

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

3

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

6

7 ¹Advanced Cardiovascular Imaging, William Harvey Research Institute, Queen Mary
8 University of London, Charterhouse Square, London, UK.

9

10 ²Clinical Pharmacology, William Harvey Research Institute, Queen Mary University of
11 London, Charterhouse Square, London, UK.

12

13 ³NIHR Barts Biomedical Research Centre, Queen Mary University of London, Charterhouse
14 Square, London, UK.

15

16 * jointly supervised the project

17

18 Correspondence to: p.b.munroe@qmul.ac.uk and s.e.petersen@qmul.ac.uk

19

20

1 **ABSTRACT**

2

3 Arterial stiffness index (ASI) is a non-invasive measure of arterial stiffness using infra-red
4 finger sensors (photoplethysmography). It is a well-suited measure for large populations as it
5 is relatively inexpensive to perform, and data can be acquired within seconds. These features
6 raise interest in using ASI as a tool to estimate cardiovascular disease risk as prior work
7 demonstrates increased arterial stiffness is associated with elevated systolic blood pressure,
8 and ASI is predictive of cardiovascular disease and mortality. We conducted genome-wide
9 association studies (GWASs) for ASI in 127,121 UK Biobank participants of European-
10 ancestry. Our primary analyses identified variants at four loci reaching genome-wide
11 significance ($P < 5 \times 10^{-8}$): *TEX41* (rs1006923; $P = 5.3 \times 10^{-12}$), *FOXO1* (rs7331212; $P = 2.2$
12 $\times 10^{-11}$), *C1orf21* (rs1930290, $P = 1.1 \times 10^{-8}$) and *MRVI1* (rs10840457, $P = 3.4 \times 10^{-8}$). Gene-
13 based testing revealed three significant genes, the most significant gene was *COL4A2* ($P =$
14 1.41×10^{-8}) encoding type IV collagen. Other candidate genes at associated loci were also
15 involved in smooth muscle tone regulation. Our findings provide new information for
16 understanding the development of arterial stiffness.

1 INTRODUCTION

2

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

19

20 Previous studies have evaluated the utility of ASI as a potential clinical marker of CVD, having
21 shown good sensitivity (87%) and specificity (87%) when differentiating between older men
22 with coronary artery disease (CAD) and younger men without CAD⁹. In a study of
23 asymptomatic middle-aged patients, the mean ASI value was significantly higher ($P = 0.002$)
24 in individuals with at least one significant (>50%) coronary stenotic plaque than those without
25 a >50% stenotic lesion¹⁰. More recently, ASI has been shown to be an independent predictor
26 of CVD, myocardial infarction and mortality in the UK Biobank cohort¹¹. During a median

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

5

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

1 RESULTS

2

3 Four loci identified for arterial stiffness index

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

15

16 We identified four genome-wide significant loci (*TEX41*, *FOXO1*, *C1orf21* and *MRVI1*) for ASI
17 in our primary analysis (Table 1 and Fig. 1). The regional association plots are indicated in
18 Fig. 2. Three of the loci (*TEX41*, *FOXO1* and *C1orf21*) were genome-wide significant in the
19 secondary GWAS that excluded the MAP adjustment (Supplementary Table S2a and
20 Supplementary Fig. S3) and in the GWAS of untransformed ASI (Supplementary Table S2b).
21 The magnitude of effect sizes and associations of the significant loci were similar across the
22 primary and secondary models. We observed no independent signals at any of the ASI loci.

23

24 The most significant associated variants in our primary GWAS were: rs1006923, an intronic
25 variant at *TEX41* ($\beta = 0.0293$, $P = 5.3 \times 10^{-12}$), rs7331212 an intronic variant at *FOXO1* ($\beta =$
26 0.0301 , $P = 2.2 \times 10^{-11}$) and rs1930290 ($\beta = -0.0230$, $P = 1.1 \times 10^{-8}$), located in the gene

1 region of the open reading frame *C1orf21*. The intronic variant rs10840457 within the *MRVI1*
2 gene region on chromosome 11 ($\beta = -0.0236$, $P = 3.4 \times 10^{-8}$) was only genome-wide
3 significant in our primary analysis, however there was some support in the secondary
4 analyses ($P = 3.3 \times 10^{-7}$) and untransformed ASI ($P = 5.4 \times 10^{-8}$).

5

6 Forty-three variants at non-overlapping loci had suggestive genome-wide significance ($P < 1$
7 $\times 10^{-5}$) with ASI in our primary analysis (Supplementary Table S3). A few variants are located
8 close to potential candidate genes of interest: rs9521719 in the *COL4A2* gene ($\beta = 0.0215$,
9 $P = 1.1 \times 10^{-8}$), rs8107744 in the *RSPH6A* gene ($\beta = -0.0276$, $P = 1.2 \times 10^{-7}$). Other variants
10 of note were rs371147897 found in the *OR4A47* gene region ($P = 1.6 \times 10^{-7}$), rs9501489
11 located in *DDR1* ($P = 4.0 \times 10^{-6}$) and rs149320025 located in the *FUCA1* gene region on
12 chromosome 1 ($P = 4.0 \times 10^{-7}$).

13

14 **Functional annotation of four ASI loci**

15 In our primary analysis, 219 candidate variants were identified by SNP2GENE function in the
16 Functional Mapping and Annotation of Genome-wide Association Studies (FUMA) platform¹⁸.
17 The majority of SNVs and their proxies ($r^2 > 0.8$) were located in introns (63%), 31% were
18 located in intergenic regions, there was one acceptor splice variant at *MRVI1*, rs11042902
19 and the remainder were located in exons and 3'-untranslated regions (3'-UTRs). Of the 219
20 variants, 151 variants mapped to the four genome-wide significant loci.

21

22 Gene-based analysis, as computed by multi-marker analysis of genomic annotation
23 (MAGMA)¹⁹, which mapped the output SNVs from BOLT-LMM to 18,666 protein coding
24 genes, identified four genes that reached the gene-wide significance threshold ($P = 2.67 \times$
25 10^{-6}). *COL4A2* was the most significant gene for ASI ($P = 1.41 \times 10^{-8}$) and *MRVI1* ($P = 3.08$
26 $\times 10^{-7}$), *FOXO1* ($P = 7.54 \times 10^{-7}$) and *FOXO3* ($P = 2.34 \times 10^{-6}$) were also significant

1 (Supplementary Fig. S4). Two further genes, *TCF20* located 2.9Mb downstream to
2 rs1006923 ($P = 5.37 \times 10^{-6}$) and *FBXO46* located on chromosome 19 ($P = 6.72 \times 10^{-6}$)
3 indicated suggestive significance. The *FBXO46* gene is located less than 67kb away from
4 rs8107744 at the *RSPH6A* locus on chromosome 19 that was significant from single variant
5 analyses using untransformed ASI values ($P = 1.4 \times 10^{-8}$, Supplementary Table S2b).

6
7 To identify further candidate genes at each locus we reviewed results from expression
8 quantitative trait loci (eQTL) analyses across 53 tissue types from Genotype-Tissue
9 Expression (GTEx) database from FUMA¹⁸. Significant eQTLs were observed for *SLC25A15*
10 at the *FOXO1* locus across several tissues including oesophagus, transformed fibroblasts
11 and sun exposed skin in the lower leg (lead SNV rs12865518, $P = 2.04 \times 10^{-8}$), *ZEB2* at the
12 *TEX41* locus in the aorta (lead SNV rs2252383, $P = 3.85 \times 10^{-5}$) and *C1orf21* and *APOBEC4*
13 at the *C1orf21* locus in nerve and brain tissue (Supplementary Table S4).

14
15 We also checked if the genetic variants identified for ASI were associated with other traits
16 using PhenoScanner²⁰. Genome-wide associations were observed for variants at three of the
17 four ASI loci. Variants at *TEX41* (rs1006923) and *MRVI1* (rs10840457) were associated with
18 blood pressure traits and ASI from analyses in UK Biobank²¹. Significant associations were
19 also observed with pulse wave peak to peak time (PPT) for rs1006923 (*TEX41*) and
20 rs7331212 at the *FOXO1* locus by the Neale lab²¹ that has publicly released GWAS results
21 for over 4,200 phenotypes found within UK Biobank. The variant at the *TEX41* locus was
22 additionally associated with CAD²² (Supplementary Table S5).

1 **DISCUSSION**

2

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

7

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

16

17 More recently, a GWAS for brachial-ankle PWV (baPWV), involving 402 Korean patients
18 (mean age 59 years, 59% male) with diagnosed CVD has been reported by Park et al.²⁴. Two
19 SNVs were found to be associated with baPWV – rs7271920 ($P = 7.15 \times 10^{-9}$) and
20 rs10125157 ($P = 8.25 \times 10^{-7}$). Neither variant was significant (rs7271920, $P = 3.51 \times 10^{-1}$;
21 rs10125157, $P = 3.10 \times 10^{-1}$) in the replication cohort in their study that included 1,206
22 individuals. We also observed non-significant results for both variants (rs7271920, $P = 0.17$;
23 rs10125157, $P = 0.81$) in our study. It is not too surprising to observe a lack of significant
24 findings for the two variants reported by Park et al.²⁴ in the UK Biobank cohort. A contributing
25 factor may be the difference in the size of the populations, and the different ethnicities, and

1 that the reported variants by Park et al.²⁴ may be false positive findings as there was no
2 replication in their study.

3

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

14

15 Our most significant variant for ASI, rs1006923 at the *TEX41* locus has previously been
16 reported to be significantly associated with CAD in a mixed population GWAS that included
17 UK Biobank participants²². It is located 129kb upstream and in low linkage disequilibrium (LD;
18 $r^2 = 0.19$) to rs1438896 a variant reported by Warren et al.²⁶ with blood pressure traits.
19 Another variant at the *TEX41* locus rs183032 is 50kb downstream to our lead variant with a
20 moderate LD ($r^2 = 0.38$) is associated with aortic stenosis in an Icelandic cohort²⁷. Helgadottir
21 (2018)²⁷ suggested the tumour growth factor- β (TGF- β) through *ZEB2*²⁷ as a candidate gene.
22 *ZEB2* has a role as a DNA-binding transcriptional repressor that interacts with the main signal
23 transducers for TGF- β receptors (SMADs). Considering that TGF- β is involved in the
24 regulation of vascular smooth muscle differentiation, as well as in the collagen up-regulation
25 in the vascular wall, changes in the expression of TGF- β have the potential to alter arterial

1 stiffness. We note additional support for *ZEB2* from eQTL data in aortic tissue in our analyses
2 (Supplementary Table S4).

3

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

15

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

18

19 At the 4th locus rs10840457 is located near the *MRVI1* (Murine Retrovirus Integration Site 1
20 Homolog, also known as IP3R-associated cGMP kinase substrate (*IRAG*)). The *MRVI1* gene
21 is responsible for encoding the *MRVI1/IRAG* protein, which is present in a number of tissues
22 including aorta and trachea³¹, and is involved in smooth muscle contractility. Specifically,
23 there is inhibition of calcium release from endoplasmic reticulum following co-expression of
24 *IRAG* and cGMP-dependent protein kinase type I β (cGKI β) in the presence of cGMP³¹. In a
25 study using *IRAG*-knockout mice, the authors concluded that signalling of cGKI β via *IRAG* is
26 a vital functional component in the regulation of smooth muscle tone and intracellular calcium

1 by nitric oxide and atrial natriuretic peptide³². At this locus *MRVI1* represents an interesting
2 candidate, but we note this is a locus that was not genome-wide significant in the secondary
3 analyses, thus further validation will be required.

4
5 Gene-based testing by MAGMA revealed *COL4A2* as the most significant gene association
6 with ASI in our cohort, with *FOXO1*, *MRVI1* and *FOXO3* also significant. The lead SNV from
7 the GWAS at the *COL4A2* locus is rs9521719 ($P = 1.1 \times 10^{-7}$). *COL4A2*, along with the
8 adjacent gene *COL4A1*, encodes the protein subunits of type IV collagen forming
9 heterotrimers. Type IV collagen is a vital structural component of basement membranes and
10 mutations in these genes are seen in disorders such as myopathy, intracerebral
11 haemorrhage and glaucoma³³. A previous GWAS reported rs3742207 located near the
12 *COL4A1* to have strong association ($P = 7.08 \times 10^{-7}$) with PWV in a cohort comprising of
13 4,221 Sardinian individuals³⁴. This variant was successfully replicated internally in 1,828
14 individuals and also in 813 Amish individuals. The PWV variant rs3742207 did not show any
15 association with ASI in our cohort ($P = 0.95$). This result may add additional support on there
16 being a difference in the genetic mechanism of ASI compared to PWV, both markers of
17 arterial stiffness.

18
19 We found the genome-wide SNV heritability of ASI to be 6%. This estimate is much lower
20 than the reported heritability for carotid-femoral PWV, which ranged between 0.36-0.40^{14,35}.
21 This dissimilarity might be explained by the differences in stiffness measure, covariates
22 included and populations. In addition, due to the methodological inconsistencies mentioned
23 above, it should be noted that the heritability of PWV would differ depending on the site of
24 measurement. Mitchell et al.¹⁴ reported moderate heritability for carotid-femoral PWV ($h^2 =$
25 0.40) in their study of 1,480 participants in the Framingham Study offspring cohort. However,
26 when using the less common approach of measuring PWV between the carotid and brachial

1 artery, the heritability estimate was lower ($h^2 = 0.09$). In twin studies, the reported heritability
2 estimates range between 0.38 and 0.53^{12,13,36} where the PWV measurements were either
3 located at the wrist (aorto-radial) or foot (aorto-dorsalis-pedis).

4
5 The main strength of our study is the very large sample size that, despite the low estimated
6 heritability for ASI, has enabled the identification of genetic variants not found in previous
7 studies. Mean arterial pressure has been shown to have a strong influence on arterial
8 stiffness¹⁵ so it was included as a covariate in the fully adjusted model, rather than systolic
9 blood pressure (SBP) and diastolic blood pressure (DBP). Despite using the largest cohort
10 to date in identifying genetic variants for arterial stiffness, the main limitation of our study is
11 that loci discovered therein require formal validation in an independent dataset. However, an
12 external study of similar sample size with ASI measurements is currently lacking. We note if
13 we use a more stringent P -value for reporting, $P \leq 1 \times 10^{-8}$, then two loci would remain
14 significant (*TEX41* and *FOXO1*). In addition, our study cohort consisted of middle-aged
15 individuals of European-ancestry, so our findings may not be generalised to other age groups
16 and ethnic populations.

17
18 In conclusion, we identified four loci significantly associated with ASI, an independent
19 predictor of CVD and mortality. The two most significant loci, *TEX41* and *FOXO1*, have SNV
20 associations that may alter arterial stiffness through blood pressure regulation and vascular
21 smooth muscle differentiation. Our results also suggest an important role of calcium in the
22 regulation of smooth muscle tone contributing to arterial stiffness. Further research will be
23 necessary to validate our discovered loci in a separate cohort and their confirmation can
24 potentially lead to the development of risk models that can be used in clinical practice.

1 **METHODS**

2

3 **Study population**

4 The UK Biobank is a large population-based prospective study of >500,000 participants aged
5 40-69 years at baseline recruited between 2006 and 2010 in England, Scotland and Wales.
6 During the initial assessment visit, a broad range of biochemical, clinical and genotype data
7 were collected and participants had a number of physical measurements. The UK Biobank
8 study was approved by the North West Multi-Centre Research Ethics Committee and all
9 enrolled individuals have provided written informed consent for collection, storage, make
10 availability of their data for health-related research. All methods were carried out in
11 accordance with the relevant guidelines and regulations.

12

13 **Arterial stiffness index measures**

14 Pulse wave ASI (UK Biobank Field 21021), measured in m/s, was derived using the pulse
15 waveform obtained at the finger (preferably index finger of the non-dependent hand though
16 can be placed on any finger or thumb) with an infra-red sensor (PulseTrace PCA 2™,
17 CareFusion, USA). The shape of the volume waveform in the finger is directly related to the
18 time it takes for the pulse waveforms to travel through the arterial tree in the lower body and
19 to be reflected back to the finger. Measurements were made by clipping the device to a finger
20 and the reading is made over 10-15 seconds. As the participants' heights were unknown until
21 after the recording of the waveform data, the actual ASI values were calculated ($ASI =$
22 $height/PPT$) by UK Biobank outside the assessment centre visit.

23

24 Data were taken at the UK Biobank Assessment Centres during the baseline recruitment
25 between 29th April 2009 and 1st October 2010 inclusive (Nmax = 169,822). UK Biobank
26 participants are free to withdraw at any time and so 15 individuals were removed based on

1 the application specific list of anonymised IDs. Individuals with absence of notch position in
2 the pulse waveform (Field 4204) were excluded (n = 25,288) as the notch in the digital volume
3 pulse waveform indicates the reflected component of the pressure transmitted and is
4 therefore required to calculate PPT. Outlier ASI values, defined as three inter-quartile ranges
5 below the first quartile or above the third quartile (n = 81), were also removed from analyses.
6 Finally, participants with missing height, weight or blood pressure measurements were
7 removed from the final analyses (n = 848).

8

9 **Covariates**

10 Blood pressure was measured using the Omron 705 IT electronic blood pressure monitor
11 (OMRON Healthcare Europe B.V. Kruisweg 577 2132 NA Hoofddorp). SBP and DBP were
12 derived as the mean of the two recorded automated measurements (UK Biobank Fields 4079,
13 4080), except for 1,141 individuals who only had one recorded reading. MAP estimation was
14 calculated using the traditional formula: $MAP = DBP + 1/3(SBP-DBP)$. Height (Field 50) was
15 measured using a Seca 202 device (Seca, Birmingham, UK). For individuals who reported
16 use of anti-hypertensive medications through questionnaires in UK biobank (Fields 6177,
17 6153), their SBP and DBP were adjusting by adding 15mmHg and 10mmHg respectively to
18 the mean recorded readings^{26,37}. The ID numbers of the arterial stiffness devices used were
19 obtained from Field 4206 in UK Biobank.

20

21 **Genotypic data**

22 Central QC and imputation of genotypic data performed by UK Biobank has been previously
23 described³⁸. Briefly, genotypic data was obtained through either UK Biobank Axiom or UK
24 BiLEVE Axiom arrays (Affymetrix Research Service Laboratory, Santa Clara, California, USA).
25 The Haplotype Reference Consortium (HRC) and the merged UK10K sequencing + 1000
26 Genomes were used as a reference panels for imputation with preference for the HRC panel

1 where SNPs were present in both panels. This study utilised the refreshed genetic dataset
2 made available by UK Biobank in July 2017³⁸.

3

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

9

10 A 4-way *k*-means clustering analysis was performed according to data from the first and
11 second principal components (PC1 and PC2) using the 'pvclust' R package (version 2.0-0)³⁹
12 to objectively identify the main ethnic groups (White, Asian, African and Chinese) within UK
13 Biobank. "White" participants were defined for those present in the "White" cluster for both
14 PC1 and PC2 analyses. They also needed to match their self-reported ancestry though
15 "mixed, "other" and "missing" were treated as being broad ethnicity. After restriction to
16 European-ancestry only, 127,121 individuals remained. Lastly, we performed SNV-level QC
17 to exclude 237,751 SNVs from genotyped SNVs in these individuals with the following
18 thresholds: minor allele frequency (MAF) of 1%, Hardy-Weinberg equilibrium (*P*-value of $1 \times$
19 10^{-6}) and missingness of 0.015 using PLINK 1.9⁴⁰. After exclusions, a final total of 546,505
20 model SNVs was generated for our GWAS analyses.

21

22 **Genome-wide association and heritability analyses**

23 We applied rank-based INT on the residuals from the regression of ASI against the phenotypic
24 covariates (i.e., all except genotyping array and PCs) before performing genetic analyses as
25 the distribution did not approximate a normal distribution (Supplementary Fig. S5). The

1 heritability of ASI explained by additive genetic variation was estimated using a variance
2 components method (BOLT-REML)⁵¹.

3

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

11

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

17

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

22

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

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

6

7 **Functional annotation**

8 Summary statistics from our primary model GWAS analyses were uploaded onto the FUMA
9 (v1.3.4c)¹⁸ web-based application (<http://fuma.ctglab.nl/>) in order to perform functional
10 annotations via its SNP2GENE function. The default FUMA settings were used whereby
11 significant SNVs ($P < 5 \times 10^{-8}$), and those that were in LD ($r^2 \geq 0.6$), with a MAF ≥ 0.01 were
12 selected for further annotation. In addition, the maximum distance between LD blocks to
13 merge into a locus was set at $< 250\text{kb}$ and the UKB release 2 European was used as the
14 reference panel population. The functional consequences for these SNVs were then obtained
15 with ANNOVAR⁴⁴. For gene-based analysis, the integrated MAGMA v1.06¹⁹ was utilised in
16 which SNVs are mapped to a gene according their genomic location before gene-level
17 associations with ASI were tested. Here, a P -value with Bonferroni adjustment of $0.05/18697$
18 $= 2.67 \times 10^{-6}$ was used to define genome-wide significance. For eQTL mapping, this was
19 implemented by the FUMA platform, GTEx v7 database⁴⁵ (<http://gtexportal.org/home/>), 53
20 tissue types was used as the gene expression reference data and the significant threshold
21 was defined as false discovery rate (FDR) < 0.05 . Genome-wide significant SNVs and those
22 in LD with $r^2 > 0.8$ were examined for previously reported GWAS associations with other traits
23 using PhenoScanner²⁰, an online search tool containing over 350 million publicly available
24 association results. The significance threshold chosen for reporting therein was P -value < 5
25 $\times 10^{-8}$.

1 **Data Availability**

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

7

1 REFERENCES

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

1. 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).
2. Scuteri, A. *et al.* Longitudinal perspective on the conundrum of central arterial stiffness, blood pressure, and aging. *Hypertension* **64**, 1219–1227 (2014).
3. Dernellis, J. & Panaretou, M. Aortic stiffness is an independent predictor of progression to hypertension in nonhypertensive subjects. *Hypertension* **45**, 426–431 (2005).
4. Liao, D. *et al.* Arterial stiffness and the development of hypertension. The ARIC study. *Hypertension* **34**, 201–6 (1999).
5. Liao, J. & Farmer, J. Arterial stiffness as a risk factor for coronary artery disease. *Curr. Atheroscler. Rep.* **16**, 387 (2014).
6. 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).
7. Laurent, S. *et al.* Aortic stiffness is an independent predictor of all-cause and cardiovascular mortality in hypertensive patients. *Hypertension* **37**, 1236–41 (2001).
8. Millasseau, S. C. *et al.* Noninvasive assessment of the digital volume pulse: Comparison with the peripheral pressure pulse. *Hypertension* **36**, 952–956 (2000).
9. 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).
10. 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).

- 1 11. Said, M. A., Eppinga, R. N., Lipsic, E., Verweij, N. & van der Harst, P. Relationship of
2 arterial stiffness index and pulse pressure with cardiovascular disease and mortality. *J.*
3 *Am. Heart Assoc.* **7**, e007621 (2018).
- 4 12. Ge, D. *et al.* Heritability of arterial stiffness in black and white American youth and young
5 adults. *Am. J. Hypertens.* **20**, 1065–72 (2007).
- 6 13. Tarnoki, A. D. *et al.* Heritability of central blood pressure and arterial stiffness. *J.*
7 *Hypertens.* **30**, 1564–1571 (2012).
- 8 14. Mitchell, G. F. *et al.* Heritability and a genome-wide linkage scan for arterial stiffness,
9 wave reflection, and mean arterial pressure: The Framingham heart study. *Circulation*
10 **112**, 194–199 (2005).
- 11 15. Sayed-Tabatabaei, F. A. *et al.* Heritability of the function and structure of the arterial
12 wall: findings of the Erasmus Rucphen Family (ERF) study. *Stroke* **36**, 2351–2356
13 (2005).
- 14 16. Levy, D. *et al.* Framingham Heart Study 100K Project: genome-wide associations for
15 blood pressure and arterial stiffness. *BMC Med. Genet.* **8 Suppl 1**, S3 (2007).
- 16 17. Logan, J. G., Engler, M. B. & Kim, H. Genetic determinants of arterial stiffness. *J*
17 *Cardiovasc. Transl. Res.* **8**, 23–43 (2015).
- 18 18. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and
19 annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
- 20 19. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-
21 set analysis of GWAS data. *PLOS Comput. Biol.* **11**, e1004219 (2015).
- 22 20. Staley, J. R. *et al.* PhenoScanner: a database of human genotype-phenotype
23 associations. *Bioinforma. Oxf. Engl.* **32**, 3207–3209 (2016).
- 24 21. Abbott, L. *et al.* UK Biobank — Neale lab. (2018). Available at: <http://www.nealelab.is/uk->
25 [biobank/](http://www.nealelab.is/uk-biobank/). (Accessed: 2nd January 2019)

- 1 22. van der Harst, P. & Verweij, N. Identification of 64 novel genetic loci provides an
2 expanded view on the genetic architecture of coronary artery disease. *Circ. Res.* **122**,
3 433–443 (2018).
- 4 23. Mitchell, G. F. *et al.* Common genetic variation in the 3'-BCL11B gene desert is
5 associated with carotid-femoral pulse wave velocity and excess cardiovascular disease
6 risk: the AortaGen Consortium. *Circ. Cardiovasc. Genet.* **5**, 81–90 (2012).
- 7 24. Park, S. *et al.* Lack of association between arterial stiffness and genetic variants by
8 genome-wide association scan. *Blood Press.* **24**, 258–261 (2015).
- 9 25. Laurent, S. *et al.* Expert consensus document on arterial stiffness: methodological
10 issues and clinical applications. *Eur. Heart J.* **27**, 2588–2605 (2006).
- 11 26. Warren, H. R. *et al.* Genome-wide association analysis identifies novel blood pressure
12 loci and offers biological insights into cardiovascular risk. *Nat. Genet.* **49**, 403–415
13 (2017).
- 14 27. Helgadóttir, A. *et al.* Genome-wide analysis yields new loci associating with aortic valve
15 stenosis. *Nat. Commun.* **9**, 987 (2018).
- 16 28. UK Biobank. UK Biobank arterial pulse-wave velocity. (2011). Available at:
17 <https://biobank.ctsu.ox.ac.uk/crystal/docs/Pulsewave.pdf>. (Accessed: 7th January 2019)
- 18 29. Qi, Y. *et al.* Novel mechanism of blood pressure regulation by forkhead box class O1-
19 mediated transcriptional control of hepatic angiotensinogen. *Hypertension* **64**, 1131–40
20 (2014).
- 21 30. Zhu, L.-H. *et al.* Mindin regulates vascular smooth muscle cell phenotype and prevents
22 neointima formation. *Clin. Sci. Lond. Engl. 1979* **129**, 129–45 (2015).
- 23 31. Schlossmann, J. *et al.* Regulation of intracellular calcium by a signalling complex of
24 IRAG, IP3 receptor and cGMP kinase I β . *Nature* **404**, 197–201 (2000).
- 25 32. Desch, M. *et al.* IRAG determines nitric oxide- and atrial natriuretic peptide-mediated
26 smooth muscle relaxation. *Cardiovasc. Res.* **86**, 496–505 (2010).

- 1 33. Kuo, D. S., Labelle-Dumais, C. & Gould, D. B. COL4A1 and COL4A2 mutations and
2 disease: insights into pathogenic mechanisms and potential therapeutic targets. *Hum.*
3 *Mol. Genet.* **21**, R97-110 (2012).
- 4 34. Tarasov, K. V *et al.* COL4A1 is associated with arterial stiffness by genome-wide
5 association scan. *Circ. Cardiovasc. Genet.* **2**, 151–8 (2009).
- 6 35. Sayed-Tabatabaei, F. A. *et al.* Heritability of the function and structure of the arterial
7 wall: Findings of the Erasmus Rucphen Family (ERF) study. *Stroke* **36**, 2351–2356
8 (2005).
- 9 36. Cecelja, M. *et al.* Arterial stiffening is a heritable trait associated with arterial dilation but
10 not wall thickening: a longitudinal study in the twins UK cohort. *Eur. Heart J.* **39**, 2282–
11 2288 (2018).
- 12 37. Tobin, M. D., Sheehan, N. A., Scurrah, K. J. & Burton, P. R. Adjusting for treatment
13 effects in studies of quantitative traits: antihypertensive therapy and systolic blood
14 pressure. *Stat. Med.* **24**, 2911–2935 (2005).
- 15 38. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data.
16 *Nature* **562**, 203–209 (2018).
- 17 39. Suzuki, R. & Shimodaira, H. Pvcust: an R package for assessing the uncertainty in
18 hierarchical clustering. *Bioinformatics* **22**, 1540–1542 (2006).
- 19 40. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer
20 datasets. *GigaScience* **4**, 7 (2015).
- 21 41. Loh, P.-R. *et al.* Efficient Bayesian mixed-model analysis increases association power in
22 large cohorts. *Nat. Genet.* **47**, 284–290 (2015).
- 23 42. R Development Core Team. R: A language and environment for statistical computing. R
24 foundation for statistical computing, Vienna, Austria. (2016). Available at: [http://www.r-](http://www.r-project.org/)
25 [project.org/](http://www.r-project.org/).

- 1 43. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics
2 identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–375 (2012).
- 3 44. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants
4 from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164–e164 (2010).
- 5 45. Aguet, F. *et al.* Genetic effects on gene expression across human tissues. *Nature* **550**,
6 204–213 (2017).

7

8

1 **FIGURES AND TABLES**

2

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

6

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

8

9

10

Table 1 Summary of genome-wide significant loci associated with arterial stiffness index

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
MRVI1	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

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

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 statement

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

Competing interests

The authors declare that they have no competing financial or non-financial interests.