



Published in final edited form as:

*Circulation*. 2018 December 18; 138(25): 2869–2880. doi:10.1161/CIRCULATIONAHA.118.035584.

## Clinical and Genetic Determinants of Varicose Veins: A Prospective, Community-Based Study of ~500,000 Individuals

Eri Fukaya, MD, PhD<sup>#1</sup>, Alyssa M. Flores, BS<sup>#1</sup>, Daniel Lindholm, MD, PhD<sup>2,3</sup>, Stefan Gustafsson, PhD<sup>4</sup>, Daniela Zanetti, PhD<sup>5</sup>, Erik Ingelsson, MD, PhD<sup>5,6</sup>, and Nicholas J. Leeper, MD<sup>#1,5,6</sup>

<sup>1</sup>Department of Surgery, Division of Vascular Surgery, Stanford University School of Medicine, Stanford, California, USA. From Stanford University School of Medicine, Stanford, CA (EF, AMF, DZ, EI, NJL); Uppsala University, Uppsala, Sweden (DL, SG).

<sup>2</sup>Department of Medical Sciences, Cardiology, Uppsala University, Uppsala, Sweden. From Stanford University School of Medicine, Stanford, CA (EF, AMF, DZ, EI, NJL); Uppsala University, Uppsala, Sweden (DL, SG).

<sup>3</sup>Uppsala Clinical Research Center, Uppsala, Sweden. From Stanford University School of Medicine, Stanford, CA (EF, AMF, DZ, EI, NJL); Uppsala University, Uppsala, Sweden (DL, SG).

<sup>4</sup>Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden. From Stanford University School of Medicine, Stanford, CA (EF, AMF, DZ, EI, NJL); Uppsala University, Uppsala, Sweden (DL, SG).

<sup>5</sup>Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA, USA. From Stanford University School of Medicine, Stanford, CA (EF, AMF, DZ, EI, NJL); Uppsala University, Uppsala, Sweden (DL, SG).

<sup>6</sup>Stanford Cardiovascular Institute, Stanford, CA, USA. From Stanford University School of Medicine, Stanford, CA (EF, AMF, DZ, EI, NJL); Uppsala University, Uppsala, Sweden (DL, SG).

# These authors contributed equally to this work.

### Abstract

**Background:** Varicose veins are a common problem with no approved medical therapies. While it is believed that varicose vein pathogenesis is multifactorial, there is a limited understanding of the genetic and environmental factors that contribute to their formation. Large-scale studies of risk factors for varicose veins may highlight important aspects of pathophysiology and identify groups at increased risk for disease.

**Methods:** We applied machine learning to agnostically search for risk factors of varicose veins in 493,519 individuals in the UK Biobank. Predictors were further studied using univariable and

---

**Corresponding author:** Nicholas J. Leeper, MD, Divisions of Vascular Surgery and Cardiovascular Medicine, Stanford University, 300 Pasteur Drive, Room H3638, Stanford, CA 94305, USA. nleeper@stanford.edu Phone: +1-650-725-5227 @alyssaflo, @ErikIngelsson, @stanfordvasc.

#### Disclosures

DL reports institutional research grants from GlaxoSmithKline and AstraZeneca, unrelated to the present project. EI is a scientific advisor for Precision Wellness and Olink Proteomics for work unrelated to the present project.

multivariable Cox regression analysis (2,441 incident events). A genome-wide association study (GWAS) of varicose veins was also performed among 337,536 unrelated individuals (9,577 cases) of white British descent, followed by eQTL and pathway analyses. Because height emerged as a new candidate risk factor, we performed Mendelian randomization analyses to assess a potential causal role for height in varicose vein development.

**Results:** Machine learning confirmed several known (age, sex, obesity, pregnancy, history of deep vein thrombosis) and identified several new risk factors for varicose vein disease, including height. After adjusting for traditional risk factors in Cox regression, greater height remained independently associated with varicose veins (HR for upper vs. lower quartile: 1.74; 95% CI: 1.51–2.01,  $P<0.0001$ ). A GWAS identified 30 new genome-wide significant loci, identifying pathways involved in vascular development and skeletal/limb biology. Mendelian randomization analysis provided evidence that increased height is causally related to varicose veins (Inverse-variance weighted: OR=1.26;  $P=2.07\times 10^{-16}$ ).

**Conclusions:** Using data from nearly half a million individuals, we present a comprehensive genetic and epidemiological study of varicose veins. We identified novel clinical and genetic risk factors which provide pathophysiological insights and could help future improvements of treatment of varicose vein disease.

## Keywords

Varicose veins; epidemiology; genetics; Genome Wide Association Study

## Introduction

Varicose veins are a common manifestation of chronic venous disease. It is estimated that more than 30 million adults in the United States have varicose veins, with interventions consuming more than \$1 billion in direct health care resources per year.<sup>1</sup> While sometimes dismissed as a cosmetic finding, varicose veins can culminate with life-limiting ulceration in up to 20% of affected individuals and have been increasingly associated with serious health risks.<sup>2</sup> Recent studies show that patients with varicose veins are up to 5-fold increased risk for developing deep vein thrombosis (DVT).<sup>3,4</sup> Moreover, the association of varicose veins with peripheral arterial disease and additional vascular diseases is also being elucidated.<sup>3,5</sup>

The factors influencing varicose vein formation are not fully understood. Epidemiological studies have established multiple risk factors such as age, female sex, pregnancy, obesity and prior DVT<sup>6–8</sup>; but several other putative factors remain unconfirmed. There is also a clear familial component to varicose vein disease, but prior genetic studies have been small in size and have provided conflicting results.<sup>9–12</sup> Accordingly, it remains difficult to identify those at risk and our lack of biological insights into pathogenesis may help explain why no approved therapies exist to prevent or delay disease progression.

To address the need for large-scale studies of varicose vein disease, we investigated the clinical, environmental, and genetic basis of varicose veins in up to 502,619 individuals from the UK Biobank, using a constellation of modalities including: (1) agnostic machine learning; (2) traditional epidemiological analyses; (3) genetic studies including a genome-

wide association study and pathway analyses; and (4) Mendelian randomization to address causality.

## Methods

All data reported in this article are made available to other researchers via application to the UK Biobank for purposes of reproducing the results or replicating the procedure. The data are publicly available at the UK Biobank repository.<sup>13</sup>

### Study design and participants:

We used data from the UK Biobank, a longitudinal cohort study of 502,619 individuals aged 40–69 years from 21 centers across the United Kingdom. Between 2006–2010, individuals aged 40–69 underwent baseline assessment and have since been carefully investigated for extensive phenotype, lifestyle, health outcomes, and genetic data. The UK Biobank protocol has been described in detail previously<sup>13</sup> and can be found online (<https://www.ukbiobank.ac.uk>).

In our observational analyses, we excluded individuals with prevalent varicose veins (n=9,100), leaving 493,519 individuals eligible. In the GWAS, we studied unrelated individuals of White British ancestry with available phenotype data (n=337,536). The UK Biobank study was approved by the North West Multi-Centre Research Ethics Committee and all participants provided written informed consent to participate in the UK Biobank.

### Definitions of potential risk factors:

To identify risk factors for varicose veins, we considered all variables from the UK Biobank, spanning domains of demographics, socioeconomic data, lifestyle behaviors, medical history, and a range of physical and cognitive measures. We excluded administrative, temporary, and unspecific ICD codes [chapters U, Z], as well as diagnoses likely to represent varicose veins, including phlebitis, thrombophlebitis, and other diseases of veins and lymphatic vessels, [ICD-10 I80–I82, I85–I89]).

A total of 2,715 predictor variables were extracted from the UK Biobank database and used in our phenome-wide scan of varicose vein risk factors (Supplemental Table 1). Medical history and smoking status were assessed using self-reported questionnaires. Blood pressure and body mass index (BMI) were measured in a standard fashion (full details available at UK Biobank data showcase: <http://biobank.ctsu.ox.ac.uk/crystal/>). Systolic blood pressure was defined as the mean of two measurements.

### Follow-up and outcome:

The primary outcome was the diagnosis of varicose veins, defined by ICD-9 (454\*) and/or ICD-10 (I-83\*). For observational analyses, prevalent cases were excluded (n=9,100), leaving 2,441 cases of incident varicose veins during a median follow-up of 6.2 years (maximum 9.1 years). For GWAS and MR analyses, both incident and prevalent events were included (n=11,541). After exclusion of individuals due to genetic relatedness or non-European descent (n=1,964), 9,577 cases were included.

**Machine learning analyses:**

For agnostic discovery of novel risk factors for varicose veins, we used a gradient boosting machine (GBM) model. Based on the input predictor variables, this machine learning algorithm consecutively fits new decision trees to provide a more accurate prediction of a response variable (i.e. outcome, varicose veins). The principle idea of this algorithm is the learning procedure which results in consecutive error-fitting, with each decision tree chosen to minimize the loss function.<sup>14</sup> In addition to classification, GBM ranks the importance (weight) of input variables to the outcome based on decision trees. We used incident cases of varicose veins as the response variable. Results are presented as variable importance. All analyses were conducted using machine learning platform (H2O 3.10.4.1). Detailed descriptions of the GBM model can be found in the Supplemental Methods.

**Statistical Analysis:**

We analyzed the associations of traditional varicose vein risk factors and the top 10 predictors identified through GBM on incident varicose veins using uni- and multivariable Cox regression analyses. For categorical predictors, Kaplan-Meier estimates were used. For continuous predictors, the predictor was entered as a restricted cubic spline to account for any non-linear associations. The top 10 novel predictors were further assessed with two multivariable-adjusted Cox models: a) adjusting for age and sex; and b) additionally adjusting for established varicose vein risk factors, including BMI, waist-hip ratio (WHR), history of DVT, and pregnancy.<sup>1,2,6-8</sup> As a secondary analysis, we assessed variables in a third model that included the first five principal components to account for genetic ancestry. The discriminatory ability of the models was assessed using c-indices. All regression analyses were conducted using R 3.3.0 and 3.3.3.

**GWAS of varicose veins:**

A GWAS was performed using imputed genotype data from the second data release of the UK Biobank (9,577 cases and 327,959 controls). A detailed description of genotyping, quality control and imputation may be found in the Supplemental Methods.

We included 10,972,371 markers belonging to the Haplotype Reference Consortium 1.1 subset with a minor allele count  $\geq 20$  in cases and controls and MaCH  $r^2 \geq 0.8$ . Assuming an additive model, the associations between genotype dosages of each marker and varicose veins were tested in logistic regression (Firth's penalized logistic regression in non-convergence) using PLINK2<sup>15</sup> and adjusting for age, sex, and the first 15 principal components of ancestry. A  $P < 5 \times 10^{-8}$  was considered to be statistically significant. Independent lead single nucleotide polymorphisms (SNPs) were identified by conditional analysis. Regional plots were created for the association test results of significant loci using Locuszoom v1.4. All other plots and tables were created in R 3.3.0.

**LD score regression:**

LD score regression (LDSC) is a tool used to estimate SNP-heritability, genetic correlation, and genomic inflation using GWAS summary datasets.<sup>16</sup> Because height appeared as a novel predictor in our analyses, we applied LDSC to estimate the genetic correlation between height and varicose veins using summary statistics from our GWAS for varicose veins and

publically available GWAS results for height from the GIANT consortium.<sup>17</sup> As a comparison, we also examined the genetic correlations with peripheral arterial disease (PAD) and traditional risk factors for varicose veins: BMI, WHR, and history of DVT. For BMI and WHR, we used GWAS summary data from the GIANT consortium limited to individuals of European ancestry.<sup>18,19</sup> For history of DVT and PAD, GWAS results from the UK Biobank were used.<sup>20</sup> We additionally applied LDSC to estimate SNP-heritability and test for genomic inflation (expected to some degree for polygenic traits and large sample sizes). SNP-heritability is reported on the liability scale to take into account sample prevalence of varicose vein disease.

### eQTL analysis

To explore the potential functional significance of our findings, we evaluated all genome-wide significant loci for evidence of expression quantitative loci (eQTL). Lookups were performed in the Genotype-Tissue expression GTEx database<sup>21</sup> (53 tissue types) and the FUMA GWAS platform was used to interrogate whole blood data repositories, BIOSQTL and blood eQTL browser.<sup>22</sup>

### Pathway and functional analysis:

We used two data-driven integrative platforms to perform pathway analyses. After clumping to identify independent loci using PLINK2<sup>15</sup>, we performed analysis using DEPICT.<sup>23</sup> We also used the FUMA GWAS platform to investigate loci for pathway and tissue enrichment.<sup>22</sup> A detailed description of pathway analysis can be found in the Supplemental Methods.

### Height and varicose veins: Mendelian randomization analysis

We performed a two-sample Mendelian randomization (MR) analysis using height-associated SNPs from the GIANT study, a GWAS meta-analysis of height in 253,288 individuals of European ancestry.<sup>17</sup> A genetic risk score (GRS) was created using height-associated SNPs as the instrumental variable, and our GWAS of varicose veins as the outcome. Our main GRS instrument included 512 SNPs without evidence for pleiotropy (See Supplemental Methods). Analyses were conducted with the R package TwoSampleMR. Power was calculated using the online tool (<https://sb452.shinyapps.io/power/>).

## Results

### Machine learning and Cox analyses

Baseline characteristics of the study sample are shown in Table 1. The 20 variables with the highest importance in the GBM model are presented in Figure 1. Age, pregnancy, history of DVT, and obesity are previously reported risk factors that were again observed in our study (Supplemental Table 2). In addition, several novel predictors of varicose veins were also identified, including leg bioimpedance, height, and history of surgery on leg arteries or other major operation. Our GBM model had similar ROC-AUCs in the training, validation, and test sets: 0.76, 0.71, and 0.70, with each set having an MSE 0.02, indicating the model had acceptable discriminatory performance and low error. A number of putative predictors which have been inconsistently reported in the literature, including hypertension (variable 238), age

at menopause (variable 405), and smoking (variable 487) were not observed among the top predictors by our machine learning approaches.

The top 10 new predictors were selected, together with previously established risk factors, for additional analysis and validation by two Cox models (age- and sex-adjusted; and adjusting for established varicose vein risk factors). Among the independently associated variables were established risk factors, such as history of DVT (HR: 2.60; 95% CI: 2.18–3.11;  $P<0.0001$ ) and BMI (HR: 1.10; 95% CI: 1.03–1.17;  $P=0.001$ , Table 2) but also novel markers, such as leg bioimpedance, greater height, and surgery on leg arteries (Figure 2, Table 2, Supplemental Figure 1). Each new predictor from our GBM model improved the predictive capacity of the Cox models when added to models including age, sex, and traditional risk factors, and remained independently associated in our secondary analysis that also adjusted for genetic ancestry (Supplemental Table 3). The largest improvement in C-index was observed for leg bioimpedance and height (from 0.56 to 0.65 and 0.61, respectively; Supplemental Table 4). Of note, debated risk factors, such as oral contraceptive use, was not found to be independently associated with varicose veins in Cox analysis; while some were, including systolic blood pressure.

### GWAS for varicose veins:

We conducted a GWAS to identify genetic variants that influence varicose vein susceptibility. Overall, the GWAS results showed modest deviation in test statistics ( $\lambda=1.09$ ). The constructed quantile-quantile (Q-Q) plot showed minimal inflation, except for in the upper tail of distribution (Supplemental Figure 2), suggesting adequate correction for population stratification. LDSC supported this and provided evidence that inflation was due to a polygenic signal in which heritability accounted for 28% of the population variance in varicose vein disease (LDSC intercept=1.032, SE=0.0088;  $h^2=0.28$ , SE=0.024).

We identified 855 new SNPs associated with varicose veins that exceeded genome-wide significance (Figure 3). Following conditional analyses, we found 30 independent genetic variants associated with varicose vein disease (Supplemental Table 5). The strongest association was located on chromosome 1 in the *CASZ1* gene, in a known blood pressure locus (rs11121615,  $P=3.71\times 10^{-65}$ ).<sup>24</sup> Other lead variants were located in the 16q24 region, in or near vascular mechanosensory channel *PIEZO1* and galactosamine-sulfatase enzyme *GALNS* (rs2911461,  $P=4.81\times 10^{-29}$ ; rs8053350  $P=8.93\times 10^{-18}$ ). Regional plots for the 30 newly discovered loci are shown in Supplemental Figure 3. Full results for all SNPs with  $P<5\times 10^{-8}$  are shown in Supplemental Table 6.

### Pathway and functional analysis

eQTL analyses identified several genes significantly regulated by varicose vein SNPs. For example, one of the lead variants, rs2861819, had several eQTL associations, including *PNO1*, *WDR92*, *PLEK*, and *PPP3R1*, which expresses calcineurin, a critical signaling component during angiogenesis.<sup>25</sup> We also identified significant eQTLs for lead variants rs2911463 (*PIEZO1*,  $P=1.50\times 10^{-5}$ ; *GALNS*,  $P=1.50\times 10^{-5}$ ) and rs3101725 (*FBN2*,  $P=9.50\times 10^{-7}$ ). These loci are particularly interesting because mutations in *GALNS* and *FBN2* have been linked to a hereditary form skeletal dysplasia<sup>26</sup> and Marfan-like syndrome



Congenital Contractural Arachnodactyly<sup>27</sup>, respectively. Detailed results from eQTL analysis are given in Supplemental Table 7.

FUMA prioritized tissues, pathways, and gene sets enriched in varicose vein loci. Tissue enrichment was greatest for adipose tissue ( $P=7.58\times 10^{-5}$ , Supplemental Table 8). Evaluation for enriched biological processes identified several gene sets related to vascular development and structure (vasculature development:  $P=1.96\times 10^{-6}$ ; endothelial cell differentiation  $P=8.95\times 10^{-6}$ ), and highlighted canonical pathways that may influence varicose vein disease: Gata3 pathway ( $P=3.06\times 10^{-6}$ ), VEGF signaling ( $P=1.11\times 10^{-5}$ ), and NFAT pathway ( $P=1.29\times 10^{-4}$ , Supplemental Table 9).

In addition, FUMA analysis also revealed that our lead SNPs were previously associated with a variety of traits, including varicose vein predictors identified in our observational analyses: WHR ( $P=2.79\times 10^{-4}$ ) and height ( $P=2.30\times 10^{-3}$ ). No significantly enriched gene sets were identified by DEPICT that met the FDR cut-off  $<0.05$ .

### Genetic correlation between height and varicose veins:

After integrating data from our GWAS and publicly available GWAS summary results on height, we found evidence for a strong genetic correlation between height and varicose veins ( $r_g=0.16$ ;  $SE=0.031$ ;  $P=2.99\times 10^{-7}$ ). This suggests that there is approximately 16% genetic overlap between height and varicose veins, which was larger than the estimated genetic correlation with BMI ( $r_g=0.095$ ,  $SE=0.032$ ,  $P=0.00260$ ) and WHR ( $r_g=0.078$ ,  $SE=0.052$ ,  $P=0.131$ ). History of DVT was found to have the greatest genetic correlation with varicose vein disease ( $r_g=0.36$ ;  $SE=0.0805$ ;  $P=5.57\times 10^{-7}$ ), as compared to PAD ( $r_g=0.20$ ,  $SE=0.22$ ,  $P=0.37$ ) and traditional factors such as BMI. The estimated genetic correlation of varicose veins and PAD was lower than that of DVT and was non-significant.

### Mendelian randomization

Because we observed a strong association of height with varicose veins, we next did multiple MR analyses to investigate causal effects. We used a GRS instrument composed of 512 independent SNPs previously associated with height<sup>17</sup>, without evidence of pleiotropy. The results provided strong evidence of a causal effect of height on varicose veins (Inverse-variance weighted, IVW:  $OR=1.26$ ,  $P=2.07\times 10^{-16}$ ; Weighted median:  $OR=1.28$ ,  $P=1.79\times 10^{-8}$ ; Maximum likelihood:  $OR=1.26$ ,  $P=9.12\times 10^{-18}$ ; MR Egger:  $OR=1.43$ ,  $P=6.74\times 10^{-6}$ , Supplemental Table 10). These results remained consistent in a leave-one-out sensitivity analysis and a funnel plot did not indicate any directional pleiotropy (Supplemental Figure 4–5). Sensitivity analyses including all 536 variants (without exclusion of potentially pleiotropic variants) provided almost identical results, but with a larger degree of heterogeneity as expected (Supplemental Table 10). At an alpha threshold of 0.05, power for MR analyses was estimated to be 100% with an OR of 1.51 per SD-increase in height (corresponding to the effect from our observational analyses).

### Discussion

Leveraging data from 493,519 individuals, this study is the largest to investigate the genetic and environmental basis of varicose veins. First, we used ‘hypothesis-free’, machine

learning approaches to identify clinical and environmental predictors of disease. This powerful method confirmed a number of well-established relationships, and identified a variety of novel risk factors which have not been previously reported. Second, we used multivariable Cox regression methods to validate our epidemiological observations, and confirm their importance for predicting disease risk. Third, we performed a GWAS and pathway analyses to discover genetic determinants of varicose veins. Finally, we used Mendelian randomization to establish height as a causal risk factor for disease.

### **Environmental predictors of varicose veins:**

Prior studies have yielded inconsistencies regarding environmental predictors of varicose veins. Machine learning confirmed several of the more accepted risk factors (e.g. age, history of DVT, obesity, pregnancy)<sup>1,2,6-8</sup>, in addition to providing insight on previously debated predictors. Age when completed full-time education (reflecting length of education) was among the most important variables identified in the GBM model, which supports the association between education and varicose veins that has been observed in smaller studies.<sup>8,28</sup> It is possible that this association is related to the ability to obtain medical treatment, as level of education may reflect socioeconomic circumstances and access to healthcare. Individuals with higher education may be more inclined to seek and receive medical care for varicose veins, and thus receive a formal diagnosis. Previous studies have also found a significant association with hormone replacement therapy,<sup>2,28</sup> which again appeared as an important variable in our GBM model and was positively related to incident varicose veins in multivariable analysis. Supplemental estrogen may alter venous wall compliance, as a previous study showed that elevated levels of endogenous estradiol are associated with increased venous distensibility.<sup>29</sup>

On the contrary, machine learning did not prioritize several other controversial factors (e.g. oral contraceptive use, Table 2)<sup>2,6,8,28</sup>, which remained consistent in subsequent formal Cox regression analyses (investigating one risk factor at the time). Given the size of this study, it is likely that these previously debated factors have limited utility as traditional risk discriminators.

Physical activity, smoking, and hypertension are other debated risk factors for varicose vein disease. While duration of moderate daily activity was not found to be a strong predictor in our GBM model and Cox analyses, days per weeks walked 10+ minutes was positively associated. This difference may reflect a strong influence of immobility and non-ambulatory status on disease, rather than routine aerobic exercise per se. Some studies have reported higher smoking rates in subjects with varicose veins, compared to those without disease.<sup>6,30</sup> However, the majority of studies have observed no association between varicose veins and smoking.<sup>31-35</sup> In our study, smoking was not associated with varicose vein development after adjusting for traditional risk factors, but became marginally associated in our secondary analysis that adjusted for genetic ancestry. Studies of the association between varicose veins and hypertension have also shown various relationships, including a positive association<sup>6</sup>, no association<sup>7</sup>, and a lower prevalence of hypertension among subjects with varicose veins.<sup>8</sup> In multivariable analysis, we observed a slight inverse association between systolic BP and varicose veins. Given this modest association of varicose veins with both smoking and blood



pressure, which could reflect reverse causation or confounding, more information is needed to evaluate their potential shared pathophysiology with varicose veins.

Because our machine learning algorithms are not restricted to *a priori* hypotheses, they had the capacity to identify novel factors associated with disease. By using this ‘agnostic’ phenome-wide approach, we were able to identify several new strong predictors, including leg bioimpedance and height. Interestingly, height was suggested to be a potential risk factor in an early epidemiological study several decades ago<sup>31</sup>, but has been inconsistently reported since then.<sup>7,28,31</sup> Bioimpedance is defined as the ability of tissue to impede electric current, and is related to the amount of fluid accumulation in body tissue.<sup>36</sup> Combined with the recent finding that leg bioimpedance is associated with risk for heart failure<sup>37</sup>, these novel predictors suggest a link between high-volume venous reflux, increased hydrostatic pressure, and resulting venous hypertension.

### Genetics of Varicose Veins

While varicose veins are known to have a strong heritable component<sup>9</sup>, their pathophysiological underpinnings remain poorly understood. To date, genetic research on varicose veins has been restricted to syndromic forms of disease (e.g. Klippel-Trenaunay Syndrome) and candidate gene studies in patients with primary venous disease.<sup>38</sup> These studies have suggested that mutations in *FOXC2*, thrombomodulin (*THBD*), and desmuslin (*SYNM*) may promote the development of varicose veins by altering vein function. However, these studies were typically small (18–700 patients), and few confirmation studies have been reported.

In our GWAS, we included nearly 10,000 cases and 300,000 controls, and thus were uniquely powered to screen the whole genome in an unbiased manner. We identified 30 genetic loci robustly associated with varicose veins, with the strongest associations occurring in the intron region of *CASZ1* (rs11121615,  $P=3.71 \times 10^{-65}$ ) which has been implicated in blood pressure<sup>24</sup> and in 16q24, where amongst other genes, *PIEZO1*, encoding a vascular mechanosensory channel, is located (rs2911463,  $P=4.81 \times 10^{-29}$ ). Disruption of *PIEZO1* has previously been shown to result in significant disorganization of the vascular system, suggesting the importance of *PIEZO1* for mature vascular development.<sup>39</sup> In addition, the lead SNP rs7773004 on 6q14 is located within 50 kB upstream to the hemochromatosis gene *HFE*, in which mutations have previously been associated with both venous ulceration and venous thromboembolism (VTE, an aggregate of DVT and pulmonary embolism).<sup>40,41</sup> These loci appear particularly interesting given the strong genetic overlap that we discovered between varicose veins and DVT. Genes regulating vascular dysfunction may promote changes in venous architecture and chronic blood stasis (and could represent new translational targets).

When comparing the genetic correlation of varicose veins with traditional risk factors and peripheral arterial disease, LD score regression demonstrated the strongest genetic correlation with history of DVT. Several genes have previously been linked to both varicose veins and DVT, including thrombomodulin (*THBD*)<sup>42</sup> and methylenetetrahydrofolate reductase (*MTHFR*).<sup>43</sup> Mutations in these genes are classically associated with inherited hypercoagulability, but patients with varicose veins have also been found to have a higher

prevalence of thrombophilias and increased levels of systemic inflammatory and pro-thrombotic markers.<sup>44–46</sup> In addition, a nationwide genetic study in Sweden demonstrated that varicose veins and VTE share familial susceptibility.<sup>47</sup> These findings similarly support the shared genetic associations between varicose veins and DVT, and suggest an overlap in their underlying pathophysiology.

To gain further biological insights from our GWAS, we performed a series of analyses to identify genes and pathways which may be dysregulated in individuals predisposed to varicose vein formation (Supplemental Tables 7–9). Interestingly, eQTL analyses suggested that several SNPs are associated with altered *in vivo* expression of genes that have been related to vascular development/integrity (e.g. *PPP3R1*<sup>25</sup>, *PIEZO1*<sup>39</sup>, *SOX18*<sup>48</sup>), limb development (e.g. *LBH*<sup>49</sup>), and conditions associated with skeletal abnormalities (e.g. *GALNS*<sup>26</sup>, *FBN2*<sup>27</sup>) (Supplemental Table 7). The potential importance of genes related to vascular development ( $P=1.96\times 10^{-6}$ ), limb development ( $P=3.55\times 10^{-3}$ ), and height ( $P=2.30\times 10^{-3}$ ) were also highlighted in pathway analyses (Supplemental Table 9). In particular, genetic variants associated with varicose veins were found to be enriched for the VEGF (vascular endothelial growth factor) pathway and NFAT (nuclear factor of activated T-cells) pathway. VEGF signaling is required for vessel formation and homeostasis, while the NFAT pathway regulates intracellular calcium signaling in osteoblasts and osteoclasts during bone growth and remodeling.<sup>50</sup> These findings warrant future studies to experimentally evaluate the reported genetic loci.

Because height was associated with varicose veins in our observational analyses, we concluded our studies with LD score regression and MR analyses which demonstrated a shared genetic etiology, and confirmed a strong causal relationship between increased height and varicose veins (Supplemental Figures 4–5, Supplemental Table 10). Interestingly, height has been found to be an independent predictor for venous thromboembolism as well.<sup>51</sup> Considering this, our findings suggest that height may have consequences on long-term risk for both varicose veins and its associated complications.

### Study Strengths and Limitations

Strengths of our study include the large sample size and utilization of combined epidemiological and genetic studies, making it the largest and most comprehensive study of varicose veins to date. Nevertheless, this study does have important limitations. The UK Biobank includes predominantly white, middle-aged to elderly individuals, so our conclusions may have limited generalizability to people of other ethnicities and age groups (such as in younger people who have not yet developed disease). Our epidemiologic analysis identified incident events with hospital ICD codes. While this allowed standardized identification of varicose vein events, the true incidence of varicose veins may also be under-represented in this dataset as diagnosis required a formal ICD code input and does not capture data from individuals who did not seek medical attention or include cases of self-reported varicose veins. Thus, it is possible that the controls in our analysis included undocumented cases of varicose veins. This may introduce bias and underestimate some of the associations driving them towards the null. It is thought that the potential bias could lead to false positive associations. In addition, limitations in characterizing candidate variables

such as physical activity, job, and diet may have altered the strength of associations in our observational analyses, which could be a reason why some of the debated risk factors did not show significant associations. Lastly, our analysis does not include the relation of our findings to disease management. The goal of this study was to identify environmental and genetic risk factors for varicose veins. Thus, although our discoveries may help predict or treat varicose veins in the future, the study did not focus on their impact on current patient care.

## Conclusions

Using a large study to investigate varicose vein disease, our study substantially extends our understanding of the genetic and environmental basis of varicose veins. We demonstrate that greater height is associated with greater susceptibility to varicose vein disease. Several risk factors with importance for predicting disease risk were identified, including previously debated (hormone replacement therapy, length of education) and other novel predictors (leg bioimpedance, history of cellulitis, surgery on leg arteries). Our discovery of genetic determinants of disease also sets a path for identification of new targets in the effort to develop therapy for varicose veins.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

Study design and conception: EF, EI, NJL. Data collection and analysis: AMF, DL, SG, DZ, EI. Data interpretation: all authors. Figure preparation: EF, AMF, DL, SG, DZ. Drafting of manuscript: EF, AMF, SG. Supervision and administrative support: EI, NJL. Critical revision and final approval of manuscript: all authors. This research has been conducted using the UK Biobank Resource under Application Number 13721 and 16097. Summary data from GWAS meta-analyses on height, body mass index, and waist-hip ratio were obtained from the GIANT consortium and downloaded from [https://portals.broadinstitute.org/collaboration/giant/index.php/GIANT\\_consortium\\_data\\_files](https://portals.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files) on December 17, 2017. Results from the GWAS for history of DVT and PAD were contributed by the Neale lab and downloaded from <https://sites.google.com/broadinstitute.org/ukbbgwasresults/> on December 17, 2017 and June 3, 2018, respectively.

This project is dedicated to the memory of Emile Mohler, a tireless champion for the field of Vascular Medicine.

### Sources of Funding

This study was performed with support from National Institutes of Health (1R01HL135313-01) and Knut and Alice Wallenberg Foundation (2013.0126).

## References:

1. Hamdan A Management of varicose veins and venous insufficiency. *JAMA*. 2012;308:2612–2621. [PubMed: 23268520]
2. Beebe-Dimmer JL, Pfeifer JR, Engle JS, Schottenfeld D. The epidemiology of chronic venous insufficiency and varicose veins. *Annals of epidemiology*. 2005;15:175–184. [PubMed: 15723761]
3. Chang SL, Huang YL, Lee MC, Hu S, Hsiao YC, Chang SW, Chang CJ and Chen PC. Association of Varicose Veins With Incident Venous Thromboembolism and Peripheral Artery Disease. *JAMA*. 2018;319:807–817. [PubMed: 29486040]

4. Müller-Bühl U, Leutgeb R, Engeser P, Achankeng EN, Szecsenyi J, Laux G. Varicose veins are a risk factor for deep venous thrombosis in general practice patients. *Vasa*. 2012;41:360–365. [PubMed: 22915533]
5. Mäkivaara LA, Ahti TM, Luukkaala T, Hakama M, Laurikka JO. Persons with varicose veins have a high subsequent incidence of arterial disease: a population-based study in Tampere, Finland. *Angiology*. 2007;58:704–709. [PubMed: 18216380]
6. Brand FN, Dannenberg AL, Abbott RD, Kannel WB. The epidemiology of varicose veins: the Framingham Study. *American journal of preventive medicine*. 1988;4:96–101. [PubMed: 3395496]
7. Sisto T, Reunanen A, Laurikka J, Impivaara O, Heliovaara M, Knekt P and Aromaa A. Prevalence and risk factors of varicose veins in lower extremities: mini-Finland health survey. *Eur J Surg*. 1995;161:405–414. [PubMed: 7548376]
8. Scott TE, LaMorte WW, Gorin DR, Menzoian JO. Risk factors for chronic venous insufficiency: a dual case-control study. *Journal of vascular surgery*. 1995;22:622–628. [PubMed: 7494366]
9. Brinsuk M, Tank J, Luft FC, Busjahn A, Jordan J. Heritability of venous function in humans. *Arteriosclerosis, thrombosis, and vascular biology*. 2004;24:207–211.
10. Ahti TM, Mäkivaara LA, Luukkaala T, Hakama M, Laurikka JO. Effect of family history on the risk of varicose veins is affected by differential misclassification. *Journal of clinical epidemiology*. 2010;63:686–690. [PubMed: 20056383]
11. Cornu-Thenard A, Boivin P, Baud JM, De Vincenzi I, Carpentier PH. Importance of the familial factor in varicose disease. Clinical study of 134 families. *The Journal of dermatologic surgery and oncology*. 1994;20:318–326. [PubMed: 8176043]
12. Krysa J, Jones GT, van Rij AM. Evidence for a genetic role in varicose veins and chronic venous insufficiency. *Phlebology / Venous Forum of the Royal Society of Medicine*. 2012;27:329–335.
13. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, Liu B, Matthews P, Ong G, Pell J, Silman A, Young A, Sprosen T, Peakman T and Collins R. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*. 2015;12:e1001779. doi: 10.1371/journal.pmed.1001779. [PubMed: 25826379]
14. Natekin A, Knoll A. Gradient boosting machines, a tutorial. *Frontiers in neurobotics*. 2013;7:21. [PubMed: 24409142]
15. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*. 2015;4:7. [PubMed: 25722852]
16. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics, Patterson N, Daly MJ, Price AL and Neale BM. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet*. 2015;47:291–295. [PubMed: 25642630]
17. Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, Chu AY, Estrada K, Luan J, Kutalik Z, Amin N, Buchkovich ML, Croteau-Chonka DC, Day FR, Duan Y, Fall T, Fehrmann R, Ferreira T, Jackson AU, Karjalainen J, Lo KS, Locke AE, Magi R, Mihailov E, Porcu E, Randall JC, Scherag A, Vinkhuyzen AA, Westra HJ, Winkler TW, Workalemahu T, Zhao JH, Absher D, Albrecht E, Anderson D, Baron J, Beekman M, Demirkan A, Ehret GB, Feenstra B, Feitosa MF, Fischer K, Fraser RM, Goel A, Gong J, Justice AE, Kanoni S, Kleber ME, Kristiansson K, Lim U, Lotay V, Lui JC, Mangino M, Mateo Leach I, Medina-Gomez C, Nalls MA, Nyholt DR, Palmer CD, Pasko D, Pechlivanis S, Prokopenko I, Ried JS, Ripke S, Shungin D, Stancakova A, Strawbridge RJ, Sung YJ, Tanaka T, Teumer A, Trompet S, van der Laan SW, van Setten J, Van Vliet-Ostaptchouk JV, Wang Z, Yengo L, Zhang W, Afzal U, Arnlov J, Arscott GM, Bandinelli S, Barrett A, Bellis C, Bennett AJ, Berne C, Blüher M, Bolton JL, Botcher Y, Boyd HA, Bruinenberg M, Buckley BM, Buyske S, Caspersen IH, Chines PS, Clarke R, Claudi-Boehm S, Cooper M, Daw EW, De Jong PA, Deelen J, Delgado G, Denny JC, Dhonukshe-Rutten R, Dimitriou M, Doney AS, Dorr M, Eklund N, Eury E, Folkersen L, Garcia ME, Geller F, Giedraitis V, Go AS, Grallert H, Grammer TB, Grassler J, Gronberg H, de Groot LC, Groves CJ, Haessler J, Hall P, Haller T, Hallmans G, Hannemann A, Hartman CA, Hassinen M, Hayward C, Heard-Costa NL, Helmer Q, Hemani G, Henders AK, Hillege HL, Hlatky MA, Hoffmann W, Hoffmann P, Holmen O, Houwing-Duistermaat JJ, Illig T, Isaacs A, James AL, Jeff J, Johansen B, Johansson A, Jolley J, Juliusdottir T, Junttila J, Kho AN, Kinnunen L, Klopp N, Kocher T, Kratzer W, Lichtner

P, Lind L, Lindstrom J, Lobbens S, Lorentzon M, Lu Y, Lyssenko V, Magnusson PK, Mahajan A, Maillard M, McArdle WL, McKenzie CA, McLachlan S, McLaren PJ, Menni C, Merger S, Milani L, Moayyeri A, Monda KL, Morken MA, Muller G, Muller-Nurasyid M, Musk AW, Narisu N, Nauck M, Nolte IM, Nothen MM, Oozageer L, Pilz S, Rayner NW, Renstrom F, Robertson NR, Rose LM, Roussel R, Sanna S, Scharnagl H, Scholtens S, Schumacher FR, Schunkert H, Scott RA, Sehmi J, Seufferlein T, Shi J, Silventoinen K, Smit JH, Smith AV, Smolonska J, Stanton AV, Stirrups K, Stott DJ, Stringham HM, Sundstrom J, Swertz MA, Syvanen AC, Tayo BO, Thorleifsson G, Tyrer JP, van Dijk S, van Schoor NM, van der Velde N, van Heemst D, van Oort FV, Vermeulen SH, Verweij N, Vonk JM, Waite LL, Waldenberger M, Wennauer R, Wilkens LR, Willenborg C, Wilsaard T, Wojczynski MK, Wong A, Wright AF, Zhang Q, Arveiler D, Bakker SJ, Beilby J, Bergman RN, Bergmann S, Biffar R, Blangero J, Boomsma DI, Bornstein SR, Bovet P, Brambilla P, Brown MJ, Campbell H, Caulfield MJ, Chakravarti A, Collins R, Collins FS, Crawford DC, Cupples LA, Danesh J, de Faire U, den Ruijter HM, Erbel R, Erdmann J, Eriksson JG, Farrall M, Ferrannini E, Ferrieres J, Ford I, Forouhi NG, Forrester T, Gansevoort RT, Gejman PV, Gieger C, Golay A, Gottesman O, Gudnason V, Gyllenstein U, Haas DW, Hall AS, Harris TB, Hattersley AT, Heath AC, Hengstenberg C, Hicks AA, Hindorf LA, Hingorani AD, Hofman A, Hovingh GK, Humphries SE, Hunt SC, Hypponen E, Jacobs KB, Jarvelin MR, Jousilahti P, Jula AM, Kaprio J, Kastelein JJ, Kayser M, Kee F, Keinanen-Kiukkaanniemi SM, Kiemenev LA, Kooner JS, Kooperberg C, Koskinen S, Kovacs P, Kraja AT, Kumari M, Kuusisto J, Lakka TA, Langenberg C, Le Marchand L, Lehtimaki T, Lupoli S, Madden PA, Mannisto S, Manunta P, Matura A, Matise TC, McKnight B, Meitinger T, Moll FL, Montgomery GW, Morris AD, Morris AP, Murray JC, Nelis M, Ohlsson C, Oldehinkel AJ, Ong KK, Ouwehand WH, Pasterkamp G, Peters A, Pramstaller PP, Price JF, Qi L, Raitakari OT, Rankinen T, Rao DC, Rice TK, Ritchie M, Rudan I, Salomaa V, Samani NJ, Saramies J, Sarzynski MA, Schwarz PE, Sebert S, Sever P, Shuldiner AR, Sinisalo J, Steinhilber V, Stolk RP, Tardif JC, Tonjes A, Tremblay A, Tremoli E, Virtamo J, Vohl MC, Electronic Medical R, Genomics C, Consortium MI, Consortium P, LifeLines Cohort S, Amouyel P, Asselbergs FW, Assimes TL, Bochud M, Boehm BO, Boerwinkle E, Bottinger EP, Bouchard C, Cauchi S, Chambers JC, Chanock SJ, Cooper RS, de Bakker PI, Dedoussis G, Ferrucci L, Franks PW, Froguel P, Groop LC, Haiman CA, Hamsten A, Hayes MG, Hui J, Hunter DJ, Hveem K, Jukema JW, Kaplan RC, Kivimaki M, Kuh D, Laakso M, Liu Y, Martin NG, Marz W, Melbye M, Moebus S, Munroe PB, Njolstad I, Oostra BA, Palmer CN, Pedersen NL, Perola M, Perusse L, Peters U, Powell JE, Power C, Quertermous T, Rauramaa R, Reinmaa E, Ridker PM, Rivadeneira F, Rotter JI, Saaristo TE, Saleheen D, Schlessinger D, Slagboom PE, Snieder H, Spector TD, Strauch K, Stumvoll M, Tuomilehto J, Uusitupa M, van der Harst P, Volzke H, Walker M, Wareham NJ, Watkins H, Wichmann HE, Wilson JF, Zanen P, Deloukas P, Heid IM, Lindgren CM, Mohlke KL, Speliotes EK, Thorsteinsdottir U, Barroso I, Fox CS, North KE, Strachan DP, Beckmann JS, Berndt SI, Boehnke M, Borecki IB, McCarthy MI, Metspalu A, Stefansson K, Uitterlinden AG, van Duijn CM, Franke L, Willer CJ, Price AL, Lettre G, Loos RJ, Weedon MN, Ingelsson E, O'Connell JR, Abecasis GR, Chasman DI, Goddard ME, Visscher PM, Hirschhorn JN and Frayling TM. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet.* 2014;46:1173–1186. [PubMed: 25282103]

18. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, Powell C, Vedantam S, Buchkovich ML, Yang J, Croteau-Chonka DC, Esko T, Fall T, Ferreira T, Gustafsson S, Kutalik Z, Luan J, Magi R, Randall JC, Winkler TW, Wood AR, Workalemahu T, Faul JD, Smith JA, Zhao JH, Zhao W, Chen J, Fehrmann R, Hedman AK, Karjalainen J, Schmidt EM, Absher D, Amin N, Anderson D, Beekman M, Bolton JL, Bragg-Gresham JL, Buyske S, Demirkan A, Deng G, Ehret GB, Feenstra B, Feitosa MF, Fischer K, Goel A, Gong J, Jackson AU, Kanoni S, Kleber ME, Kristiansson K, Lim U, Lotay V, Mangino M, Leach IM, Medina-Gomez C, Medland SE, Nalls MA, Palmer CD, Pasko D, Pechlivanis S, Peters MJ, Prokopenko I, Shungin D, Stancakova A, Strawbridge RJ, Sung YJ, Tanaka T, Teumer A, Trompet S, van der Laan SW, van Setten J, Van Vliet-Ostaptchouk JV, Wang Z, Yengo L, Zhang W, Isaacs A, Albrecht E, Arnlöv J, Arscott GM, Attwood AP, Bandinelli S, Barrett A, Bas IN, Bellis C, Bennett AJ, Berne C, Blagieva R, Bluher M, Bohringer S, Bonnycastle LL, Bottcher Y, Boyd HA, Bruinenberg M, Caspersen IH, Chen YI, Clarke R, Daw EW, de Craen AJM, Delgado G, Dimitriou M, Doney ASF, Eklund N, Estrada K, Eury E, Folkersen L, Fraser RM, Garcia ME, Geller F, Giedraitis V, Gigante B, Go AS, Golay A, Goodall AH, Gordon SD, Gorski M, Grabe HJ, Grallert H, Grammer TB, Grassler J, Gronberg H,



Groves CJ, Gusto G, Haessler J, Hall P, Haller T, Hallmans G, Hartman CA, Hassinen M, Hayward C, Heard-Costa NL, Helmer Q, Hengstenberg C, Holmen O, Hottenga JJ, James AL, Jeff JM, Johansson A, Jolley J, Juliusdottir T, Kinnunen L, Koenig W, Koskenvuo M, Kratzer W, Laitinen J, Lamina C, Leander K, Lee NR, Lichtner P, Lind L, Lindstrom J, Lo KS, Lobbens S, Lorbeer R, Lu Y, Mach F, Magnusson PKE, Mahajan A, McArdle WL, McLachlan S, Menni C, Merger S, Mihailov E, Milani L, Moayyeri A, Monda KL, Morken MA, Mulas A, Muller G, Muller-Nurasyid M, Musk AW, Nagaraja R, Nothen MM, Nolte IM, Pilz S, Rayner NW, Renstrom F, Rettig R, Ried JS, Ripke S, Robertson NR, Rose LM, Sanna S, Scharnagl H, Scholtens S, Schumacher FR, Scott WR, Seufferlein T, Shi J, Smith AV, Smolonska J, Stanton AV, Steinthorsdottir V, Stirrups K, Stringham HM, Sundstrom J, Swertz MA, Swift AJ, Syvanen AC, Tan ST, Tayo BO, Thorand B, Thorleifsson G, Tyrer JP, Uh HW, Vandenput L, Verhulst FC, Vermeulen SH, Verweij N, Vonk JM, Waite LL, Warren HR, Waterworth D, Weedon MN, Wilkens LR, Willenborg C, Wilsgaard T, Wojczynski MK, Wong A, Wright AF, Zhang Q, LifeLines Cohort S, Brennan EP, Choi M, Dastani Z, Drong AW, Eriksson P, Franco-Cereceda A, Gadin JR, Gharavi AG, Goddard ME, Handsaker RE, Huang J, Karpe F, Kathiresan S, Keildson S, Kiryluk K, Kubo M, Lee JY, Liang L, Lifton RP, Ma B, McCarroll SA, McKnight AJ, Min JL, Moffatt MF, Montgomery GW, Murabito JM, Nicholson G, Nyholt DR, Okada Y, Perry JRB, Dorajoo R, Reinmaa E, Salem RM, Sandholm N, Scott RA, Stolk L, Takahashi A, Tanaka T, van 't Hooft FM, Vinkhuyzen AAE, Westra HJ, Zheng W, Zondervan KT, Consortium AD, Group A-BW, Consortium CAD, Consortium CK, Glc Icbp, Investigators M, Mu TC, Consortium MI, Consortium P, ReProGen C, Consortium G, International Endogene C, Heath AC, Arveiler D, Bakker SJL, Beilby J, Bergman RN, Blangero J, Bovet P, Campbell H, Caulfield MJ, Cesana G, Chakravarti A, Chasman DI, Chines PS, Collins FS, Crawford DC, Cupples LA, Cusi D, Danesh J, de Faire U, den Ruijter HM, Dominiczak AF, Erbel R, Erdmann J, Eriksson JG, Farrall M, Felix SB, Ferrannini E, Ferrieres J, Ford I, Forouhi NG, Forrester T, Franco OH, Gansevoort RT, Gejman PV, Gieger C, Gottesman O, Gudnason V, Gyllensten U, Hall AS, Harris TB, Hattersley AT, Hicks AA, Hindorf LA, Hingorani AD, Hofman A, Homuth G, Hovingh GK, Humphries SE, Hunt SC, Hypponen E, Illig T, Jacobs KB, Jarvelin MR, Jockel KH, Johansen B, Jousilahti P, Jukema JW, Jula AM, Kaprio J, Kastelein JJP, Keinanen-Kiukkaanniemi SM, Kiemene LA, Knekt P, Kooner JS, Kooperberg C, Kovacs P, Kraja AT, Kumari M, Kuusisto J, Lakka TA, Langenberg C, Marchand LL, Lehtimaki T, Lyssenko V, Mannisto S, Marette A, Matise TC, McKenzie CA, McKnight B, Moll FL, Morris AD, Morris AP, Murray JC, Nelis M, Ohlsson C, Oldehinkel AJ, Ong KK, Madden PAF, Pasterkamp G, Peden JF, Peters A, Postma DS, Pramstaller PP, Price JF, Qi L, Raitakari OT, Rankinen T, Rao DC, Rice TK, Ridker PM, Rioux JD, Ritchie MD, Rudan I, Salomaa V, Samani NJ, Saramies J, Sarzynski MA, Schunkert H, Schwarz PEH, Sever P, Shuldiner AR, Sinisalo J, Stolk RP, Strauch K, Tonjes A, Tregouet DA, Tremblay A, Tremoli E, Virtamo J, Vohl MC, Volker U, Waeber G, Willemsen G, Witteman JC, Zillikens MC, Adair LS, Amouyel P, Asselbergs FW, Assimes TL, Bochud M, Boehm BO, Boerwinkle E, Bornstein SR, Bottinger EP, Bouchard C, Cauchi S, Chambers JC, Chanock SJ, Cooper RS, de Bakker PIW, Dedoussis G, Ferrucci L, Franks PW, Froguel P, Groop LC, Haiman CA, Hamsten A, Hui J, Hunter DJ, Hveem K, Kaplan RC, Kivimaki M, Kuh D, Laakso M, Liu Y, Martin NG, Marz W, Melbye M, Metspalu A, Moebus S, Munroe PB, Njolstad I, Oostra BA, Palmer CNA, Pedersen NL, Perola M, Perusse L, Peters U, Power C, Quertermous T, Rauramaa R, Rivadeneira F, Saaristo TE, Saleheen D, Sattar N, Schadt EE, Schlessinger D, Slagboom PE, Snieder H, Spector TD, Thorsteinsdottir U, Stumvoll M, Tuomilehto J, Uitterlinden AG, Uusitupa M, van der Harst P, Walker M, Wallaschofski H, Wareham NJ, Watkins H, Weir DR, Wichmann HE, Wilson JF, Zanen P, Borecki IB, Deloukas P, Fox CS, Heid IM, O'Connell JR, Strachan DP, Stefansson K, van Duijn CM, Abecasis GR, Franke L, Frayling TM, McCarthy MI, Visscher PM, Scherag A, Willer CJ, Boehnke M, Mohlke KL, Lindgren CM, Beckmann JS, Barroso I, North KE, Ingelsson E, Hirschhorn JN, Loos RJF and Speliotes EK. Genetic studies of body mass index yield new insights for obesity biology. *Nature*. 2015;518:197–206. [PubMed: 25673413]

19. Heid IM, Jackson AU, Randall JC, Winkler TW, Qi L, Steinthorsdottir V, Thorleifsson G, Zillikens MC, Speliotes EK, Magi R, Workalemahu T, White CC, Bouatia-Naji N, Harris TB, Berndt SI, Ingelsson E, Willer CJ, Weedon MN, Luan J, Vedantam S, Esko T, Kilpelainen TO, Kutalik Z, Li S, Monda KL, Dixon AL, Holmes CC, Kaplan LM, Liang L, Min JL, Moffatt MF, Molony C, Nicholson G, Schadt EE, Zondervan KT, Feitosa MF, Ferreira T, Lango Allen H, Weyant RJ, Wheeler E, Wood AR, Estrada K Magic, Goddard ME, Lettre G, Mangino M, Nyholt DR, Purcell



S, Smith AV, Visscher PM, Yang J, McCarroll SA, Nemesh J, Voight BF, Absher D, Amin N, Aspelund T, Coin L, Glazer NL, Hayward C, Heard-Costa NL, Hottenga JJ, Johansson A, Johnson T, Kaakinen M, Kapur K, Ketkar S, Knowles JW, Kraft P, Kraja AT, Lamina C, Leitzmann MF, McKnight B, Morris AP, Ong KK, Perry, JR, Peters MJ, Polasek O, Prokopenko I, Rayner NW, Ripatti S, Rivadeneira F, Robertson NR, Sanna S, Sovio U, Surakka I, Teumer A, van Wingerden S, Vitart V, Zhao JH, Cavalcanti-Proenca C, Chines PS, Fisher E, Kulzer, JR, Lecoeur C, Narisu N, Sandholt C, Scott LJ, Silander K, Stark K, Tammesoo ML, Teslovich TM, Timpson NJ, Watanabe RM, Welch R, Chasman DI, Cooper MN, Jansson JO, Kettunen J, Lawrence RW, Pellikka N, Perola M, Vandenput L, Alavere H, Almgren P, Atwood LD, Bennett AJ, Biffar R, Bonnycastle LL, Bornstein SR, Buchanan TA, Campbell H, Day IN, Dei M, Dorr M, Elliott P, Erdos MR, Eriksson JG, Freimer NB, Fu M, Gaget S, Geus EJ, Gjesing AP, Grallert H, Grassler J, Groves CJ, Guiducci C, Hartikainen AL, Hassanali N, Havulinna AS, Herzig KH, Hicks AA, Hui J, Igl W, Jousilahti P, Jula A, Kajantie E, Kinnunen L, Kolcic I, Koskinen S, Kovacs P, Kroemer HK, Krzely V, Kuusisto J, Kvaloy K, Laitinen J, Lantieri O, Lathrop GM, Lokki ML, Luben RN, Ludwig B, McArdle WL, McCarthy A, Morken MA, Nelis M, Neville MJ, Pare G, Parker AN, Peden JF, Pichler I, Pietilainen KH, Platou CG, Pouta A, Ridderstrale M, Samani NJ, Saramies J, Sinisalo J, Smit JH, Strawbridge RJ, Stringham HM, Swift AJ, Teder-Laving M, Thomson B, Usala G, van Meurs JB, van Ommen GJ, Vatin V, Volpato CB, Wallaschofski H, Walters GB, Widen E, Wild SH, Willemsen G, Witte DR, Zgaga L, Zitting P, Beilby JP, James AL, Kahonen M, Lehtimäki T, Nieminen MS, Ohlsson C, Palmer LJ, Raitakari O, Ridker PM, Stumvoll M, Tonjes A, Viikari J, Balkau B, Ben-Shlomo Y, Bergman RN, Boeing H, Smith GD, Ebrahim S, Froguel P, Hansen T, Hengstenberg C, Hveem K, Isomaa B, Jorgensen T, Karpe F, Khaw KT, Laakso M, Lawlor DA, Marre M, Meitinger T, Metspalu A, Midthjell K, Pedersen O, Salomaa V, Schwarz PE, Tuomi T, Tuomilehto J, Valle TT, Wareham NJ, Arnold AM, Beckmann JS, Bergmann S, Boerwinkle E, Boomsma DI, Caulfield MJ, Collins FS, Eiriksdottir G, Gudnason V, Gyllenstein U, Hamsten A, Hattersley AT, Hofman A, Hu FB, Illig T, Iribarren C, Jarvelin MR, Kao WH, Kaprio J, Launer LJ, Munroe PB, Oostra B, Penninx BW, Pramstaller PP, Psaty BM, Quertermous T, Rissanen A, Rudan I, Shuldiner AR, Soranzo N, Spector TD, Syvanen AC, Uda M, Uitterlinden A, Volzke H, Vollenweider P, Wilson JF, Witteman JC, Wright AF, Abecasis GR, Boehnke M, Borecki IB, Deloukas P, Frayling TM, Groop LC, Haritunians T, Hunter DJ, Kaplan RC, North KE, O'Connell JR, Peltonen L, Schlessinger D, Strachan DP, Hirschhorn JN, Assimes TL, Wichmann HE, Thorsteinsdottir U, van Duijn CM, Stefansson K, Cupples LA, Loos RJ, Barroso I, McCarthy MI, Fox CS, Mohlke KL and Lindgren CM. Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nat Genet.* 2010;42:949–960. [PubMed: 20935629]

20. Neale lab. Rapid GWAS of Thousands of Phenotypes for 337,000 Samples in the UK Biobank, 2017 Available from: <https://sites.google.com/broadinstitute.org/ukbbgwasresults/>.
21. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nature genetics.* 2013;45:580–585. [PubMed: 23715323]
22. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nature communications.* 2017;8:1826.
23. Pers TH, Karjalainen JM, Chan Y, Westra HJ, Wood AR, Yang J, Lui JC, Vedantam S, Gustafsson S, Esko T, Frayling T, Speliotes EK. Genetic Investigation of ATC, Boehnke M, Raychaudhuri S, Fehrmann RS, Hirschhorn JN and Franke L. Biological interpretation of genome-wide association studies using predicted gene functions. *Nat Commun.* 2015;6:5890. [PubMed: 25597830]
24. Levy D, Ehret GB, Rice K, Verwoert GC, Launer LJ, Dehghan A, Glazer NL, Morrison AC, Johnson AD, Aspelund T, Aulchenko Y, Lumley T, Kottgen A, Vasani RS, Rivadeneira F, Eiriksdottir G, Guo X, Arking DE, Mitchell GF, Mattace-Raso FU, Smith AV, Taylor K, Scharpf RB, Hwang SJ, Sijbrands EJ, Bis J, Harris TB, Ganesh SK, O'Donnell CJ, Hofman A, Rotter JI, Coresh J, Benjamin EJ, Uitterlinden AG, Heiss G, Fox CS, Witteman JC, Boerwinkle E, Wang TJ, Gudnason V, Larson MG, Chakravarti A, Psaty BM and van Duijn CM. Genome-wide association study of blood pressure and hypertension. *Nat Genet.* 2009;41:677–687. [PubMed: 19430479]
25. Graef IA, Chen F, Chen L, Kuo A, Crabtree GR. Signals transduced by Ca(2+)/calineurin and NFATc3/c4 pattern the developing vasculature. *Cell.* 2001;105:863–875. [PubMed: 11439183]
26. Fukuda S, Tomatsu S, Masue M, Sukegawa K, Iwata H, Ogawa T, Nakashima Y, Hori T, Yamagishi A, Hanyu Y. Mucopolysaccharidosis type IVA. N-acetylgalactosamine-6-sulfate sulfatase exonic

- point mutations in classical Morquio and mild cases. *J Clin Invest*. 1992;90:1049–1053. [PubMed: 1522213]
27. Putnam EA, Zhang H, Ramirez F, Milewicz DM. Fibrillin-2 (FBN2) mutations result in the Marfan-like disorder, congenital contractural arachnodactyly. *Nature genetics*. 1995;11:456–458. [PubMed: 7493032]
  28. Lee AJ, Evans CJ, Allan PL, Ruckley CV, Fowkes FG. Lifestyle factors and the risk of varicose veins: Edinburgh Vein Study. *Journal of Clinical Epidemiology*. 2003;56:171–179. [PubMed: 12654412]
  29. Ciardullo AV, Panico S, Bellati C, Rubba P, Rinaldi S, Iannuzzi A, Cioffi V, Iannuzzo G and Berrino F. High endogenous estradiol is associated with increased venous distensibility and clinical evidence of varicose veins in menopausal women. *J Vasc Surg*. 2000;32:544–549. [PubMed: 10957662]
  30. Scott TE, Mendez MV, LaMorte WW, Cupples LA, Vokonas PS, Garcia RI and Menzoian JO. Are varicose veins a marker for susceptibility to coronary heart disease in men? Results from the Normative Aging Study. *Ann Vasc Surg* 2004;18:459–464. [PubMed: 15156364]
  31. Abramson JH, Hopp C, Epstein LM. The epidemiology of varicose veins. A survey in western Jerusalem. *Journal of epidemiology and community health*. 1981;35:213–217. [PubMed: 7328382]
  32. Komsuoglu B, Goldeli O, Kulan K, Cetinarslan B, Komsuoglu SS. Prevalence and risk factors of varicose veins in an elderly population. *Gerontology*. 1994;40:25–31. [PubMed: 8034200]
  33. Malhotra SL. An epidemiological study of varicose veins in Indian railroad workers from the South and North of India, with special reference to the causation and prevention of varicose veins. *Int J Epidemiol* 1972;1:177–183. [PubMed: 4669192]
  34. Hirai M, Naiki K, Nakayama R. Prevalence and risk factors of varicose veins in Japanese women. *Angiology* 1990;41:228–232. [PubMed: 2310052]
  35. Carpentier PH, Maricq HR, Biro C, Poncot-Makinen CO, Franco A: Prevalence, risk factors, and clinical patterns of chronic venous disorders of lower limbs: A population-based study in France. *J Vasc Surg* 2004;40:650–659. [PubMed: 15472591]
  36. Khalil SF, Mohktar MS, Ibrahim F. The theory and fundamentals of bioimpedance analysis in clinical status monitoring and diagnosis of diseases. *Sensors (Basel)*. 2014;14:10895–10928. [PubMed: 24949644]
  37. Lindholm D, Fukaya E, Leeper NJ, Ingelsson E. Bioimpedance and New-Onset Heart Failure: A Longitudinal Study of >500 000 Individuals From the General Population. *J Am Heart Assoc*. 2018;7:e008970. [PubMed: 29959136]
  38. Anwar MA, Georgiadis KA, Shalhoub J, Lim CS, Gohel MS, Davies AH. A review of familial, genetic, and congenital aspects of primary varicose vein disease. *Circulation Cardiovascular genetics*. 2012;5:460–466. [PubMed: 22896013]
  39. Li J, Hou B, Tumova S, Muraki K, Bruns A, Ludlow MJ, Sedo A, Hyman AJ, McKeown L, Young RS, Yuldasheva NY, Majeed Y, Wilson LA, Rode B, Bailey MA, Kim HR, Fu Z, Carter DA, Bilton J, Imrie H, Ajuh P, Dear TN, Cubbon RM, Kearney MT, Prasad RK, Evans PC, Ainscough JF and Beech DJ. Piezo1 integration of vascular architecture with physiological force. *Nature*. 2014;515:279–282. [PubMed: 25119035]
  40. Zamboni P, Izzo M, Tognazzo S, Carandina S, De Palma M, Catozzi L, Caggiati A, Scapoli G, Gemmati D. The overlapping of local iron overload and HFE mutation in venous leg ulcer pathogenesis. *Free Radic Biol Med*. 2006;15:1869–1873.
  41. Brown K, Luddington R, Taylor SA, Lillicrap DP and Baglin TP. Risk of venous thromboembolism associated with the common hereditary haemochromatosis Hfe gene (C282Y) mutation. *Br J Haematol*. 1999;105:95–97. [PubMed: 10233369]
  42. Le Flem L, Mennen L, Aubry ML, Aiach M, Scarabin PY, Emmerich J, Alhenc-Gelas M. Thrombomodulin promoter mutations, venous thrombosis, and varicose veins. *Arterioscler Thromb Vasc Biol*. 2001; 21:445–451. [PubMed: 11231927]
  43. Sverdlova AM, Bubnova NA, Baranovskaya SS, Vasina VI, Avitisjan AO, Schwartz EI. Prevalence of the methylenetetrahydrofolate reductase (MTHFR) C677T mutation in patients with varicose veins of lower limbs. *Mol Genet Metab*. 1998;63:35–36. [PubMed: 9538515]

44. Poredos P, Spirkoska A, Rucigaj T, Fareed J, Jezovnik MK. Do blood constituents in varicose veins differ from the systemic blood constituents? *Eur J Vasc Endovasc Surg*. 2015;50:250–256. [PubMed: 26100448]
45. Bergan JJ, Schmid-Schönbein GW, Smith PD, Nicolaides AN, Boisseau MR, Eklof B. Chronic venous disease. *N Engl J Med*. 2006;355:488–498. [PubMed: 16885552]
46. Darvall KA, Sam RC, Adam DJ, Silverman SH, Fegan CD, Bradbury AW. Higher prevalence of thrombophilia in patients with varicose veins and venous ulcers than controls. *J Vasc Surg*. 2009; 49:1235–1241. [PubMed: 19307088]
47. Zoller B, Ji J, Sundquist J, Sundquist K. Venous thromboembolism and varicose veins share familial susceptibility: a nationwide family study in Sweden. *Journal of the American Heart Association*, 2014;3:e000850. [PubMed: 25158864]
48. Pennisi D, Gardner J, Chambers D, Hosking B, Peters J, Muscat G, Abbott C, Koopman P. Mutations in Sox18 underlie cardiovascular and hair follicle defects in ragged mice. *Nat Genet*. 2000;24:434–437. [PubMed: 10742113]
49. Briegel KJ, Joyner AL. Identification and characterization of Lbh, a novel conserved nuclear protein expressed during early limb and heart development. *Developmental biology*. 2001;233:291–304. [PubMed: 11336496]
50. Winslow MM, Pan M, Starbuck M, Gallo EM, Deng L, Karsenty G and Crabtree GR. Calcineurin/NFAT signaling in osteoblasts regulates bone mass. *Dev Cell*. 2006;10:771–782. [PubMed: 16740479]
51. Zoller B, Ji J, Sundquist J, Sundquist K. Body Height and Incident Risk of Venous Thromboembolism: A Cosibling Design. *Circulation Cardiovascular genetics*. 2017;10:e001651. [PubMed: 28874396]

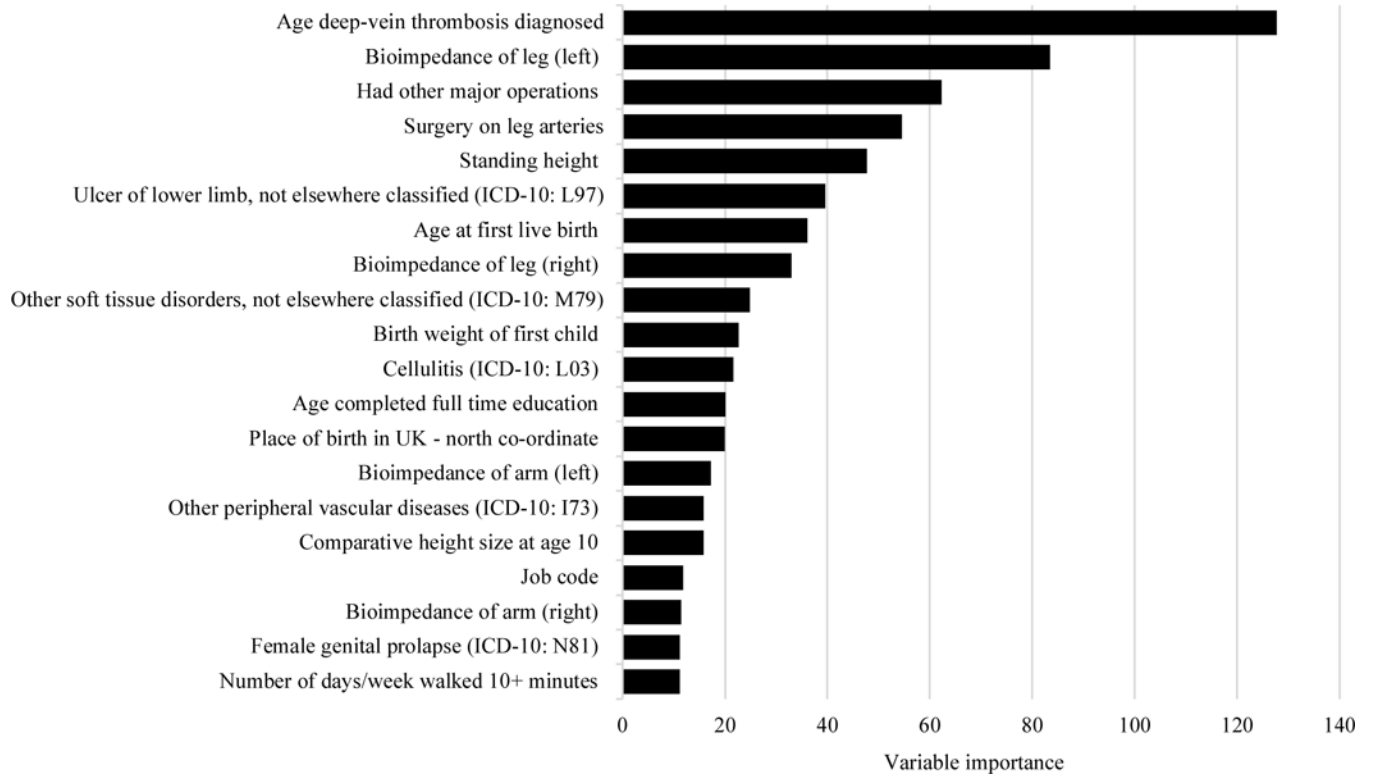
## Clinical Perspective

### What is new?

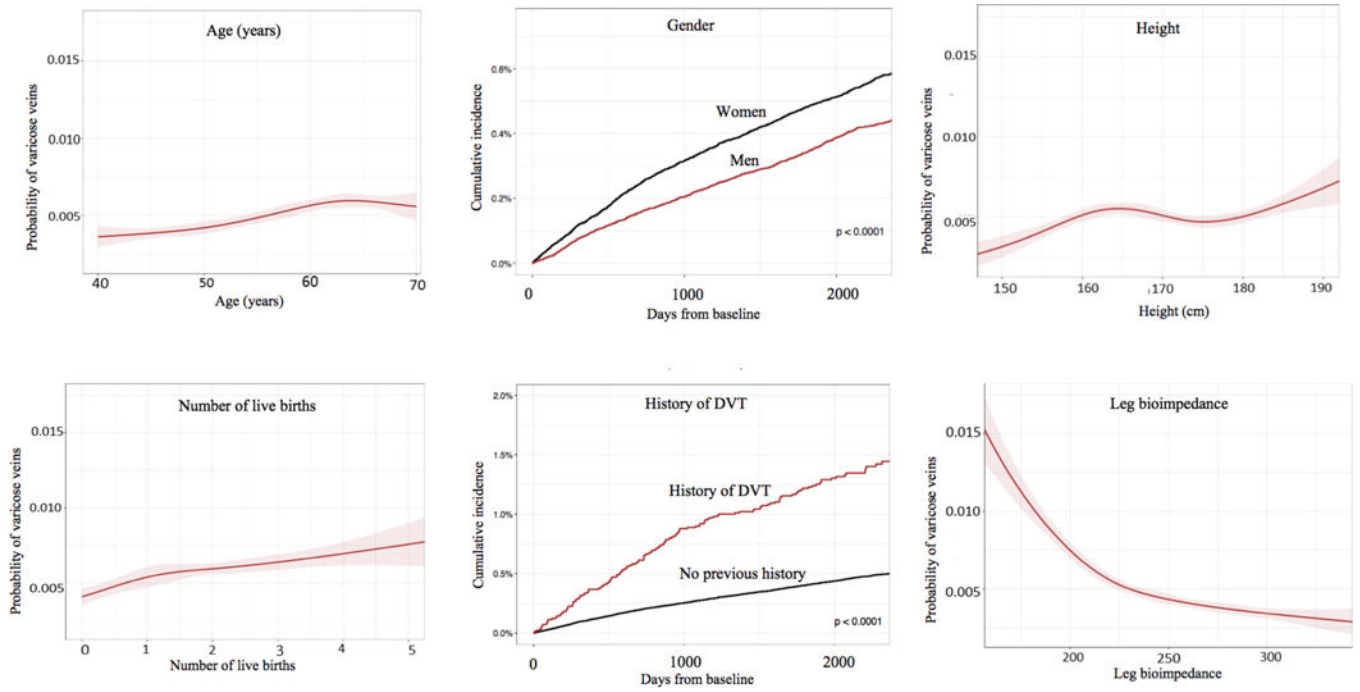
- In this population-based study of ~500,00 individuals, greater height appeared as a novel predictor of varicose vein disease in machine learning analyses, and was independently associated in multivariable-adjusted Cox regression.
- Using Mendelian randomization, we demonstrate that greater height has a causal role in varicose vein development.
- We identify 30 genetic loci in the first large-scale genome-wide analysis of varicose veins, including 337,536 individuals (9,577 cases and 327,959 controls), and discover a strong genetic correlation between varicose veins and deep vein thrombosis.

### What are the clinical implications?

- Varicose veins are a common condition that are increasingly being associated with serious health risks such as venous thromboembolism, yet little is known about the genetic and environmental risk factors for varicose vein disease.
- Our study demonstrates the causal role of height, and identifies novel genetic and clinical associations with varicose veins.
- This knowledge greatly expands our understanding of disease pathophysiology and may help future improvements in the management of varicose veins and their associated complications.

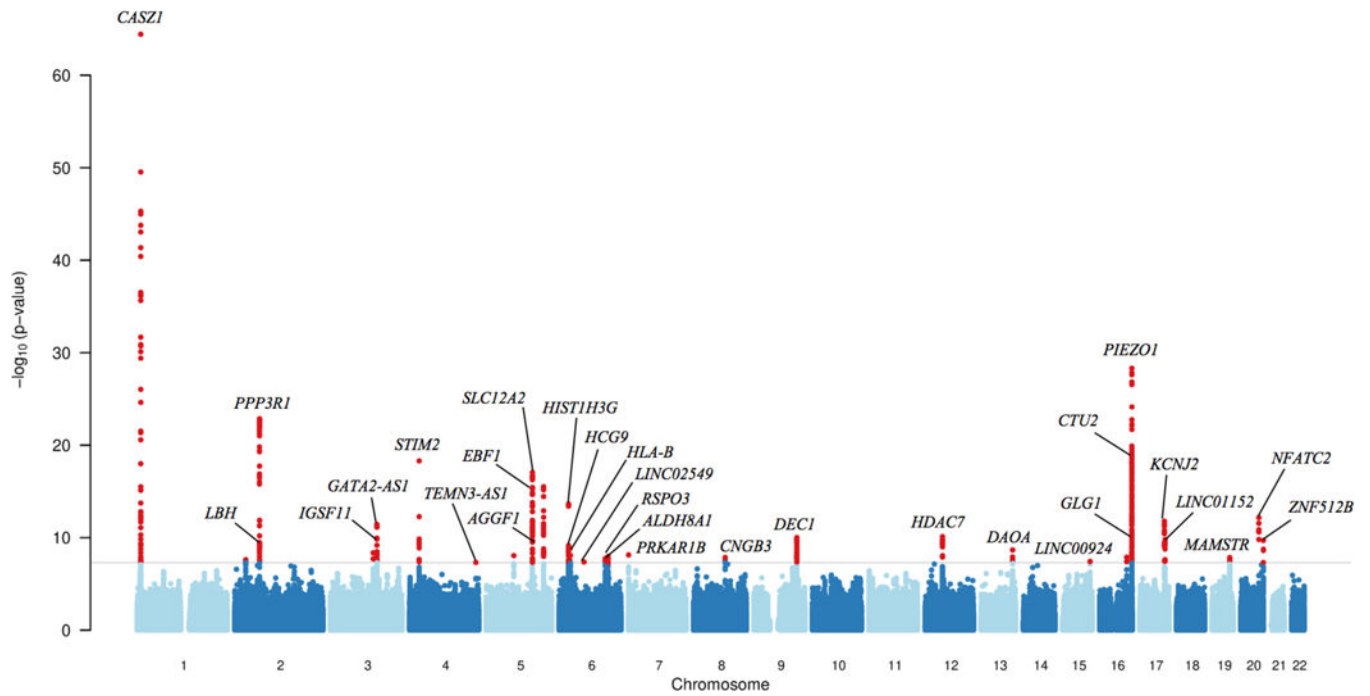


**Figure 1:** Variable importance for the top 20 variables in the gradient boosting model for varicose veins.



**Figure 2:** Univariable analysis demonstrating relationships between varicose veins and other variables.





**Figure 3:** Manhattan plot of GWAS analysis. Gray line represents the threshold of genome-wide significance,  $P < 5 \times 10^{-8}$ . Gene names correspond to the gene in closest proximity to the variant.

**Table 1:**

## Baseline characteristics

Characteristic	
Age (years; N = 493,519)	58 (50–63)
Sex (female; N = 493,519)	54% (226,280)
Height (cm; N = 491,016)	168 (162–175)
BMI (kg/m <sup>2</sup> ; N = 490,513)	26.8 (24.1–29.9)
Waist-hip ratio (N = 491,317)	0.87 (0.80–0.94)
History of DVT (N = 491,435)	2% (9,722)
Leg bioimpedance (left; N = 483,580)	247 (224–271)
Current smoker (N = 490,618)	11% (52,146)
Duration of moderate daily activity (N = 361,856)	40 (20–60)
<b>Female-specific factors</b>	
Number of live births (N = 272,989)	1.82 (1.01–2.63)
Age at live first birth (years; N = 180,038)	25 (22–28)
History of oral contraceptive use (N = 263,309)	3% (6912)
History of hormone replacement therapy (N = 263,309)	7% (19,319)

For continuous variables, medians and interquartile ranges are reported; for categorical variables, percentages and frequencies. N is the number of non-missing values. Moderate daily activity is defined as 10+ continuous minutes of activities such as carrying light loads, cycling at normal pace; does not include walking.

**Table 2:**

Cox models of the top novel predictors from the GBM approach, along with debated predictors for varicose veins.

Variable		Age- and sex-adjusted HR (95%CI)	Fully adjusted* HR (95%CI)
Traditional	History of DVT	2.71 (2.28–3.23)	2.60 (2.18–3.11)
	Birth weight of first child <sup>‡</sup>	1.23 (1.05–1.44)	1.21 (1.03–1.42)
	Number of live births	1.12 (0.99–1.27)	1.12 (0.99–1.27)
	BMI	1.11 (1.00–1.24)	1.11 (0.99–1.24)
	Waist-hip ratio	1.10 (0.97–1.26)	1.01 (0.88–1.16)
Novel	Leg bioimpedance	0.56 (0.51–0.63)	0.43 (0.39–0.50)
	Height	1.70 (1.48–1.96)	1.74 (1.51–2.01)
	Major operations	0.75 (0.69–0.82)	0.77 (0.70–0.84)
	Surgery on leg arteries	3.71 (2.46–5.62)	3.50 (2.30–5.33)
	Ulceration of lower limb (ICD-10: L97)	4.65 (2.36–9.16)	3.00 (1.36–6.61)
	Other soft tissue disorders (ICD-10: M79)	2.01 (1.58–2.55)	1.66 (1.30–2.12)
	Cellulitis (ICD-10: L03)	2.09 (1.52–2.87)	1.79 (1.29–2.48)
Debated	Smoking	1.04 (0.92–1.19)	1.04 (0.91–1.19)
	Systolic blood pressure	0.81 (0.72–0.90)	0.79 (0.71–0.88)
	Age completed full time education	0.80 (0.70–0.91)	0.82 (0.72–0.93)
	Oral contraceptive use	0.67 (0.44–1.02)	0.68 (0.44–1.04)
	Hormone replacement therapy	1.25 (1.04–1.49)	1.25 (1.04–1.50)
	Days per week walked >10 minutes	1.32 (1.18–1.47)	1.33 (1.19–1.48)
	Duration of moderate daily activity	1.00 (0.87–1.14)	1.00 (0.88–1.15)

DVT: deep vein thrombosis. BMI: body mass index. HR: hazard ratio. CI: confidence interval. Associations between candidate variables and varicose veins are reported for as HR and 95% CI for the upper vs. lower quartile for continuous variables, or presence vs. absence for binary variables. Separate models were used for each variable.

\* Fully adjusted model includes age, gender, BMI, history of DVT, and waist-hip ratio. Pregnancy-specific factors, oral contraceptive use, and hormone replacement therapy were only assessed in female study participants.

<sup>‡</sup>Weight in pounds.