

SCIENTIFIC REPORTS

OPEN

Genome-wide association study identifies loci for arterial stiffness index in 127,121 UK Biobank participants

Kenneth Fung¹, Julia Ramírez², Helen R. Warren^{2,3}, Nay Aung¹, Aaron M. Lee¹, Evan Tzanis^{2,3}, Steffen E. Petersen^{1,3} & Patricia B. Munroe^{2,3}

Arterial stiffness index (ASI) is a non-invasive measure of arterial stiffness using infra-red finger sensors (photoplethysmography). It is a well-suited measure for large populations as it is relatively inexpensive to perform, and data can be acquired within seconds. These features raise interest in using ASI as a tool to estimate cardiovascular disease risk as prior work demonstrates increased arterial stiffness is associated with elevated systolic blood pressure, and ASI is predictive of cardiovascular disease and mortality. We conducted genome-wide association studies (GWASs) for ASI in 127,121 UK Biobank participants of European-ancestry. Our primary analyses identified variants at four loci reaching genome-wide significance ($P < 5 \times 10^{-8}$): *TEX41* (rs1006923; $P = 5.3 \times 10^{-12}$), *FOXO1* (rs7331212; $P = 2.2 \times 10^{-11}$), *C1orf21* (rs1930290, $P = 1.1 \times 10^{-8}$) and *MRVI1* (rs10840457, $P = 3.4 \times 10^{-8}$). Gene-based testing revealed three significant genes, the most significant gene was *COL4A2* ($P = 1.41 \times 10^{-8}$) encoding type IV collagen. Other candidate genes at associated loci were also involved in smooth muscle tone regulation. Our findings provide new information for understanding the development of arterial stiffness.

Arterial stiffness measures have been reported as independent markers of vascular ageing^{1,2}, hypertension^{3,4}, cardiovascular disease (CVD)^{5,6} and mortality^{6,7}. Carotid-femoral pulse wave velocity (PWV) is the reference standard method for measuring arterial stiffness. However, an alternative and more convenient non-invasive method is to record the digital blood volume waveforms using infra-red finger sensors (photoplethysmography)⁸, where measurements can be recorded in a seated position rather than supine position required for carotid-femoral PWV recordings. This automatic technique is able to detect the waveform formed by the digital volume pulse, which is created by two components. First, pressure is transmitted from the left ventricle to the finger (direct component) whilst the second component is due to the transmission of pressure from the heart to the lower body (reflected component) via the aorta. The digital volume pulse can therefore be visualised as a dicrotic waveform, and the interval between the peaks of the direct and reflected components can be recorded to derive the arterial stiffness index (ASI), when divided into the individual's height. Higher ASI values reflect arterial walls with greater stiffness, due to the earlier arrival of wave reflection. ASI has been shown to have close agreement with other techniques measuring arterial stiffness, including PWV ($r = 0.58$, $P < 0.01$) and augmentation index ($r = 0.80$, $P < 0.01$)⁹.

Previous studies have evaluated the utility of ASI as a potential clinical marker of CVD, having shown good sensitivity (87%) and specificity (87%) when differentiating between older men with coronary artery disease (CAD) and younger men without CAD⁹. In a study of asymptomatic middle-aged patients, the mean ASI value was significantly higher ($P = 0.002$) in individuals with at least one significant (>50%) coronary stenotic plaque than those without a >50% stenotic lesion¹⁰. More recently, ASI has been shown to be an independent predictor

¹Advanced Cardiovascular Imaging, William Harvey Research Institute, Queen Mary University of London, Charterhouse Square, London, UK. ²Clinical Pharmacology, William Harvey Research Institute, Queen Mary University of London, Charterhouse Square, London, UK. ³NIHR Barts Biomedical Research Centre, Queen Mary University of London, Charterhouse Square, London, UK. Steffen E. Petersen and Patricia B. Munroe jointly supervised this work. Correspondence and requests for materials should be addressed to S.E.P. (email: s.e.petersen@qmul.ac.uk) or P.B.M. (email: p.b.munroe@qmul.ac.uk)

Locus	SNV	CHR	BP	EA	EAF	P	β	SE
C1orf21	rs1930290	1	184272584	T	0.553	1.1E-08	-0.023	0.004
TEX41	rs1006923	2	145775399	T	0.677	5.3E-12	0.029	0.004
MRVII	rs10840457	11	10675738	A	0.314	3.4E-08	-0.024	0.004
FOXO1	rs7331212	13	41185309	G	0.737	2.2E-11	0.030	0.005

Table 1. Summary of genome-wide significant loci associated with arterial stiffness index. Locus indicates the name of the gene in closest proximity to the most significant SNV. SNV = single-nucleotide variant; CHR = chromosome; BP = base pair position (Build 37); EA = effect allele; EAF = effect allele frequency; P = P-value (standard infinitesimal mixed model); β = effect-size estimates on an inverse-normal transformed scale. SE = standard error.

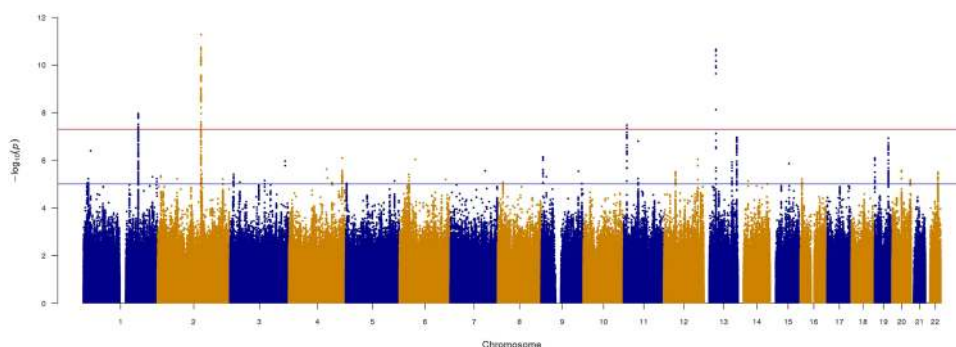


Figure 1. Manhattan plot for arterial stiffness index (ASI) in UK Biobank. The red line indicates the P-value threshold for genome-wide significance (5×10^{-8}) while the blue line indicates P-value threshold for suggestive significance (1×10^{-5}).

of CVD, myocardial infarction and mortality in the UK Biobank cohort¹¹. During a median follow-up period of 2.8 years, the risk of CVD and myocardial infarction for individuals with higher ASI were 27% and 38% higher respectively. For all-cause mortality, each standard deviation change in ASI has a hazard ratio of 1.08 (95% confidence interval [CI], 1.05–1.12) in a multi-variant adjusted model¹¹.

Our current knowledge of the biological factors and pathways contributing to arterial stiffness is limited. Heritability studies of arterial stiffness using PWV measurements have suggested moderate genetic contribution with estimates up to 0.53 in twin studies^{12,13} and ranging between 0.26 and 0.40 in population studies^{14,15}. The identification of single nucleotide variants (SNVs) may improve the current understanding of the mechanisms controlling arterial stiffness and may be a useful addition to the development of disease risk models. Previous genome-wide association studies (GWASs) for arterial stiffness were mostly performed using PWV as the phenotype. One of the first arterial stiffness GWASs was on 644 individuals from the Framingham Heart Study, using carotid-brachial PWV as the phenotype, and no variants reached genome-wide significance¹⁶. Subsequent genetic studies have reported inconsistent findings but none have studied ASI as the phenotypic measure¹⁷. The wealth of data available within the UK Biobank study offers the opportunity to produce more robust findings in a large population cohort. We therefore performed a GWAS to identify SNVs associated with ASI and explored the biological mechanisms underlying arterial stiffness.

Results

Four loci identified for arterial stiffness index. In total, 143,590 UK Biobank participants met the inclusion criteria to derive ASI at the baseline visit (Supplementary Fig. S1). After genetic quality control (QC) exclusions, three GWASs were performed on 127,121 individuals of European ancestry (age 56 ± 8.1 , 48.1% males) with a mean ASI of 9.0 m/s. Our primary GWAS was performed on rank based inverse normal transformed (INT) residuals and included mean arterial pressure (MAP) as a covariate. We also performed analyses excluding MAP and on untransformed ASI. The baseline visit characteristics of the individuals included in the GWAS are summarised in Supplementary Table S1. Following the GWAS, we reviewed the results and quantile-quantile (QQ) plots. There was minimal genomic inflation in test statistics ($\lambda = 1.097$ for all models) under polygenic inheritance (Supplementary Fig. S2). The genome-wide SNV heritability of ASI in our cohort was estimated at 6.1% (standard error 0.4%).

We identified four genome-wide significant loci (*TEX41*, *FOXO1*, *C1orf21* and *MRVII*) for ASI in our primary analysis (Table 1 and Fig. 1). The regional association plots are indicated in Fig. 2. Three of the loci (*TEX41*, *FOXO1* and *C1orf21*) were genome-wide significant in the secondary GWAS that excluded the MAP adjustment (Supplementary Table S2a and Supplementary Fig. S3) and in the GWAS of untransformed ASI (Supplementary Table S2b). The magnitude of effect sizes and associations of the significant loci were similar across the primary and secondary models. We observed no independent signals at any of the ASI loci.

The most significant associated variants in our primary GWAS were: rs1006923, an intronic variant at *TEX41* ($\beta = 0.0293$, $P = 5.3 \times 10^{-12}$), rs7331212 an intronic variant at *FOXO1* ($\beta = 0.0301$, $P = 2.2 \times 10^{-11}$) and rs1930290

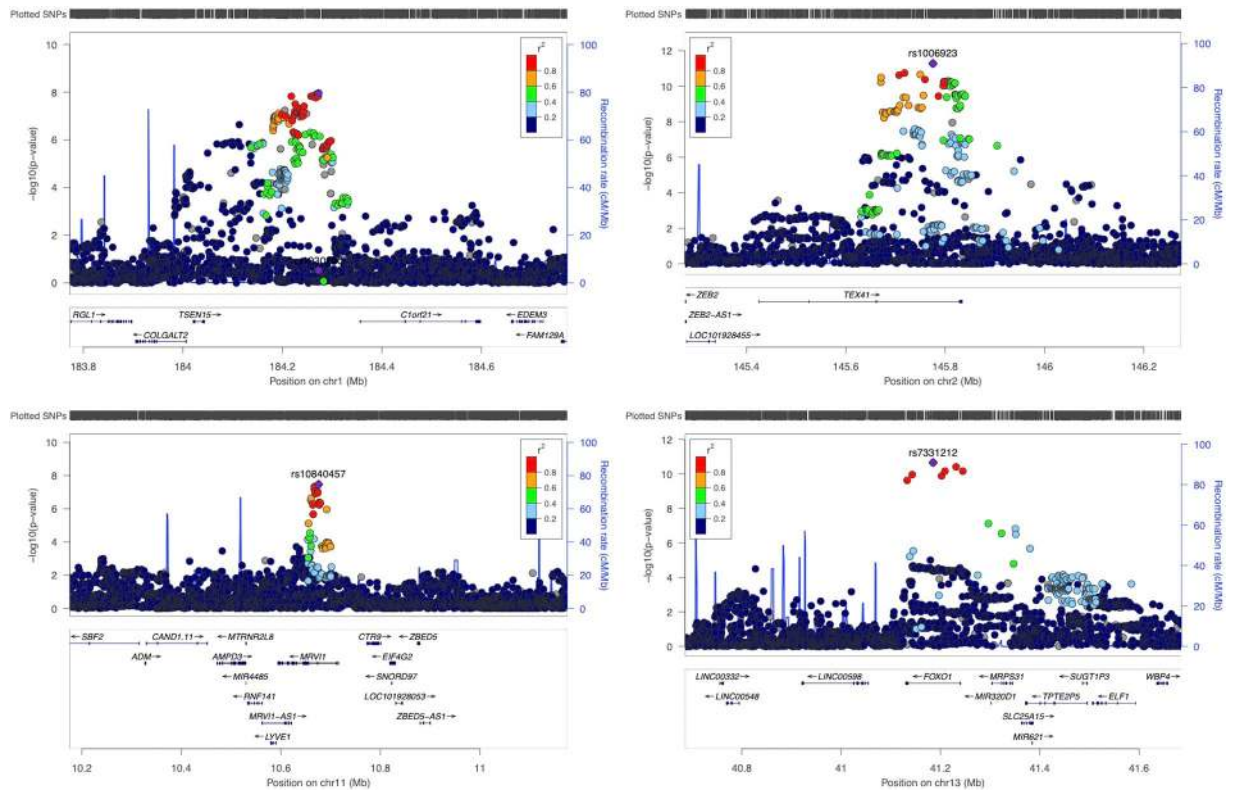


Figure 2. LocusZoom plots for arterial stiffness index loci ($P < 5 \times 10^{-8}$).

($\beta = -0.0230$, $P = 1.1 \times 10^{-8}$), located in the gene region of the open reading frame *C1orf21*. The intronic variant rs10840457 within the *MRVII* gene region on chromosome 11 ($\beta = -0.0236$, $P = 3.4 \times 10^{-8}$) was only genome-wide significant in our primary analysis, however there was some support in the secondary analyses ($P = 3.3 \times 10^{-7}$) and untransformed ASI ($P = 5.4 \times 10^{-8}$).

Forty-three variants at non-overlapping loci had suggestive genome-wide significance ($P < 1 \times 10^{-5}$) with ASI in our primary analysis (Supplementary Table S3). A few variants are located close to potential candidate genes of interest: rs9521719 in the *COL4A2* gene ($\beta = 0.0215$, $P = 1.1 \times 10^{-8}$), rs8107744 in the *RSPH6A* gene ($\beta = -0.0276$, $P = 1.2 \times 10^{-7}$). Other variants of note were rs371147897 found in the *OR4A47* gene region ($P = 1.6 \times 10^{-7}$), rs9501489 located in *DDR1* ($P = 4.0 \times 10^{-6}$) and rs149320025 located in the *FUCA1* gene region on chromosome 1 ($P = 4.0 \times 10^{-7}$).

Functional annotation of four ASI loci. In our primary analysis, 219 candidate variants were identified by SNP2GENE function in the Functional Mapping and Annotation of Genome-wide Association Studies (FUMA) platform¹⁸. The majority of SNVs and their proxies ($r^2 > 0.8$) were located in introns (63%), 31% were located in intergenic regions, there was one acceptor splice variant at *MRVII*, rs11042902 and the remainder were located in exons and 3'-untranslated regions (3'-UTRs). Of the 219 variants, 151 variants mapped to the four genome-wide significant loci.

Gene-based analysis, as computed by multi-marker analysis of genomic annotation (MAGMA)¹⁹, which mapped the output SNVs from BOLT-LMM to 18,666 protein coding genes, identified four genes that reached the gene-wide significance threshold ($P = 2.67 \times 10^{-6}$). *COL4A2* was the most significant gene for ASI ($P = 1.41 \times 10^{-8}$) and *MRVII* ($P = 3.08 \times 10^{-7}$), *FOXO1* ($P = 7.54 \times 10^{-7}$) and *FOXO3* ($P = 2.34 \times 10^{-6}$) were also significant (Supplementary Fig. S4). Two further genes, *TCF20* located 2.9 Mb downstream to rs1006923 ($P = 5.37 \times 10^{-6}$) and *FBXO46* located on chromosome 19 ($P = 6.72 \times 10^{-6}$) indicated suggestive significance. The *FBXO46* gene is located less than 67 kb away from rs8107744 at the *RSPH6A* locus on chromosome 19 that was significant from single variant analyses using untransformed ASI values ($P = 1.4 \times 10^{-8}$, Supplementary Table S2b).

To identify further candidate genes at each locus we reviewed results from expression quantitative trait loci (eQTL) analyses across 53 tissue types from Genotype-Tissue Expression (GTEx) database from FUMA¹⁸. Significant eQTLs were observed for *SLC25A15* at the *FOXO1* locus across several tissues including oesophagus, transformed fibroblasts and sun exposed skin in the lower leg (lead SNV rs12865518, $P = 2.04 \times 10^{-8}$), *ZEB2* at the *TEX41* locus in the aorta (lead SNV rs2252383, $P = 3.85 \times 10^{-5}$) and *C1orf21* and *APOBEC4* at the *C1orf21* locus in nerve and brain tissue (Supplementary Table S4).

We also checked if the genetic variants identified for ASI were associated with other traits using PhenoScanner²⁰. Genome-wide associations were observed for variants at three of the four ASI loci. Variants at *TEX41* (rs1006923) and *MRVII* (rs10840457) were associated with blood pressure traits and ASI from analyses in

UK Biobank²¹. Significant associations were also observed with pulse wave peak to peak time (PPT) for rs1006923 (*TEX41*) and rs7331212 at the *FOXO1* locus by the Neale lab²¹ that has publicly released GWAS results for over 4,200 phenotypes found within UK Biobank. The variant at the *TEX41* locus was additionally associated with CAD²² (Supplementary Table S5).

Discussion

Our main finding was the identification of four genome-wide significant loci for ASI in a large European-ancestry based population cohort, despite the relatively low (0.06) heritability observed for ASI. We further followed-up with several candidate genes using bioinformatics analyses at each of the identified loci.

Previous genetic association studies for arterial stiffness, as summarised by Logan *et al.*¹⁷, have mainly used PWV as the phenotypic measure of stiffness and the reported findings were limited. A meta-analysis including GWAS results of 20,634 individuals from 9 discovery cohorts and of 5,306 individuals from two replication cohorts, all from European ancestry, identified one locus on chromosome 14 in the 3'-*BCL11B* gene desert. This gene desert was shown to be associated with the carotid-femoral PWV (rs7152623, discovery $P = 2.8 \times 10^{-10}$, replication $P = 1.4 \times 10^{-6}$)²³. However, variants at this locus were not statistically significantly associated with ASI ($P = 0.08$) in our dataset.

More recently, a GWAS for brachial-ankle PWV (baPWV), involving 402 Korean patients (mean age 59 years, 59% male) with diagnosed CVD has been reported by Park *et al.*²⁴. Two SNVs were found to be associated with baPWV – rs7271920 ($P = 7.15 \times 10^{-9}$) and rs10125157 ($P = 8.25 \times 10^{-7}$). Neither variant was significant (rs7271920, $P = 3.51 \times 10^{-1}$; rs10125157, $P = 3.10 \times 10^{-1}$) in the replication cohort in their study that included 1,206 individuals. We also observed non-significant results for both variants (rs7271920, $P = 0.17$; rs10125157, $P = 0.81$) in our study. It is not too surprising to observe a lack of significant findings for the two variants reported by Park *et al.*²⁴ in the UK Biobank cohort. A contributing factor may be the difference in the size of the populations, and the different ethnicities, and that the reported variants by Park *et al.*²⁴ may be false positive findings as there was no replication in their study.

The lack of replication of results on PWV in UK Biobank and across PWV GWASs may also be due to the lack of methodological standardisation to derive PWV, as it can be measured at different sites such as carotid-femoral or brachial-ankle. Our findings suggest that ASI and PWV may have different aetiologies and thus may provide independent data on the underlying biological mechanisms and potential cardiovascular risk factors. A previous expert consensus statement on arterial stiffness described PWV as a measure of regional stiffness, while ASI is seen as a surrogate marker of stiffness through wave reflection assessments²⁵. In other words, whilst PWV is determined mainly by the speed at which waveform travels, for ASI, other factors such as the reflective point would also impact on its measurements leading to potential differences in their aetiologies.

Our most significant variant for ASI, rs1006923 at the *TEX41* locus has previously been reported to be significantly associated with CAD in a mixed population GWAS that included UK Biobank participants²². It is located 129 kb upstream and in low linkage disequilibrium (LD; $r^2 = 0.19$) to rs1438896 a variant reported by Warren *et al.*²⁶ with blood pressure traits. Another variant at the *TEX41* locus rs183032 is 50 kb downstream to our lead variant with a moderate LD ($r^2 = 0.38$) is associated with aortic stenosis in an Icelandic cohort²⁷. Helgadóttir (2018)²⁷ suggested the tumour growth factor- β (TGF- β) through *ZEB2*²⁷ as a candidate gene. *ZEB2* has a role as a DNA-binding transcriptional repressor that interacts with the main signal transducers for TGF- β receptors (SMADs). Considering that TGF- β is involved in the regulation of vascular smooth muscle differentiation, as well as in the collagen up-regulation in the vascular wall, changes in the expression of TGF- β have the potential to alter arterial stiffness. We note additional support for *ZEB2* from eQTL data in aortic tissue in our analyses (Supplementary Table S4).

At the second ASI locus, *FOXO1* represents a good candidate gene, rs7331212 is located within the O class of the forkhead family of transcription factors. There is genome-wide significant association of the same variant with pulse wave PPT²¹, which is the time interval between the peak values of the direct and reflected components of the pulse waveform using to calculate ASI²⁸. Although the specific function of *FOXO1* gene is not well described, it may play a role in blood pressure regulation. Specifically, lack of *FOXO1* has been shown to reduce expression of angiotensinogen, which is a precursor of angiotensin II that mediates vasoconstriction, in knockout mouse models²⁹. Furthermore, *FOXO1* is involved in the signalling axis that regulates mindin³⁰, which has a role in neointima formation where there is vascular smooth muscle cell proliferation. Importantly, the significance of the association between our lead variant at *FOXO1* and ASI persisted after adjustment for MAP in our study.

At our third ASI locus, there are few candidate genes. *C1orf21* is an uncharacterised protein-coding gene that at present has not been functionally annotated.

At the 4th locus rs10840457 is located near the *MRVII* (Murine Retrovirus Integration Site 1 Homolog, also known as IP3R-associated cGMP kinase substrate (*IRAG*)). The *MRVII* gene is responsible for encoding the *MRVII/IRAG* protein, which is present in a number of tissues including aorta and trachea³¹, and is involved in smooth muscle contractility. Specifically, there is inhibition of calcium release from endoplasmic reticulum following co-expression of *IRAG* and cGMP-dependent protein kinase type I β (cGKI β) in the presence of cGMP³¹. In a study using *IRAG*-knockout mice, the authors concluded that signalling of cGKI β via *IRAG* is a vital functional component in the regulation of smooth muscle tone and intracellular calcium by nitric oxide and atrial natriuretic peptide³². At this locus *MRVII* represents an interesting candidate, but we note this is a locus that was not genome-wide significant in the secondary analyses, thus further validation will be required.

Gene-based testing by MAGMA revealed *COL4A2* as the most significant gene association with ASI in our cohort, with *FOXO1*, *MRVII* and *FOXO3* also significant. The lead SNV from the GWAS at the *COL4A2* locus is rs9521719 ($P = 1.1 \times 10^{-7}$). *COL4A2*, along with the adjacent gene *COL4A1*, encodes the protein subunits of type IV collagen forming heterotrimers. Type IV collagen is a vital structural component of basement membranes and mutations in these genes are seen in disorders such as myopathy, intracerebral haemorrhage and glaucoma³³. A

previous GWAS reported rs3742207 located near the *COL4A1* to have strong association ($P = 7.08 \times 10^{-7}$) with PWV in a cohort comprising of 4,221 Sardinian individuals³⁴. This variant was successfully replicated internally in 1,828 individuals and also in 813 Amish individuals. The PWV variant rs3742207 did not show any association with ASI in our cohort ($P = 0.95$). This result may add additional support on there being a difference in the genetic mechanism of ASI compared to PWV, both markers of arterial stiffness.

We found the genome-wide SNV heritability of ASI to be 6%. This estimate is much lower than the reported heritability for carotid-femoral PWV, which ranged between 0.36–0.40^{14,35}. This dissimilarity might be explained by the differences in stiffness measure, covariates included and populations. In addition, due to the methodological inconsistencies mentioned above, it should be noted that the heritability of PWV would differ depending on the site of measurement. Mitchell *et al.*¹⁴ reported moderate heritability for carotid-femoral PWV ($h^2 = 0.40$) in their study of 1,480 participants in the Framingham Study offspring cohort. However, when using the less common approach of measuring PWV between the carotid and brachial artery, the heritability estimate was lower ($h^2 = 0.09$). In twin studies, the reported heritability estimates range between 0.38 and 0.53^{12,13,36} where the PWV measurements were either located at the wrist (aorto-radial) or foot (aorto-dorsalis-pedis).

The main strength of our study is the very large sample size that, despite the low estimated heritability for ASI, has enabled the identification of genetic variants not found in previous studies. Mean arterial pressure has been shown to have a strong influence on arterial stiffness¹⁵ so it was included as a covariate in the fully adjusted model, rather than systolic blood pressure (SBP) and diastolic blood pressure (DBP). Despite using the largest cohort to date in identifying genetic variants for arterial stiffness, the main limitation of our study is that loci discovered therein require formal validation in an independent dataset. However, an external study of similar sample size with ASI measurements is currently lacking. We note if we use a more stringent P -value for reporting, $P \leq 1 \times 10^{-8}$, then two loci would remain significant (*TEX41* and *FOXO1*). In addition, our study cohort consisted of middle-aged individuals of European-ancestry, so our findings may not be generalised to other age groups and ethnic populations.

In conclusion, we identified four loci significantly associated with ASI, an independent predictor of CVD and mortality. The two most significant loci, *TEX41* and *FOXO1*, have SNV associations that may alter arterial stiffness through blood pressure regulation and vascular smooth muscle differentiation. Our results also suggest an important role of calcium in the regulation of smooth muscle tone contributing to arterial stiffness. Further research will be necessary to validate our discovered loci in a separate cohort and their confirmation can potentially lead to the development of risk models that can be used in clinical practice.

Methods

Study population. The UK Biobank is a large population-based prospective study of >500,000 participants aged 40–69 years at baseline recruited between 2006 and 2010 in England, Scotland and Wales. During the initial assessment visit, a broad range of biochemical, clinical and genotype data were collected and participants had a number of physical measurements. The UK Biobank study was approved by the North West Multi-Centre Research Ethics Committee and all enrolled individuals have provided written informed consent for collection, storage, make availability of their data for health-related research. All methods were carried out in accordance with the relevant guidelines and regulations.

Arterial stiffness index measures. Pulse wave ASI (UK Biobank Field 21021), measured in m/s, was derived using the pulse waveform obtained at the finger (preferably index finger of the non-dependent hand though can be placed on any finger or thumb) with an infra-red sensor (PulseTrace PCA 2TM, CareFusion, USA). The shape of the volume waveform in the finger is directly related to the time it takes for the pulse waveforms to travel through the arterial tree in the lower body and to be reflected back to the finger. Measurements were made by clipping the device to a finger and the reading is made over 10–15 seconds. As the participants' heights were unknown until after the recording of the waveform data, the actual ASI values were calculated ($ASI = \text{height}/PPT$) by UK Biobank outside the assessment centre visit.

Data were taken at the UK Biobank Assessment Centres during the baseline recruitment between 29th April 2009 and 1st October 2010 inclusive ($N_{\text{max}} = 169,822$). UK Biobank participants are free to withdraw at any time and so 15 individuals were removed based on the application specific list of anonymised IDs. Individuals with absence of notch position in the pulse waveform (Field 4204) were excluded ($n = 25,288$) as the notch in the digital volume pulse waveform indicates the reflected component of the pressure transmitted and is therefore required to calculate PPT. Outlier ASI values, defined as three inter-quartile ranges below the first quartile or above the third quartile ($n = 81$), were also removed from analyses. Finally, participants with missing height, weight or blood pressure measurements were removed from the final analyses ($n = 848$).

Covariates. Blood pressure was measured using the Omron 705 IT electronic blood pressure monitor (OMRON Healthcare Europe B.V. Kruisweg 577 2132 NA Hoofddorp). SBP and DBP were derived as the mean of the two recorded automated measurements (UK Biobank Fields 4079, 4080), except for 1,141 individuals who only had one recorded reading. MAP estimation was calculated using the traditional formula: $MAP = DBP + 1/3(SBP-DBP)$. Height (Field 50) was measured using a Seca 202 device (Seca, Birmingham, UK). For individuals who reported use of anti-hypertensive medications through questionnaires in UK biobank (Fields 6177, 6153), their SBP and DBP were adjusting by adding 15 mmHg and 10 mmHg respectively to the mean recorded readings^{26,37}. The ID numbers of the arterial stiffness devices used were obtained from Field 4206 in UK Biobank.

Genotypic data. Central QC and imputation of genotypic data performed by UK Biobank has been previously described³⁸. Briefly, genotypic data was obtained through either UK Biobank Axiom or UK BiLEVE

Axiom arrays (Affymetrix Research Service Laboratory, Santa Clara, California, USA). The Haplotype Reference Consortium (HRC) and the merged UK10K sequencing +1000 Genomes were used as a reference panels for imputation with preference for the HRC panel where SNPs were present in both panels. This study utilised the refreshed genetic dataset made available by UK Biobank in July 2017³⁸.

Genetic QC was performed in 164,835 individuals who underwent assessment of arterial stiffness using finger photoplethysmography and had available genotypic data. This process excluded participants with either high missingness or high heterozygosity defined by UK Biobank in their genotypic data ($n = 345$), as well as those with mis-match between self-reported and inferred sex from the genotypes ($n = 116$).

A 4-way k -means clustering analysis was performed according to data from the first and second principal components (PC1 and PC2) using the 'pvcust' R package (version 2.0–0)³⁹ to objectively identify the main ethnic groups (White, Asian, African and Chinese) within UK Biobank. "White" participants were defined for those present in the "White" cluster for both PC1 and PC2 analyses. They also needed to match their self-reported ancestry though "mixed", "other" and "missing" were treated as being broad ethnicity. After restriction to European-ancestry only, 127,121 individuals remained. Lastly, we performed SNV-level QC to exclude 237,751 SNVs from genotyped SNVs in these individuals with the following thresholds: minor allele frequency (MAF) of 1%, Hardy-Weinberg equilibrium (P -value of 1×10^{-6}) and missingness of 0.015 using PLINK 1.9⁴⁰. After exclusions, a final total of 546,505 model SNVs was generated for our GWAS analyses.

Genome-wide association and heritability analyses. We applied rank-based INT on the residuals from the regression of ASI against the phenotypic covariates (i.e., all except genotyping array and PCs) before performing genetic analyses as the distribution did not approximate a normal distribution (Supplementary Fig. S5). The heritability of ASI explained by additive genetic variation was estimated using a variance components method (BOLT-REML)⁵¹.

We performed three GWASs using BOLT-LMM (v2.3.2) software⁴¹ across our 546,505 model SNVs and ~9.9 million imputed SNVs with $MAF \geq 1\%$ and imputation information (INFO) score ≥ 0.3 . Our primary ASI GWAS was on the INT residuals including the following covariates: age, age², sex, weight, genotyping array (UK Biobank vs UK BiLEVE), device used to obtain pulse waveform, smoking status (current vs non-current smokers), MAP and first 10 principal components (PCs). As a sensitivity analyses, we also performed a GWAS using untransformed ASI values and same covariates as those included in our primary analysis.

A secondary GWAS where MAP was excluded as a covariate was also performed. All GWASs assumed a linear mixed infinitesimal model method under an additive genetic model implemented in BOLT-LMM. This method accounts for cryptic population structure and allows the inclusion of related individuals permitting greater power compared to principal component analysis⁴¹.

Genome-wide significance was defined as P -value $\leq 5 \times 10^{-8}$ and R (version 3.5.1) statistical Software⁴² was used to generate Manhattan plots and QQ plots. Regional association plots were made for genome-wide significant loci using the LocusZoom web-based platform (locuszoom.org) with the hg19/1000 Genomes (Nov 2014) EUR Build.

Conditional analysis was performed for genome-wide significant locus to detect independent association signals using an approximate conditional and joint multiple-SNP (COJO) analysis implemented in genome-wide complex trait analysis (GCTA) tool⁴³. A secondary signal was declared if all three of the following conditions were met:

- (i) original P -value of newly identified variant was $< 1 \times 10^{-6}$
- (ii) ratio between the lead SNV and secondary association P -values on a $-\log_{10}$ scale is 1.5 or less, (i.e., $-\log_{10}(P \text{ lead}) / -\log_{10}(P \text{ sec}) < 1.5$)
- (iii) ratio between the original association and conditional association P -values on a $-\log_{10}$ scale is 1.5 or less (i.e., $-\log_{10}(P) / -\log_{10}(P \text{ cond}) < 1.5$)

Functional annotation. Summary statistics from our primary model GWAS analyses were uploaded onto the FUMA (v1.3.4c)¹⁸ web-based application (<http://fuma.ctglab.nl/>) in order to perform functional annotations via its SNP2GENE function. The default FUMA settings were used whereby significant SNVs ($P < 5 \times 10^{-8}$), and those that were in LD ($r^2 \geq 0.6$), with a $MAF \geq 0.01$ were selected for further annotation. In addition, the maximum distance between LD blocks to merge into a locus was set at < 250 kb and the UKB release 2 European was used as the reference panel population. The functional consequences for these SNVs were then obtained with ANNOVAR⁴⁴. For gene-based analysis, the integrated MAGMA v1.06¹⁹ was utilised in which SNVs are mapped to a gene according their genomic location before gene-level associations with ASI were tested. Here, a P -value with Bonferroni adjustment of $0.05/18697 = 2.67 \times 10^{-6}$ was used to define genome-wide significance. For eQTL mapping, this was implemented by the FUMA platform, GTEx v7 database⁴⁵ (<http://gtexportal.org/home/>), 53 tissue types was used as the gene expression reference data and the significant threshold was defined as false discovery rate (FDR) < 0.05 . Genome-wide significant SNVs and those in LD with $r^2 > 0.8$ were examined for previously reported GWAS associations with other traits using PhenoScanner²⁰, an online search tool containing over 350 million publicly available association results. The significance threshold chosen for reporting therein was P -value $< 5 \times 10^{-8}$.

Data Availability

The UK Biobank Resource is available, via application, to all bona fide researchers undertaking health-related research that is in the public interest. Summary data are available online: www.ukbiobank.ac.uk/data-showcase. Information on accessing the genetic and phenotype data used in this analysis can be found at www.ukbiobank.ac.uk/using-the-resource/.

References

- Benetos, A. *et al.* Telomere length as an indicator of biological aging: the gender effect and relation with pulse pressure and pulse wave velocity. *Hypertension* **37**, 381–5 (2001).
- Scuteri, A. *et al.* Longitudinal perspective on the conundrum of central arterial stiffness, blood pressure, and aging. *Hypertension* **64**, 1219–1227 (2014).
- Dernellis, J. & Panaretou, M. Aortic stiffness is an independent predictor of progression to hypertension in nonhypertensive subjects. *Hypertension* **45**, 426–431 (2005).
- Liao, D. *et al.* Arterial stiffness and the development of hypertension. The ARIC study. *Hypertension* **34**, 201–6 (1999).
- Liao, J. & Farmer, J. Arterial stiffness as a risk factor for coronary artery disease. *Curr. Atheroscler. Rep.* **16**, 387 (2014).
- Vlachopoulos, C., Aznaouridis, K. & Stefanadis, C. Prediction of cardiovascular events and all-cause mortality with arterial stiffness: a systematic review and meta-Analysis. *J. Am. Coll. Cardiol.* **55**, 1318–1327 (2010).
- Laurent, S. *et al.* Aortic stiffness is an independent predictor of all-cause and cardiovascular mortality in hypertensive patients. *Hypertension* **37**, 1236–41 (2001).
- Millasseau, S. C. *et al.* Noninvasive assessment of the digital volume pulse: Comparison with the peripheral pressure pulse. *Hypertension* **36**, 952–956 (2000).
- Woodman, R. J. *et al.* Assessment of central and peripheral arterial stiffness: Studies indicating the need to use a combination of techniques. *Am. J. Hypertens.* **18**, 249–260 (2005).
- Kao, Y.-T. *et al.* Arterial stiffness index and coronary artery plaques in patients with subclinical coronary atherosclerosis. *Acta Cardiol. Sin.* **31**, 59–65 (2015).
- Said, M. A., Eppinga, R. N., Lipsic, E., Verweij, N. & van der Harst, P. Relationship of arterial stiffness index and pulse pressure with cardiovascular disease and mortality. *J. Am. Heart Assoc.* **7**, e007621 (2018).
- Ge, D. *et al.* Heritability of arterial stiffness in black and white American youth and young adults. *Am. J. Hypertens.* **20**, 1065–72 (2007).
- Tarnoki, A. D. *et al.* Heritability of central blood pressure and arterial stiffness. *J. Hypertens.* **30**, 1564–1571 (2012).
- Mitchell, G. F. *et al.* Heritability and a genome-wide linkage scan for arterial stiffness, wave reflection, and mean arterial pressure: The Framingham heart study. *Circulation* **112**, 194–199 (2005).
- Sayed-Tabatabaei, F. A. *et al.* Heritability of the function and structure of the arterial wall: findings of the Erasmus Rucphen Family (ERF) study. *Stroke* **36**, 2351–2356 (2005).
- Levy, D. *et al.* Framingham Heart Study 100 K Project: genome-wide associations for blood pressure and arterial stiffness. *BMC Med. Genet.* **8**(Suppl 1), S3 (2007).
- Logan, J. G., Engler, M. B. & Kim, H. Genetic determinants of arterial stiffness. *J. Cardiovasc. Transl. Res.* **8**, 23–43 (2015).
- Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
- de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLOS Comput. Biol.* **11**, e1004219 (2015).
- Staley, J. R. *et al.* PhenoScanner: a database of human genotype-phenotype associations. *Bioinforma. Oxf. Engl.* **32**, 3207–3209 (2016).
- Abbott, L. *et al.* UK Biobank — Neale lab. (2018). Available at, <http://www.nealelab.is/uk-biobank/> (Accessed: 2nd January 2019).
- van der Harst, P. & Verweij, N. Identification of 64 novel genetic loci provides an expanded view on the genetic architecture of coronary artery disease. *Circ. Res.* **122**, 433–443 (2018).
- Mitchell, G. F. *et al.* Common genetic variation in the 3'-BCL11B gene desert is associated with carotid-femoral pulse wave velocity and excess cardiovascular disease risk: the AortaGen Consortium. *Circ. Cardiovasc. Genet.* **5**, 81–90 (2012).
- Park, S. *et al.* Lack of association between arterial stiffness and genetic variants by genome-wide association scan. *Blood Press.* **24**, 258–261 (2015).
- Laurent, S. *et al.* Expert consensus document on arterial stiffness: methodological issues and clinical applications. *Eur. Heart J.* **27**, 2588–2605 (2006).
- Warren, H. R. *et al.* Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. *Nat. Genet.* **49**, 403–415 (2017).
- Helgadottir, A. *et al.* Genome-wide analysis yields new loci associating with aortic valve stenosis. *Nat. Commun.* **9**, 987 (2018).
- UK Biobank. UK Biobank arterial pulse-wave velocity. (2011). Available at, <https://biobank.ctsu.ox.ac.uk/crystal/docs/Pulsewave.pdf> (Accessed: 7th January 2019).
- Qi, Y. *et al.* Novel mechanism of blood pressure regulation by forkhead box class O1-mediated transcriptional control of hepatic angiotensinogen. *Hypertension* **64**, 1131–40 (2014).
- Zhu, L.-H. *et al.* Mindin regulates vascular smooth muscle cell phenotype and prevents neointima formation. *Clin. Sci. Lond. Engl.* **1979**(129), 129–45 (2015).
- Schlossmann, J. *et al.* Regulation of intracellular calcium by a signalling complex of IRAG, IP3 receptor and cGMP kinase I β . *Nature* **404**, 197–201 (2000).
- Desch, M. *et al.* IRAG determines nitric oxide- and atrial natriuretic peptide-mediated smooth muscle relaxation. *Cardiovasc. Res.* **86**, 496–505 (2010).
- Kuo, D. S., Labelle-Dumais, C. & Gould, D. B. COL4A1 and COL4A2 mutations and disease: insights into pathogenic mechanisms and potential therapeutic targets. *Hum. Mol. Genet.* **21**, R97–110 (2012).
- Tarasov, K. V. *et al.* COL4A1 is associated with arterial stiffness by genome-wide association scan. *Circ. Cardiovasc. Genet.* **2**, 151–8 (2009).
- Sayed-Tabatabaei, F. A. *et al.* Heritability of the function and structure of the arterial wall: Findings of the Erasmus Rucphen Family (ERF) study. *Stroke* **36**, 2351–2356 (2005).
- Cecelja, M. *et al.* Arterial stiffening is a heritable trait associated with arterial dilation but not wall thickening: a longitudinal study in the twins UK cohort. *Eur. Heart J.* **39**, 2282–2288 (2018).
- Tobin, M. D., Sheehan, N. A., Scurrah, K. J. & Burton, P. R. Adjusting for treatment effects in studies of quantitative traits: antihypertensive therapy and systolic blood pressure. *Stat. Med.* **24**, 2911–2935 (2005).
- Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
- Suzuki, R. & Shimodaira, H. Pvcust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* **22**, 1540–1542 (2006).
- Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**, 7 (2015).
- Loh, P.-R. *et al.* Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat. Genet.* **47**, 284–290 (2015).
- R Development Core Team. R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. Available at, <http://www.r-project.org/> (2016).
- Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–375 (2012).
- Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164–e164 (2010).
- Aguet, F. *et al.* Genetic effects on gene expression across human tissues. *Nature* **550**, 204–213 (2017).

Acknowledgements

K.F. is supported by The Medical College of Saint Bartholomew's Hospital Trust, an independent registered charity that promotes and advances medical and dental education and research at Barts and The London School of Medicine and Dentistry. J.R. acknowledges support from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007–2013) under REA grant agreement n° 608765 and the Medical Research Council grant MR/N025083/1. N.A. is supported by a Wellcome Trust Research Training Fellowship (203553/Z/16/Z). We also acknowledge support from the National Institute for Health Research (NIHR) Cardiovascular Biomedical Research Centre at Barts and Queen Mary University of London, UK. A.M.L. and S.E.P. acknowledge support from the "SmartHeart" EPSRC programme grant (www.nihr.ac.uk; EP/P001009/1). This project was enabled through access to the MRC eMedLab Medical Bioinformatics infrastructure, supported by the Medical Research Council (Grant Number MR/L016311/1). This research has been conducted using the UK Biobank Resource (Application 2964). The authors wish to thank all UK Biobank participants and staff.

Author Contributions

K.F. contributed in the conceptualisation, formal analysis, methodology, interpretation of the results and writing the original draft. J.R. contributed in the formal analysis and critically revised the manuscript. H.R.W., A.M.L. and E.T. contributed in the data curation and resources. N.A. contributed in methodology. S.E.P. was involved in the conceptualisation, interpretation of results, supervision and critically revised the original manuscript. P.B.M. was involved in conceptualisation, methodology, formal analysis, interpretation of results, supervision and critically revised the original manuscript. All authors reviewed the manuscript draft.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-019-45703-0>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019