.0005). At nominal significance ($p<0.05$), we highlight associations of genetically predicted aorta diameter and an

increased risk with hypertension, atrial fibrillation, and heart failure, and genetically predicted arch width with heart failure.

For these analyses, we obtained results that were similar in significance and directions of effect when applying weighted median and Egger methods. Analyses indicated adequate instrument strength (mean F-statistic>10).

A

Exposure	Outcome	Cohort	Pval	q	Estimate (95% CI)
Lower descending curvature	SVS	MEGASTROKE	<0.001	<0.001	7.22 (6.19 to 8.24)
Lower descending curvature	CKD	CKDGen	<0.001	<0.001	2.53 (2.14 to 2.92)
Lower descending curvature	LAS	MEGASTROKE	<0.001	<0.001	7.52 (6.11 to 8.93)
Lower descending curvature	CES	MEGASTROKE	<0.001	<0.001	4.86 (3.30 to 6.43)
Linear ascending arch length	LAS	MEGASTROKE	<0.001	0.015	-0.53 (-0.83 to -0.23)
Max aorta diameter	HYPTENS	FinnGen	0.006	0.150	0.28 (0.08 to 0.48)
Max aorta diameter	AF	FinnGen	0.015	0.322	0.34 (0.07 to 0.61)
Arch width	HF	HERMES	0.032	0.498	0.30 (0.03 to 0.58)
Max descending arch	LAS	MEGASTROKE	0.033	0.498	-1.54 (-2.95 to -0.13)
Arch width	HF	FinnGen	0.035	0.498	0.28 (0.02 to 0.54)
Descending aorta curvature	HF	FinnGen	0.037	0.498	10.72 (0.63 to 20.80)
Lower descending curvature	HF	HERMES	0.046	0.568	0.36 (0.01 to 0.71)



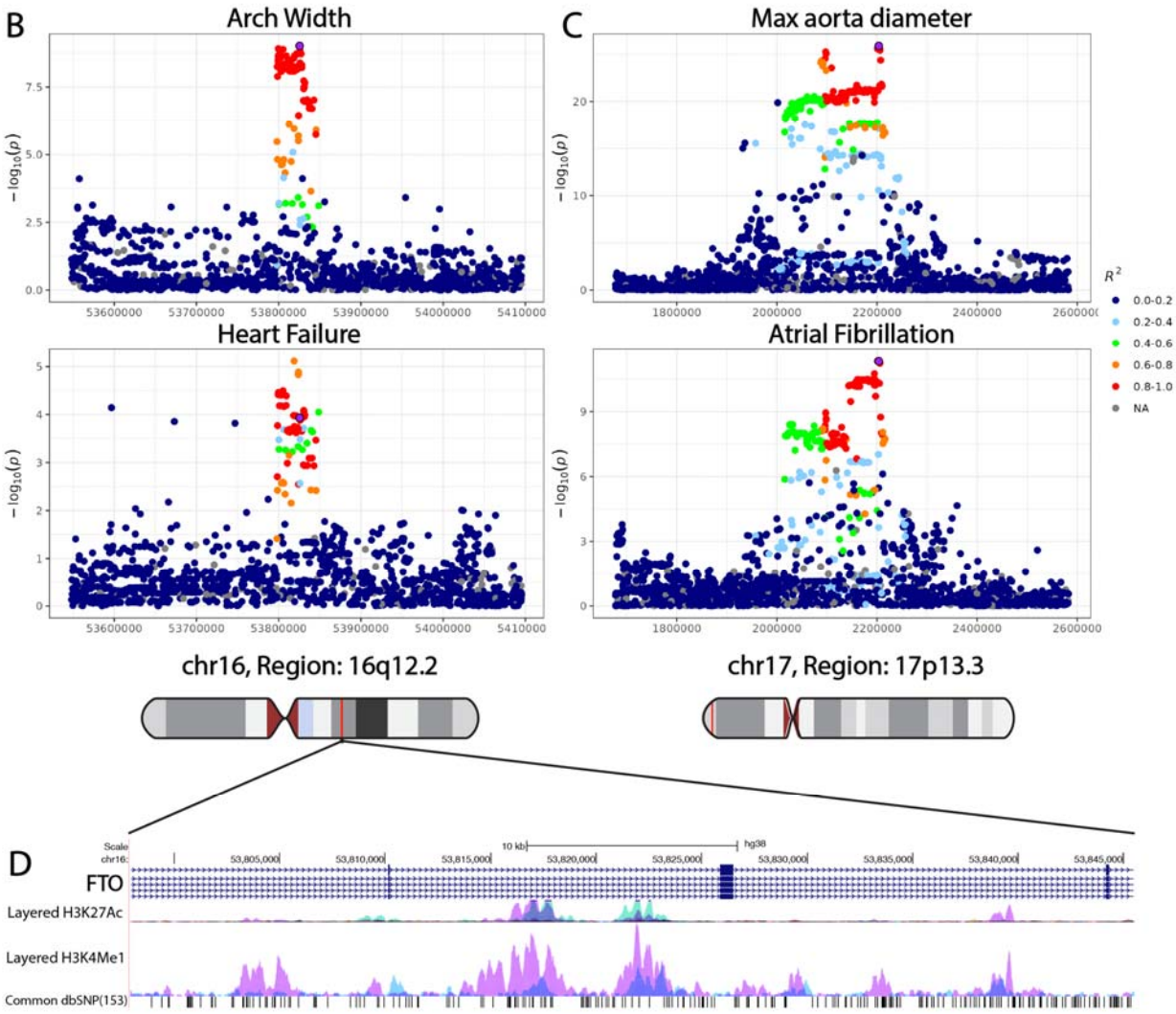


Figure 5. (A) Select AGPs and their causal link to small vessel stroke (SVS), chronic kidney disease (CKD), large artery stroke (LAS), cardioembolic stroke (CES), hypertension (HYPTENS), atrial fibrillation (AF), and heart failure (HF); (B) colocalization plot between arch width (region: 16q12.2) and heart failure showing Bayesian colocalization (PPH4 = 0.85); (C) colocalization plot between maximum aortic diameter (region: 17p13.3) and atrial fibrillation with Bayesian colocalization (PPH4 = 0.97); (D) FTO gene region localized on arch width loci with modification tracks (H3K4Me3 and H3K27Ac).

Bayesian colocalization

Using AGP GWAS results and cardiovascular disease outcomes (heart failure, hypertension, atrial fibrillation, cardiomyopathy, aortic stenosis, coronary artery disease, chronic kidney disease (CKD), vascular dementia, Alzheimer disease, stroke), we performed Bayesian colocalization analyses to identify shared causal variants between each AGP and outcome. Colocalization analyses identified 28 AGPs to have shared causal variants (PPH4 > 0.75) with hypertension, including arch width, unfolding of the aortic arch, tortuosity of the aorta, as well as aortic diameter. 12 AGPs were colocalized with heart failure, including several phenotypes of the arch (arch width, arch width-height ratio, arch unfolding), as well as aortic diameter, whereas 17 AGPs were colocalized with nonischemic cardiomyopathy. 16 AGPs colocalized with atrial fibrillation including centerline length of the aorta, arch width, aortic tortuosity, and more specifically tortuosity of the ascending arch. There was evidence of colocalization across 26 AGPs across all categories of stroke, and several AGPs being colocalized to specific subtypes of stroke including, large artery stroke, any ischemic stroke, and cardioembolic stroke

(**Supplementary Table 19**). More specifically, one genomic locus of arch width (region: 16q12.2) was colocalized (PPH4 > 0.75) with heart failure, atrial fibrillation, and hypertension (**Fig. 5.B, Supplemental Fig. 11, Supplemental Table 20**).

Gene, gene-set analysis, and drug-gene interactions

We identified candidate genes and gene-set associations for their relationship with aortic geometric structures using MAGMA and identified expression quantitative trait loci (eQTL) in aortic tissue using GTEx v8. Additionally, we queried our GWAS results to identify any significant exonic variants using FUMA. After Bonferroni correction(P<1.42e-9) MAGMA identified 125 significant gene-level associations (**Supplemental Table 21**), whereas FUMA identified 34 genes with significant exonic SNP variations (**Supplemental Table 22**). In the tissue specific gene-set analysis, we identified enriched gene-expression($P<5x10^{-8}$) among aortic artery, coronary artery, tibial artery, as well as the uterus, cervix, and fallopian tube. Some genes were identified across multiple AGPs, as well as being localized to specific subsegments of the aorta (**Fig. 6A**). *AMH*, *DOT1L*, *LSM2*, *PBX2*, *SF3A2*, and *PLEKHJ1* were identified primarily in AGPs of the descending aorta segments. Cellular communication network factor 3 (NOV) was identified with aortic length, arch width, and aortic diameter and has previously been associated with hypertension[14]. Fibrillin-2 (*FBN2*) was associated primarily with AGP GWAS related to the structure of the aortic arch (**Figs. 6A-B**), with two variants (rs530936609 and rs73346354) overlapping enhancer histone modification tracks. *FBN2* has previously been associated with CVD, coronary artery disease, as well as calcium-channel blocker and beta-blocker use[15-18].

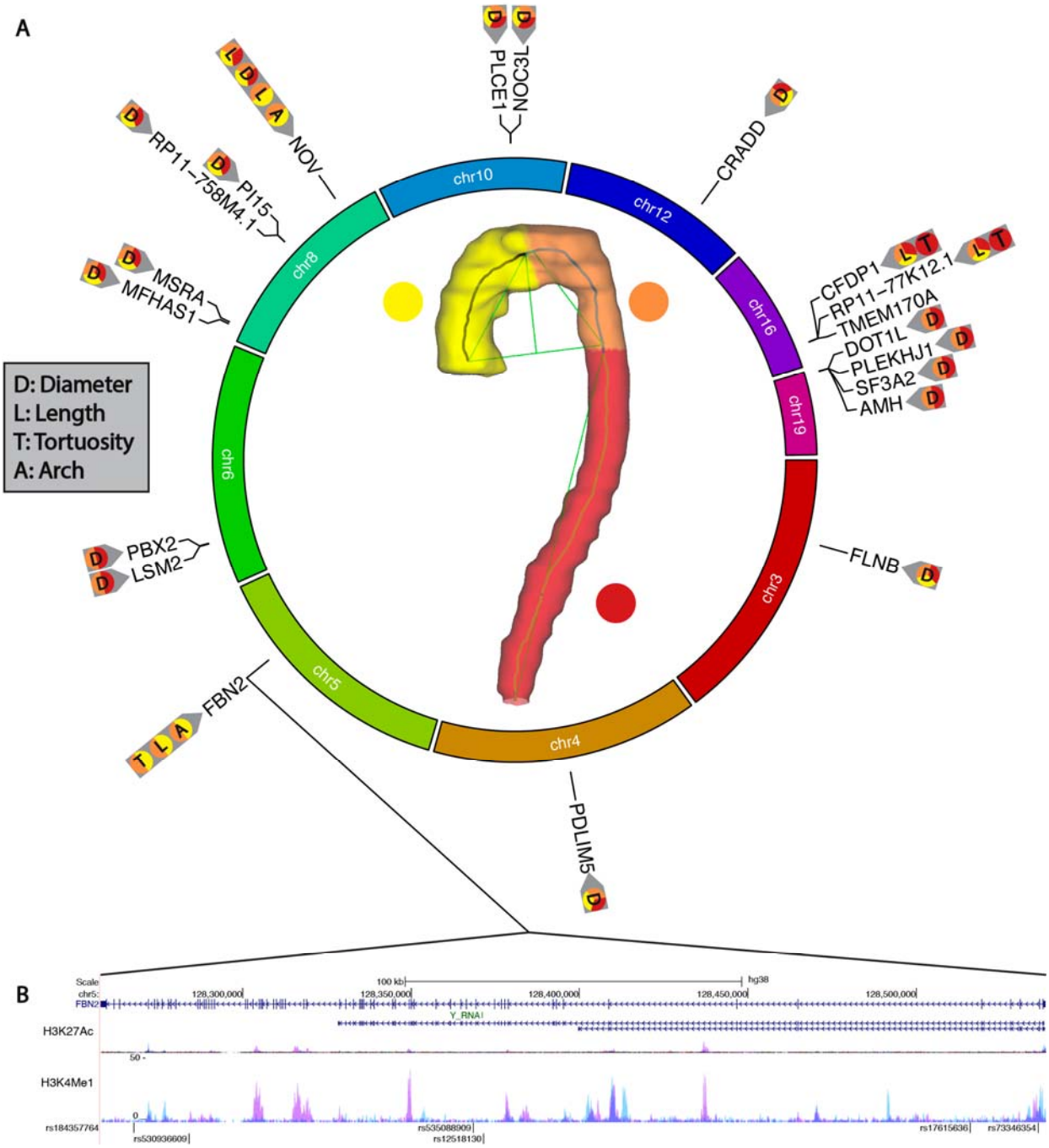


Figure 6. (A) Top 20 most common gene-level associations (P<1.28x10^{-9}) mapped to their corresponding genome location and their corresponding AGP. (B) FBN2 genetic locus, with

corresponding histone modification tracks (H3K4Me3 and H3K27Ac) and corresponding SNP variants.

Identification of candidate genes for drug target prioritization

We queried our results to identify AGP-associated genes that would be potential targets for therapeutic drugs using the drug-gene interaction database (DGIB)[19]. Of the 125 genes identified in our analyses, 11 genes were targets for 39 cardiovascular and blood pressure drugs (**Supplemental Table 23**). Ten drugs were calcium channel blockers (known antihypertensive agents) which targeted 4 genes (*ATXN2*, *UNIPROT:Q08289*, *CACNB2*, *EHMT2*). Four drugs were angiotensin converting enzyme (ACE) inhibitors, which are effective for the treatment of several cardiovascular conditions (including hypertension, heart failure and chronic kidney disease), which targeted 3 genes (*DOT1L, EHMT2, HSPA4*). Four alkaloids and 1 alpha-adrenergic agonist targeted 1 gene, namely *EMHT2*.

3. Discussion

This study leveraged aortic geometry imaging data and genetics from two large independent cohorts, the UKB and the PMBB, to comprehensively characterize AGPs, their underlying genetic architecture, and their putative causal relationships with disease outcomes. We quantified the clinical and epidemiologic correlates, prognostic associations, and genetic architecture of aortic 3D-geometry across multiple relevant phenotype categories. Using a GWAS discovery cohort of 46,456 participants in the UK Biobank, we validated our results on three separate validation cohorts of the UKB and two external cohorts from the PMBB. UKB validation results replicated 36 unique genomic loci associated with AGPs. This is the first study to examine the

genetic architecture of aortic 3D-geometric phenotypes, including aortic length, tortuosity, curvature, as well as arch-specific geometric indices, namely, arch width and height, as well as aortic diameter across the entire thoracic aorta. We report the first GWAS investigating 3D-AGPs of length, curvature, and tortuosity, which identified 237 genetic loci for these novel AGPs.

In observational analyses, we identified various AGPs that were associated with prevalent and new-onset hypertension, atrial fibrillation, cardiomyopathy, and heart failure. Specifically, we identified independent associations between the 3-dimensional unfolding of the aortic arch (characterized by arch widening and elongation) with incident cardiomyopathy and heart failure. This is likely related to the impact of these geometric changes on pulsatile left ventricular afterload and ventricular-arterial interactions[20]. Furthermore, arch width was associated with an increased risk of hypertension and atrial fibrillation.

Although the relationships between aortic stiffness and geometry as determinants of pulsatile hemodynamics are intertwined[1,20], only the former has been widely studied as a mediator of CVD and target organ damage. Studies regarding biologic and genetic determinants of aortic geometry have been limited to cross-sectional diameter, whereas 3D-aortic geometry has been assessed in only a few relatively small studies due to its complexity and to the scarcity of available data. Notably, previous smaller cross-sectional observational studies have previously highlighted links between aging and 3D-aortic geometry (including longer and wider aortic arch) and between some aortic 3D-phenotypes and the presence of type-B dissection[21]. Similarly, a prior study among 100 adults free of overt CVD demonstrated a relationship between age-related unfolding of the aortic arch and decreased ascending aortic distensibility, increased aortic arch stiffness (assessed via pulse wave velocity) and increased central blood pressures[20].

We demonstrate, for the first time, that 3D-geometric patterns reflective of structural aortic degeneration characterized by arch unfolding and descending aortic lengthening, are predictive for new-onset hypertension, heart failure, cardiomyopathy, and atrial fibrillation. This is consistent with the increasingly recognized role of the aorta as a key determinant of left ventricular afterload (via its impact on premature wave reflections and pulsatile characteristic impedance[22,23]. Further research is required to assess whether therapeutically targeting the hemodynamic impact of aortic structural degeneration can prevent these cardiovascular diseases. Moreover, we identify 3D-structural aortic degeneration itself as a quantifiable phenotypic domain that can be readily assessed with non-invasive imaging, and that has important clinical and epidemiologic implications.

Despite the recognized importance of the aorta, data regarding the molecular determinants of aortic structural degeneration are scarce[1] and derived mostly from animal studies or non-genetic observational studies. Genetic determinants of aortic geometry can provide important clues regarding its pathogenesis. We identified several genes that were significantly associated with multiple AGPs across several subsegments of the aorta (**Fig. 6**). NOV was identified across all AGP categories, as well as several subsegments of the aorta. Furthermore, NOV has been linked to increased inflammation, as well as degradation of cartilage[24]. Waldman *et al.* found that inhibition of NOV expression restores the function of anti-inflammatory, thermogenic, and mitochondrial gene expression, resulting in improved vascular function in mice[25]. FBN2 is a protein-encoder gene responsible for the production of fibrillin-2 and is essential for maintaining elastin deposits and the elastic fiber system.[26] FBN2 was primarily associated with AGP structures in the aortic arch. It has also been documented that FBN2 loss of function mutations result in a reduction of fibrillin-2 and disorganized smooth muscle cells in

epithelial tube formation, as well as a shortened tracheal tube length[27]. To our knowledge, FBN2 has not previously been identified in aortic GWAS results, suggesting that it is primarily associated with 3-dimensional geometric properties of the aorta.

Our work showed that AGPs had moderate to high heritability, with h^2 ranging from 23%, for length AGPs (namely thoracic aortic length) to 41% for diameter phenotypes (namely for lower descending aortas). Heritabilities of aortic structure have been previously reported in the literature, which are comparable to our findings[28]. Our GWAS identified novel susceptibility loci for AGPs. Previous GWAS have been conducted only on ascending aortic area[29]. Observational studies have also previously established numerous links between aortic imaging-derived phenotypes and clinical outcomes; however, these associations are prone to residual confounding and reverse causation.

Our identification of genetic instruments for aortic 3D-geometry allowed us to perform Mendelian randomization analyses to assess the putative causal effect of AGPs on disease outcomes. MR analyses suggested a detrimental effect of increased curvature of the descending aorta and increased propensity to SVS, CKD, LAS and CES, in addition to a potentially protective effect of increased linear ascending arch with LAS. These associations are highly relevant and biologically plausible. The susceptibility for small vessel stroke and chronic kidney disease is likely mediated by the impact of aortic wall degeneration on pressure and flow pulsatility in the arterial system. Increased pulsatility tends to damage target organs that exhibit low-resistance microvascular beds, such as the kidney and the brain[1]. The association between AGPs and the risk of large artery stroke may be related to its impact on the risk of atrial fibrillation (which leads to embolic strokes), on 3D flow patterns in the aorta (which may favor embolization to the brain) and to the impact of pulsatility on intracerebral plaque rupture.

However, these potential mechanisms are only hypothetical at present, and further research is required to assess the mechanistic and biologic basis behind these potentially causal associations.

Our study has several strengths. To our knowledge, this is the first study to combine an observational and genetic approach to comprehensively characterize 3-D aortic geometry and investigate the potential causal role of AGPs in disease incidence and progression. We also leverage data from two independent large biobanks and validate findings across multiple ancestry groups. Additionally, our study design leverages genetic variation as instrumental variables to assess causality between AGPs and disease outcomes. Compared with conventional epidemiological research, the MR approach is more robust to biases related to environmental confounding, reverse causation, and measurement error. Various sensitivity methods were employed to ensure the robustness of our assumptions in the MR model.

Our work also has limitations. Firstly, the imaging modality varied between the UKB (CMR) and the PMBB (CT). Additionally, the participants in the PMBB were undergoing imaging for medical evaluation, whereas UKB participants were randomly selected. Nevertheless, the replication of findings across these two cohorts with different inclusion and exclusion criteria suggests robust associations. Lastly, causal inferences made through our MR analyses should be interpreted cautiously, as MR measures the lifelong effect of genetic variants, therefore its estimates should not be directly translated to assume the effect of short-term clinical interventions on aortic function.

This is the first study to comprehensively assess aortic geometry using imaging data and a genomic-pronged approach to infer causality with disease outcomes. We provide evidence for potentially causal relationships between increased lower descending curvature and the risk of

CKD and stroke subtypes SVS, LAS and CES, and a potentially protective effect of increased linear ascending arch length and risk of LAS.

4. Methods

Data Sources

The UKB is composed of >500,000 participating individuals aged 37-73 years at the time of recruitment, who underwent various questionnaires, physical measurements, biological sampling (blood and urine), and genome sequencing across 22 assessment centers in the UK[30]. A subset of participants were invited to complete an additional examination that included magnetic resonance imaging of the heart[10].The PMBB is composed of 174,712 consenting patients of the Penn Medicine health network, with a subset of 44,000 participants with available genotyping data[11]. Additionally, all participants medical records including imaging results are de-identified and linked to their identifier.

Aortic segmentation and 3-dimensional geometric phenotyping

Data from a subset of individuals (n=49,282) enrolled in the UKB who underwent cardiac magnetic resonance imaging (CMR) were included for this discovery analysis[10]. A variation of the U-Net convolutional neural network (CNN) segmentation architecture was used to segment the aorta[31]. The model was trained on manually segmented images from CMR scans of 194 participants and validated on an independent subset of 39 MRI scans. The model used an encoder-decoder U-Net architecture with 2-dimensional convolutional layers of 16, 32, 64, 128, and 256 (**Fig. 2A**). The network was trained using the dice coefficient as the loss function and was optimized using the Adam optimizer. Geometric and intensity augmentations were performed randomly on the training data. The model had an average dice coefficient of 0.934

(0.013). Axial CMR images slices were processed individually before being combined to form the 3-dimensional aorta segmentation. The voxelized aorta segmentation was then post-processed using erosion and dilation morphological operations with a kernel size of 3 voxels to remove any holes in the segmentation. Mesh generation and centerline extraction were performed using the VMTK toolkit[32]. The centerline was then interpolated to have 100 points using linear interpolation.

Identification of aortic segments, namely, the arch, ascending arch, descending arch, descending aorta, and lower descending aorta was accomplished through exploiting properties of the vertical axis (z-axis) of the aorta centerline to identify **(1)** the apex of the aorta, and **(2)** the end of the aortic arch. The apex of the aortic arch was identified as the maximum centerline point on the vertical axis. The aortic arch was extracted through identifying the centerline point on the descending aorta with smallest Euclidean distance on the vertical axis with the aortic root. Centerline length was computed as the summation of distances between consecutive points. Minimal linear length was computed as the distance between the first and last point within each aortic segment. Tortuosity was computed as the ratio of centerline length to minimal linear length minus 1. Aortic curvature was calculated using the VMTK toolkit[32]. The aortic radius was computed across every point on the centerline by calculating the nearest distance from the centerline to the aorta mesh and doubled to compute short-axis aortic diameter. AGPs were computed similarly on the PMBB participants using computed tomography (CT) scans. The aorta was segmented using a previously developed algorithm[33]. To identify the thoracic aorta, the segmentation was cut at the T12 vertebral level.

In addition to cross-sectional diameter, centerline length and curvature, various key AGPs were computed from the overall aorta and from specific segments using the 3D-image

segmentation, as shown in **Figure 2B and Supplemental Figure 1**. We computed length, diameter and 3D-curvature and tortuosity metrics. Given that the arch is expected to be curved, standard curvature and tortuosity metrics do not characterize its expected geometry. Therefore, in order to characterize arch geometry, we used arch length, arch height, arch width, the dimensionless width-to-height ratio, and the arch unfolding index:

$$\textit{Unfolding Index} = -((\textit{Arch Length} / \textit{Arch Width})-1)$$

Allometric scaling

Given the normal expected relationship between body length and aortic length, we investigated the allometric relation between aortic length and body height using sex-adjusted models, as previously described.[34] The normal (physiologic) body height-aorta length relationship was assessed using a subset of UKB individuals (n=4,483) who were free of cardiovascular disease at the time of imaging who also met the following criteria: **(1)** BMI >18 and <25 kg/m^2, **(2)** low-density lipoprotein (LDL) cholesterol <130 mg/dL, **(3)** high-density lipoprotein (HDL) cholesterol >50 (female) or >40 (male) mg/dL, **(4)** triglycerides <200 mg/dL, **(5)** hemoglobin A1c <42 mmol/mol, and **(6)** never a smoker. Subjects whose centerline length measurements fell outside the 95% confidence interval (CI) were discarded to avoid erroneous stochasticity. The allometric relation is shown in **Supplemental Figure 2**. The allometric power that best represented the physiologic relationship between aortic length and body height was 0.44. Hence, subsequent analyses of length used both unindexed aortic length (cm) and allometrically normalized to height (cm) raised to the power of 0.44.

Survival analyses

We evaluated the relationship between AGPs and incident new-onset cardiovascular diseases to which aortic structural degeneration is hypothesized to contribute, namely, cardiomyopathy, heart failure, atrial fibrillation, hypertension, and cerebrovascular events, among participants who were free of the outcome of interest at the time of aortic imaging. Multivariable Cox proportional-hazards models were used to identify AGPs that were significant predictors of these cardiovascular diseases after adjustment for the following risk factors: (1) age, (2) sex, (3) BMI, (4) LDL cholesterol, (5) HDL cholesterol, (6) triglycerides, (7) systolic blood pressure, and (7) diastolic blood pressure. For an intuitive comparison of the multiple novel metrics of aortic 3D-geometry, all AGP variables were normalized to Z-scores. Prognostic performance was quantified using the inverse probability of censoring weighting (IPCW) estimation of the cumulative time-dependent ROC curve[35]. AUC values were reported for the 25th, 50th, and 75th quantile for the time to event for each disease. All analyses were repeated for sex-specific models.

Genome-wide association studies

Exploratory genetic analysis was performed on 46,456 UKB participants with European genetic ancestry who underwent CMR imaging. GWAS were performed on 51 AGPs using imputed autosomal genome data with the following quality control measures: (1) exclusion of subjects who had >10% genotype missingness, (2) exclusion of single nucleotide polymorphisms (SNPs) with a minor allele frequency (MAF) <1% and with minor allele counts <100, and (3) exclusion of SNPs that failed the Hardy-Weinberg test ($P < 10^{-15}$). SNPs with an imputation INFO score <0.8 were excluded from further analyses. GWAS analyses were performed using a two-step linear mixed effect model from REGENIE[36] and were adjusted for (1) age (at the time of CMR),

(2) sex, and (3) the first 10 genetic principal components. GWAS analyses were performed separately on three sub cohorts of UKB participants according to their self-reported genetic ancestry, namely (1) European (n=46,456), (2) Asian (n=663), and (3) African (n=305). We first validated our results using 6,197 additional UKB imaging studies, we then performed an additional external validation on 8,066 imaging results acquired from the PMBB[37]. Both validation studies were adjusted using the same covariates. SNP-based heritability estimates were calculated using GCTA[38] Haseman-Elston regression analysis for SNPs on both the entire UKB cohort and the UKB European ancestry cohort. LDSC heritability analysis was also performed using GWAS summary statistic data[39].

We used FUMA[40] (version 1.5.2) to define independent, lead and candidate variants, and to facilitate the functional annotation of GWAS results. First, independent significant SNPs were identified as SNPs below $P < 5x10^{-8}$ and not in high linkage disequilibrium (LD) with each other ($r^2<0.6$). Then, SNPs in LD with an independent SNP were retained as candidate variants, forming LD blocks. LD was derived using the 1000 Genomes project reference population. Next, LD blocks within 250 kb of each other were merged into one locus. Independent lead SNPs were then identified through a second clumping of the independent variants at LD $r^2<0.1$. Additionally, we used FUMA to compare the genome-wide lead SNPs identified in our AGPs GWAS against results from the National Human Genome Research Institute-European Bioinformatics Institute (NHGRI-EBI) GWAS catalog[41], a publicly available GWAS database.

LD score regression

To evaluate the extent of shared common variant genetic architectures between AGPs and disease outcomes, we computed the genetic correlations (r_g) between AGPs and disease outcome traits

using the LDSC software tool[39] (v1.0.1). The LD scores were computed from the 1000 Genomes European reference data and were incorporated into the LDSC software.

Downstream analyses to prioritize genes and tissues relevant to AGPs.

We also implemented the Multi-marker Analysis of GenoMic Annotation[42] (MAGMA) (v1.10) function within FUMA to conduct gene-based, gene-set, and gene-property analyses on the full GWAS summary statistics for all AGPs. We analyzed 18,987 protein coding genes to identify candidate genes and gene-set associations for their relationship with AGPs. Gene-property analyses were conducted to identify tissue specificity of each AGP, testing relationships between tissue-specific gene expression and phenotype-gene associations. This was also implemented using the FUMA *in-silico* pipeline, which incorporates expression quantitative-trait loci (eQTL) data for 30 tissues from Genotype-Tissue Expression (GTEx) v8.

Two-sample Mendelian randomization

We conducted two-sample MR to explore associations between genetically predicted aortic geometry phenotypes and the risk of target organ damage phenotypes. MR is an epidemiological approach that leverages genetic variation as instrumental variables to provide causal estimates of an exposure on an outcome of interest. There are 3 core assumptions to leveraging genetic instruments to conduct robust MR experiments: (1) the genetic instruments must be associated with the exposure of interest (the relevance assumption), (2) There are no confounders of the association between the genetic instrument(s) and outcomes of interest (the independence assumption), (3) the genetic instruments are only related to the outcome via the exposure (the exclusion restriction assumption). We selected independent genetic variants that are significantly

associated with AGPs at genome-wide level by performing LD clumping with a threshold of r^2=0.001 based on the 1000 Genomes Reference Panel using the European ancestry population as a reference. We estimated the strength of each instrument by calculating the F statistic, using the formula: formula $F = (\beta^2 / \text{standard error}^2)$[43]. An F statistic >10 was considered of high strength[44].

We used GWAS summary statistics for primary outcomes from the FinnGen (Version 10) to ensure there would be no sample overlap between exposure and outcome datasets. MetaGWAS were further used for stroke subtypes from the MEGASTROKE Consortium[45], for heart failure (HF) from the HERMES cohort[45,46], and for chronic kidney disease (CKD) and estimated glomerular filtration rate (eGFR) from the CKDGen cohort.

We used the inverse-variance weighted (IVW) as our main MR method, and the MR-Egger[47] and weighted median methods[48] were applied as sensitivity analyses, as they make different assumptions about the presence of invalid instruments and pleiotropy[49]. We estimated the I^2 statistic to detect heterogeneity among the MR estimates obtained from multiple genetic variants. We adjusted for multiple testing using the false discovery rate (FDR) method. Estimates are reported as odds ratios (ORs) and 95% CI for the putative effects of each exposure.

Bayesian colocalization analyses

We assessed genetic colocalization as sensitivity analyses to confirm that the exposure with a putative causal effect on outcomes (as indicated by MR) were regulated by the same causal variant[50]. The colocalization method calculates posterior probabilities (PP) for the competing hypotheses of H_0 (no causal variants, H_1 (causal variant for trait 1), H_2 (causal variant for trait 2), H_3 (distinct causal variants for traits 1 and 2), and H_4 (shared causal variants for traits 1 and 2,

which supports a causal association). If the PP that one SNP is shared for traits 1 and 2 was greater than 0.75, we regarded it as a colocalization signal, supporting a causal effect of the protein on the outcome (PPH4 > 0.75). Priors were calibrated to p1 (probability that the variant is associated with AGPs only), and p2 (probability that the variant is associated with outcomes only) at 10^{-4}, and p12 (probability that an arbitrary variant is associated with the trait for both AGPs and outcomes) at 10^{-5}. Bayesian colocalization was performed for each chromosome as well as for each genomic locus identified by FUMA.

Analyses linking GWAS results to drug targets, drug-gene interactions and disease

We queried our results to identify potential drug-gene interactions that would be using the drug-gene interaction database (DGIB)[19]. We used the CLUE Touchstone tool (https://clue.io/touchstone) to check for correlations between drug signatures and gene knockdown.

Data availability:

UK Biobank data is available to researchers following UKB application approval and IRB approval. Researchers can apply at https://www.ukbiobank.ac.uk/. Access to Penn Medicine BioBank data is provided to investigators at the University of Pennsylvania. External access can be provided through collaboration with an investigator at the university.

Code availability:

All code for aortic geometric phenotype (AGP) extraction can be accessed at https://github.com/cams2b/ukb_aorta/. The provided repository contains pretrained weights for U-Net segmentation as well as the required morphological pipeline for extraction AGPs.

Disclosures:

J.A.C. has recently consulted for Bayer, Sanifit, Fukuda-Denshi, Bristol Myers Squibb, JNJ, Edwards Life Sciences, Merck and the Galway-Mayo Institute of Technology. He received University of Pennsylvania research grants from the NIH, Fukuda-Denshi, Bristol Myers Squibb and Microsoft. He is named as an inventor in a University of Pennsylvania patent for the use of inorganic nitrates/nitrites in Heart Failure with Preserved Ejection Fraction. He has received research device loans from Atcor Medical, Fukuda-Denshi, Uscom, NDD Medical Technologies, Microsoft and MicroVision Medical. The remaining authors have nothing to disclose.

Acknowledgements

We acknowledge the Penn Medicine BioBank (PMBB) for providing data and thank the patient-participants of Penn Medicine who consented to participate in this research program. We would also like to thank the Penn Medicine BioBank team and Regeneron Genetics Center for providing genetic variant data for analysis. The PMBB is approved under IRB protocol# 813913 and supported by Perelman School of Medicine at University of Pennsylvania, a gift from the Smilow family, and the National Center for Advancing Translational Sciences of the National Institutes of Health under CTSA award number UL1TR001878.

Supplemental Note

Penn Medicine BioBank Banner Author List and Contribution Statements

PMBB Leadership Team
Daniel J. Rader, M.D., Marylyn D. Ritchie, Ph.D.
Contribution: All authors contributed to securing funding, study design and oversight. All authors reviewed the final version of the manuscript.

Patient Recruitment and Regulatory Oversight
JoEllen Weaver, Nawar Naseer, Ph.D., M.P.H., Giorgio Sirugo, M.D., P.h.D., Afiya Poindexter, Yi-An Ko, Ph.D., Kyle P. Nerz

Contributions: JW manages patient recruitment and regulatory oversight of study. NN manages participant engagement, assists with regulatory oversight, and researcher access. GS assists with researcher access. AP, YK, KPN perform recruitment and enrollment of study participants.

Lab Operations
JoEllen Weaver, Meghan Livingstone, Fred Vadivieso, Stephanie DerOhannessian, Teo Tran, Julia Stephanowski, Salma Santos, Ned Haubein, P.h.D., Joseph Dunn

Contribution: JW, ML, FV, SD conduct oversight of lab operations. ML, FV, AK, SD, TT, JS, SS perform sample processing. NH, JD are responsible for sample tracking and the laboratory information management system.

Clinical Informatics
Anurag Verma, Ph.D., Colleen Morse Kripke, M.S. DPT, MSA, Marjorie Risman, M.S., Renae Judy, B.S., Colin Wollack, M.S.

Contribution: All authors contributed to the development and validation of clinical phenotypes used to identify study subjects and (when applicable) controls.

Genome Informatics
Anurag Verma Ph.D., Shefali S. Verma, Ph.D., Scott Damrauer, M.D., Yuki Bradford, M.S., Scott Dudek, M.S., Theodore Drivas, M.D., Ph.D.

Contribution: AV, SSV, and SD are responsible for the analysis, design, and infrastructure needed to quality control genotype and exome data. YB performs the analysis. TD and AV provides variant and gene annotations and their functional interpretation of variants.

References

1. Chirinos, J.A., Segers, P., Hughes, T. & Townsend, R. Large-Artery Stiffness in Health and Disease: JACC State-of-the-Art Review. *Journal of the American College of Cardiology* **74**, 1237-1263 (2019).
2. Vasan, R.S., Song, R.J., Xanthakis, V. & Mitchell, G.F. Aortic Root Diameter and Arterial Stiffness: Conjoint Relations to the Incidence of Cardiovascular Disease in the Framingham Heart Study. *Hypertension* **78**, 1278-1286 (2021).
3. Pirruccello, J.P. *et al.* Deep learning enables genetic analysis of the human thoracic aorta. *Nature Genetics* **54**, 40-51 (2022).
4. Benjamins, J.W. *et al.* Genomic insights in ascending aortic size and distensibility. *eBioMedicine* **75**, 103783 (2022).
5. Bai, W. *et al.* A population-based phenome-wide association study of cardiac and aortic structure and function. *Nature Medicine* **26**, 1654-1662 (2020).
6. Bouke, P.A. *et al.* Aortic elongation part I: the normal aortic ageing process. *Heart* **104**, 1772 (2018).
7. Samuel, H. *et al.* Aortic elongation part II: the risk of acute type A aortic dissection. *Heart* **104**, 1778 (2018).
8. Aung, N. *et al.* Genome-wide association analysis reveals insights into the genetic architecture of right ventricular structure and function. *Nature Genetics* **54**, 783-791 (2022).
9. Gomes, B. *et al.* Genetic architecture of cardiac dynamic flow volumes. *Nature Genetics* (2023).
10. Littlejohns, T.J. *et al.* The UK Biobank imaging enhancement of 100,000 participants: rationale, data collection, management and future directions. *Nature Communications* **11**, 2624 (2020).
11. Verma, A. *et al.* The Penn Medicine BioBank: Towards a Genomics-Enabled Learning Healthcare System to Accelerate Precision Medicine in a Diverse Population. *Journal of Personalized Medicine* **12**, 1974 (2022).
12. Vasan, R.S. *et al.* Genetic Variants Associated With Cardiac Structure and Function: A Meta-analysis and Replication of Genome-wide Association Data. *JAMA* **302**, 168-178 (2009).
13. Tadros, R. *et al.* Shared genetic pathways contribute to risk of hypertrophic and dilated cardiomyopathies with opposite directions of effect. *Nature Genetics* **53**, 128-134 (2021).
14. Slavin, T.P., Feng, T., Schnell, A., Zhu, X. & Elston, R.C. Two-marker association tests yield new disease associations for coronary artery disease and hypertension. *Human Genetics* **130**, 725-733 (2011).
15. Kichaev, G. *et al.* Leveraging Polygenic Functional Enrichment to Improve GWAS Power. *The American Journal of Human Genetics* **104**, 65-75 (2019).
16. Aragam, K.G. *et al.* Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants. *Nature Genetics* **54**, 1803-1815 (2022).
17. Wu, Y. *et al.* Genome-wide association study of medication-use and associated disease in the UK Biobank. *Nature Communications* **10**, 1891 (2019).
18. Sakaue, S. *et al.* A cross-population atlas of genetic associations for 220 human phenotypes. *Nature Genetics* **53**, 1415-1424 (2021).

19. Freshour, S.L. *et al.* Integration of the Drug–Gene Interaction Database (DGIdb 4.0) with open crowdsource efforts. *Nucleic Acids Research* **49**, D1144-D1151 (2020).
20. Redheuil, A. *et al.* Age-Related Changes in Aortic Arch Geometry: Relationship With Proximal Aortic Function and Left Ventricular Mass and Remodeling. *Journal of the American College of Cardiology* **58**, 1262-1270 (2011).
21. Qiu, P. *et al.* Changes in aortic arch geometry and the risk for Stanford B dissection. *Journal of Thoracic Disease* **12**, 7193-7201 (2020).
22. Chirinos, J.A. & Segers, P. Noninvasive Evaluation of Left Ventricular Afterload. *Hypertension* **56**, 555-562 (2010).
23. Chirinos, J.A. & Segers, P. Noninvasive Evaluation of Left Ventricular Afterload. *Hypertension* **56**, 563-570 (2010).
24. Kuwahara, M. *et al.* CCN3 (NOV) Drives Degradative Changes in Aging Articular Cartilage. in *International Journal of Molecular Sciences* Vol. 21 (2020).
25. Waldman, M. *et al.* Silencing the Adipocytokine NOV: A Novel Approach to Reversing Oxidative Stress-Induced Cardiometabolic Dysfunction. in *Cells* Vol. 11 (2022).
26. Putnam, E.A., Zhang, H., Ramirez, F. & Milewicz, D.M. Fibrillin–2 (FBN2) mutations result in the Marfan–like disorder, congenital contractural arachnodactyly. *Nature Genetics* **11**, 456-458 (1995).
27. Wenguang, Y. *et al.* Fibrillin-2 is a key mediator of smooth muscle extracellular matrix homeostasis during mouse tracheal tubulogenesis. *European Respiratory Journal* **53**, 1800840 (2019).
28. Zhao, B. *et al.* Heart-brain connections: Phenotypic and genetic insights from magnetic resonance images. *Science* **380**, abn6598.
29. Francis, C.M. *et al.* Genome-wide associations of aortic distensibility suggest causality for aortic aneurysms and brain white matter hyperintensities. *Nature Communications* **13**, 4505 (2022).
30. Sudlow, C. *et al.* UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLOS Medicine* **12**, e1001779 (2015).
31. Ronneberger, O., Fischer, P. & Brox, T. *U-Net: Convolutional Networks for Biomedical Image Segmentation*, 234-241 (2015).
32. Antiga, L. *et al.* An image-based modeling framework for patient-specific computational hemodynamics. *Medical & Biological Engineering & Computing* **46**, 1097-1112 (2008).
33. Wasserthal, J. *et al.* TotalSegmentator: Robust Segmentation of 104 Anatomic Structures in CT Images. *Radiology: Artificial Intelligence* **5**, e230024 (2023).
34. Chirinos, J.A. *et al.* Left Ventricular Mass. *Hypertension* **56**, 91-98 (2010).
35. Robins, J.M. & Finkelstein, D.M. Correcting for Noncompliance and Dependent Censoring in an AIDS Clinical Trial with Inverse Probability of Censoring Weighted (IPCW) Log-Rank Tests. *Biometrics* **56**, 779-788 (2004).
36. Mbatchou, J. *et al.* Computationally efficient whole-genome regression for quantitative and binary traits. *Nature Genetics* **53**, 1097-1103 (2021).
37. Chae, A. *et al.* Strategies for Implementing Machine Learning Algorithms in the Clinical Practice of Radiology. *Radiology* **310**, e223170 (2024).
38. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics* **42**, 565-569 (2010).

39. Bulik-Sullivan, B.K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics* **47**, 291-295 (2015).
40. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nature Communications* **8**, 1826 (2017).
41. Sollis, E. *et al.* The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. *Nucleic Acids Research* **51**, D977-D985 (2023).
42. de Leeuw, C.A., Mooij, J.M., Heskes, T. & Posthuma, D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLOS Computational Biology* **11**, e1004219 (2015).
43. Bowden, J. *et al.* Improving the accuracy of two-sample summary-data Mendelian randomization: moving beyond the NOME assumption. *Int J Epidemiol* **48**, 728-742 (2019).
44. Burgess, S., Thompson, S.G. & Collaboration, C.C.G. Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol* **40**, 755-64 (2011).
45. Malik, R. *et al.* Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet* **50**, 524-537 (2018).
46. Lumbers, R.T. *et al.* The genomics of heart failure: design and rationale of the HERMES consortium. *ESC Heart Fail* **8**, 5531-5541 (2021).
47. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* **44**, 512-25 (2015).
48. Bowden, J., Davey Smith, G., Haycock, P.C. & Burgess, S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol* **40**, 304-14 (2016).
49. Davies, N.M., Holmes, M.V. & Davey Smith, G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* **362**, k601 (2018).
50. Giambartolomei, C. *et al.* Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet* **10**, e1004383 (2014).