01/04/2017

Taller height as a risk factor for venous thromboembolism: a Mendelian randomization meta-analysis

N. S. Roetker,[*] S. M. Armasu,[†] J. S. Pankow,[*] P. L. Lutsey,[*] W. Tang,[*] M. A. Rosenberg,[‡] T. M. Palmer,[§] R. F. MacLehose,[*] S. R. Heckbert,[¶] M. Cushman,[**,††] M. de Andrade,[†] A. R. Folsom[*]

Short Running Head: Height and Venous Thromboembolism

[*]*Division of Epidemiology & Community Health, School of Public Health, University of Minnesota, Minneapolis, MN, USA;* [†]*Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN, USA;* [‡]*Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA, USA;* [§]*Department of Mathematics and Statistics, Lancaster University, Lancaster, UK;* [¶]*Department of Epidemiology, University of Washington, Seattle, WA, USA;* [**]*Department of Medicine, University of Vermont, Burlington, VT, USA;* [††]*Department of Pathology, Larner College of Medicine at the University of Vermont, Burlington, VT, USA*

Correspondence: Nicholas S. Roetker, Division of Epidemiology & Community Health, School of Public Health, University of Minnesota, 1300 South 2nd Street, Suite 300, Minneapolis, MN 55454, USA. Phone: +1 612-624-9261; Fax: +1 612-624-0315; E-mail: roetk001@umn.edu

Word count: 3,924 (from first word of Introduction to last word of Conclusion); 249 (summary)

Essentials

- Observational data suggests taller people have a higher risk of venous thromboembolism (VTE).

- We used Mendelian randomization techniques to further explore this association in three studies.

- Risk of VTE increased by 30–40% for each 10 cm increment in height.

- Height was more strongly associated with deep vein thrombosis than with pulmonary embolism.

Summary. *Background*: Taller height is associated with greater risk of venous thromboembolism (**VTE**). *Objectives*: We used instrumental variable (**IV**) techniques (Mendelian randomization) to further explore this relationship. *Methods*: Participants of European ancestry were included from two cohort studies [Atherosclerosis Risk in Communities (**ARIC**) study and Cardiovascular Health Study (**CHS**)] and one case-control study [Mayo Clinic VTE Study (**Mayo**)]. We created two weighted genetic risk scores (**GRS**) for height; the full GRS included 668 single nucleotide polymorphisms (**SNPs**) from a previously published meta-analysis and the restricted GRS included a subset of 362 SNPs not associated with weight independently of height. Standard logistic regression and IV models were used to estimate odds ratios (**ORs**) for VTE per 10 cm increment in height. ORs were pooled across the three studies using inverse variance weighted random effects meta-analysis. *Results*: Among 9143 ARIC and 3180 CHS participants free of VTE at baseline, there were 367 and 109 incident VTE events. There were 1143 VTE cases and 1292 controls included from Mayo. The pooled ORs from non-IV models and models using the full and restricted GRSs as IVs were 1.27 (95% CI: 1.11, 1.46),

1.34 (95% CI: 1.04, 1.73), and 1.45 (95% CI: 1.04, 201) per 10 cm greater height, respectively.

Conclusions: Taller height is associated with an increased risk of VTE in adults of European ancestry. Possible explanations for this association, including that taller people may have greater venous surface area, greater number of venous valves, or greater hydrostatic pressure, need to be explored further.

3

Introduction

Venous thromboembolism (**VTE**), comprising deep vein thrombosis (**DVT**) and pulmonary embolism (**PE**), occurs in 350 000 to 600 000 people and contributes to over 100 000 deaths in the U.S. each year [1]. Taller adult height is associated with greater risk of VTE [2–11]. For example, the Longitudinal Investigation of Thromboembolism Etiology (**LITE**) found that each 6.3 cm increment in height was independently associated with a 1.14 (95% confidence interval 1.05, 1.24) greater risk of VTE [9]. The physiologic explanation underpinning this increased risk is not yet fully understood; some proposed mechanisms include greater venous surface area, greater number of venous valves, and greater hydrostatic pressure in taller people [9].

It is possible that uncontrolled confounding by other risk factors explains the association of height with VTE, so additional techniques in causal inference would be useful to pursue. Genome-wide association studies have found many genetic variants related to height [12–20]. The most recent meta-analysis by Wood *et al.* found that 697 single nucleotide polymorphisms (**SNPs**) explained 20% of the heritability (and 16% of the overall variance) of height [20]. The high amount of variance in height explained by these SNPs, which should not be influenced by traditional confounding factors or by reverse causality, provides a unique opportunity to apply Mendelian randomization methods to explore the relationship between height and VTE. We used SNPs from the Wood *et al.* meta-analysis [20] as an instrumental variable (**IV**) for height, allowing us to better estimate the association between height and risk of VTE. We implemented these techniques in three separate studies (Atherosclerosis Risk in Communities study, the Cardiovascular Health Study, and the Mayo Clinic VTE Study) and then combined these results using meta-analytic techniques.

Methods

Study Population

The Longitudinal Investigation of Thromboembolism Etiology (LITE) is a combination of two population-based cohort studies: the Atherosclerosis Risk in Communities (**ARIC**) study [21] and the Cardiovascular Health Study (**CHS**) [22]. The study design and methods of LITE were described previously [23,24]. Briefly, ARIC began in 1987-89 with enrollment of 15 792 men and women aged 45 to 64 years, who were predominately black or white, and held follow-up examinations in 1990-92, 1993-95, 1996-98, and 2011-13, in addition to yearly telephone contact. CHS began in 1989-1990 with enrollment of 5201 mostly white men and women aged ≥65 years; an additional 687 African Americans were enrolled in CHS in 1992-93. Over the first 10 years of the study, participants in CHS were contacted every 6 months with either a telephone interview or clinic visit, and then by telephone interview every 6 months thereafter. The Mayo Clinic VTE Study (**Mayo**) is a case-control study of outpatients, including VTE cases referred to Mayo Clinic (n=1488) and controls undergoing general medical examinations at Mayo Clinic (n=1439). Cases and controls were selected for study participation from 1994-2009 and 2004-2009, respectively, using frequency matching on age, sex, state of residence, and prior myocardial infarction and stroke status [25]. Each study was approved by the institutional review board of its participating study center, and all participants provided written informed consent.

Height and Covariate Measurement

Baseline standing height was measured by trained technicians with participants standing as straight as possible without shoes and with back and heels against a vertical mounted centimeter ruler, to the nearest centimeter (cm) in ARIC and nearest half-cm in CHS. Weight was measured using a balance scale. Body mass index (**BMI**) was calculated as weight (in kilograms) divided

by height (in meters) squared. Waist circumference was measured at the level of the umbilicus to the nearest cm. In order to confirm that potential confounders were not associated with the IV for height (a necessary assumption of a valid IV), we additionally consider other measured risk factors for VTE. Blood samples were drawn by staff and blood measurements were performed in study-specific central laboratories. Diabetes was defined as fasting blood glucose ≥7 mmol/L (≥126 mg/dL), non-fasting blood glucose ≥11.1 mmol/L (≥200 mg/dL), a self-reported physician's diagnosis of diabetes, or reported use of a diabetes medication. Glomerular filtration rate was estimated from creatinine using an existing algorithm [26]. Factor VIII and activated partial thromboplastin time (available in ARIC only) were assessed as previously described [27,28]. C-reactive protein was measured via high-sensitivity immunoassay using Visit 2 (1990-92) plasma samples in ARIC [29] and baseline plasma samples in CHS [30]. In Mayo, height and weight were ascertained by self-report at the time of study enrollment for both cases and controls, and BMI was calculated as weight (in kilograms) divided by height (in meters) squared.

VTE Occurrence

Hospitalizations were identified by participants or proxy reports during annual contact by phone in ARIC and semi-annual contact by phone or clinic visit in CHS. Additional hospitalizations were found in ARIC through 2011 and CHS through 2001 by surveillance of local hospital discharge lists and Medicare records, respectively. Copies of hospital records were obtained for hospitalizations with International Classification of Disease (ICD) discharge codes indicating possible VTE events [23]. Validation of VTE events was performed by two physicians using standardized criteria requiring positive imaging tests for diagnosis of DVT and PE [23]. VTE events in Mayo were initially identified by a physician in the hematology clinic and

6

validated by presence of imaging tests or by pathological examination of a surgically removed thrombus, as described previously [25]. DVT events in ARIC and CHS were restricted to those occurring in the lower extremity or vena cava. DVT events in Mayo also included those occurring in the arm.

Genotyping and Imputation

Details on genotyping, quality control, and imputation have been described previously [31–33]. Briefly, genotyping was performed in ARIC, CHS, and Mayo using the Affymetrix 6.0, Illumina 370CNV, and Illumina Human 660W Quad genome-wide SNP arrays, respectively. Before imputation, samples were filtered based on call rates, and SNPs were filtered based on call rates, mean allele frequencies, and Hardy Weinberg equilibrium *P*-values (see Table S1). Imputation to ≈2.5 million autosomal SNPs in the HapMap Phase II sample of European ancestry (CEU) was performed using MACH [34] v1.0.16 in ARIC and Mayo and BIMBAM [35] v0.99 in CHS.

Genetic Instrument for Height

The 697 height-related SNPs and their respective weights used for creating the weighted genetic risk score (**GRS**) IV were obtained from a genome-wide association meta-analysis study of height in 253 288 people of European ancestry [20]. This meta-analysis was performed using only the SNPs in the HapMap II sample; since 8 of the SNPs were unavailable in the 1000 Genomes reference panel, we too elected to use SNPs from HapMap II. The SNP weights are external to Mayo but are not external to ARIC and CHS since the two latter studies were included in the original meta-analysis [together ARIC and CHS comprised 11 338 (4.5%) of the

total sample of 253 288]. In order to minimize possible linkage disequilibrium among SNPs selected for the GRS, we excluded one of two SNPs from any correlated pair of SNPs with $r^2 >$ 0.1 as listed in Supplementary Table 1 from Wood *et al.* [20] (the first of the two correlated SNPs was excluded based on the listed order). We found that 29 SNPs were in possible linkage disequilibrium, leaving a total of 668 SNPs with which to create the "full GRS" (see Table S1S2, Column A for the list of SNPs used in the full GRS). Exploratory analyses indicated that the full GRS was associated with weight, independently of height, in ARIC and CHS (violating a key IV assumption); therefore, as a secondary analysis we created a "restricted GRS" that was uncorrelated with weight independently of height. The set of SNPs used to create the restricted GRS was determined using the ARIC cohort; in linear regression models (separately for each SNP), weight was regressed on each SNP and height, and the 306 SNPs showing the most statistically significant associations with weight ($P < 0.4$) independently of height were excluded. This alpha level was chosen because it was the smallest cut-off for which the restricted GRS was no longer associated with weight independently of height in ARIC. This left a subset of 362 SNPs (see Table S1S2, Column B for the list of SNPs used in the restricted GRS). Each GRS was created by taking $\sum$(beta value * estimated # of allele copies) over all of the SNPs, using the beta values listed in Column G of Supplementary Table 1 from Wood *et al.* [20].

Statistical Analysis

Only persons of European ancestry who provided written informed consent for DNA testing were included in this study. In ARIC and CHS respectively, there were 9349 and 3362 eligible participants with genotyping data available; we excluded those who had missing height information ($n = 5$ and 6), reported a VTE prior to baseline ($n = 160$ and 165), or were taking

anticoagulants at baseline ($n = 41$ and 11), leaving a maximum of 9143 and 3180 participants for the present analyses, respectively. In Mayo, there were 2522 eligible participants with genotyping data available; after excluding those who had missing height information ($n = 117$) there was a maximum of 2435 participants for the present analyses.

To assess the strength of the GRS IVs, we used the F-statistic from a linear model regressing height on the GRS. If the GRS is associated with risk factors for VTE other than height, it would violate assumptions necessary for the IV analysis. To assess this, we estimated proportion of variance ($\mathbf{R^2}$) of each covariate explained by the GRS using linear regression. For Mayo, we only used controls to assess the association between covariates and the GRS because VTE is a relatively rare disease and we considered the controls to be a better representation of the general population compared to the combined group of cases and controls [36].

To examine the possibility of a nonlinear relationship between height and VTE, we initially fit logistic regression models ~~in ARIC~~ with VTE as the outcome and modeling height using restricted cubic splines. Since the association between VTE and height was approximately linear in the logit scale (see Figure S1), we did not transform height in subsequent analyses. We estimated odds ratios (OR) and 95% confidence intervals (CI) for VTE per 10 cm increment in height using three separate logistic models. The first model was non-IV logistic regression adjusted for age, sex, body mass index, and study site in ARIC and CHS and for age, sex, body mass index, Minnesota residence, and history of stroke and myocardial infarction in Mayo. The other two models were IV logistic structural mean models via a generalized method of moments estimator [37], using the full GRS and restricted GRS as the IVs for height in each respective model. Each of these models was fit separately to each of the three datasets. IV models using the full GRS and restricted GRS as the IVs were also run for DVT and PE separately. IV models

were unadjusted for ARIC and CHS and were adjusted for the study matching variables for Mayo. As an alternative to using the restricted GRS IV model, we fit a model using the full GRS as the IV and adjusted for weight as a covariate in order to assess the possibility that weight mediates the height—VTE association.

Although using a GRS as an IV typically avoids the problem of weak instrument bias, there may still be inconsistent effect estimation by the individual variants comprising the GRS [38]. Thus, we performed sensitivity analyses to test for the influence of bias from potential directional pleiotropy or heterogeneity of individual SNP IVs using three robust IV methods: inverse variance weighted (**IVW**), MR-Egger regression, and weighted median estimation [38,39]. Summary IV estimates and standard errors for each SNP (using the ratio estimator [40]) were analyzed using the Stata package *mrrobust* to obtain robust IV OR estimates for the full and restricted set of SNPs for each of the three studies. Evidence for directional pleiotropy or heterogeneity of individual IV estimates was assessed by evaluating the I^2 index and Q test [41] from the IVW analysis and the significance of the intercept term from MR-Egger regression, as well as by visual inspection of funnel plots for asymmetry. The standard errors used in the funnel plots were estimated using the delta method.[42]

For each of the three sets of results for the main VTE analysis, four sets of results for the DVT and PE analysis, and six sets of results for the sensitivity analysis, inverse variance weighted random-effects meta-analysis was performed to obtain a pooled OR and 95% CI across the three studies. Analyses were performed in Stata, version 12.1 (ARIC and CHS) and R, version 3.1.1 (Mayo) [43,44].

Results

Baseline characteristics are summarized for ARIC and CHS in Table 1 and for Mayo in Table 2. Notably, participants in CHS were on average older and shorter compared with those in ARIC and Mayo, and mean BMI was higher in Mayo compared with ARIC and CHS. In each study, height was strongly associated with the IVs (all F-statistics >105 and >63 for the full and restricted GRS, respectively, where F<10 is indicative of a weak IV [45]; Tables 1 and 2). Other anthropometric factors, including BMI, weight, and waist circumference, were also associated with the full or restricted GRS IVs (Tables 1 and 2). However, these anthropometric covariates were not associated with the restricted GRS IV after adjustment for height (all P-values > 0.2; however, adjustment for height could potentially create bias). The GRS IVs also were not generally associated with other common VTE risk factors in the three different studies (Tables 1 and 2). Additionally, a separate GRS for VTE, consisting of 5 SNPs in *F5* (Leiden), *F2*, *F11*, *FGG*, and *ABO* genes [46], was not associated with the full or restricted GRS IVs in ARIC (R^2=0.00; P>0.5).

Among 9143 ARIC and 3180 CHS participants, there were 367 and 109 incident VTE events after 23 and 12 years of follow-up, respectively. There were 1143 VTE cases and 1292 controls included from Mayo. Combining the study-specific non-IV logistic regression model results, the summary OR for VTE per 10 cm increment in height was 1.27 (95% CI: 1.11, 1.46) with moderate heterogeneity between studies (I^2=42%, P=0.18; Figure 1). Sex-stratified point estimates were similar in men and women in each study and by meta-analysis (all P for interaction >0.36, data not shown). The summary ORs obtained by combining study-specific IV model results were similar whether using the full GRS IV [OR: 1.34 (95% CI: 1.04, 1.73)] or the restricted GRS IV [OR: 1.45 (95% CI: 1.04, 2.01)] with little to no evidence of heterogeneity between studies [I^2=22% (P=0.28) and I^2=0% (P=0.98), respectively; Figure 1]. There was

slightly less precision in the estimate of the pooled OR using the restricted GRS IV potentially because inclusion of fewer SNPs resulted in a weaker instrument. The summary OR obtained from the full GRS IV model additionally adjusted for weight was slightly attenuated [OR: 1.24 (95% CI: 0.96, 1.61)]. Combining study-specific IV model results using the restricted GRS IV, height was associated slightly more strongly with DVT [OR: 1.46 (95% CI: 0.98, 2.16)] than with PE [1.20 (0.79, 1.82); see Figure ~~S1~~S2).

~~Sensitivity analyses estimated summary ORs from robust IV analyses that take a weighted summary of the individual effects of multiple SNP IVs [OR (95% CI) for IVW, MR-Egger regression, and weighted median using the set of 362 restricted SNPs: 1.15 (0.96, 1.39), 1.34 (0.99, 1.82), and 1.15 (0.87, 1.54), respectively; $I^2=0$ for each model; see Figure S2] were similar but slightly attenuated in comparison to the analyses using a single GRS as the IV.~~ Results from sensitivity analyses using robust IV estimators that take a weighted summary of the individual effects of the SNPs are shown in Figure S3. When using the set of 362 restricted SNPs as individual IVs, the summary OR using MR-Egger regression [OR: 1.34 (0.99, 1.82); $I^2=0$] was similar to the estimate obtained when using the restricted GRS as a single IV. By comparison, the summary ORs using IVW [OR: 1.15 (95% CI: 0.96, 1.39); $I^2=0$] and weighted median [OR: 1.15 (95% CI: 0.87, 1.54); $I^2=0$] methods were fairly attenuated. Within each study, there was little to no evidence of heterogeneity between SNP IVs in the full or restricted set of SNPs in IVW models [I^2 from 0% to 7%; all Q test P-values $\geq$ 0.16; see Figures ~~S3~~S4-~~S8~~S9]. Funnel plots did not display obvious evidence of heterogeneity or directional pleiotropy using either set of SNPs, which was supported by formal testing of MR-Egger regression intercept terms in ARIC and Mayo (P-values > 0.2) but not in CHS (P-values $\leq$ 0.02; see Figures ~~S3~~S4-~~S8~~S9). Pooled across studies, the summary MR-Egger regression intercept term did not show evidence

of directional pleiotropy [intercept: -0.006 (95% CI: -0.020, 0.008) using full set of SNPs and -0.008 (95% CI: -0.024, 0.007) using restricted set of SNPs].

Discussion

In this meta-analysis we found a consistent positive relationship between height and VTE whether estimating the association of height with VTE using standard logistic regression and adjusting for relevant confounders or using IV analysis with a single weighted GRS, which suggested the risk of VTE increased by approximately 30-40% for each 10 cm increment in height. Sensitivity analyses using robust IV methods with multiple IVs (i.e., individual SNPs) generally supported our main findings; however the associations were attenuated.

Previous studies have found that taller height is associated with lower risk of most cardiovascular disease outcomes and total mortality [10], but taller height is associated with higher risk of VTE [2–11], atrial fibrillation [47], and abdominal aortic aneurysm [48]. Previous Mendelian randomization analyses provide additional support for an association of taller height with lower risk of coronary heart disease [49,50] but not with risk of stroke [49]. To our knowledge, this is the first study using Mendelian randomization to identify taller height as a risk factor for VTE.

Vessel wall damage, blood flow changes, and hypercoagulability are the three important factors for the pathogenesis of thrombosis [51]. A recent analysis in LITE found that the association of height with VTE risk was not mediated by markers of hemostasis or inflammation, suggesting that taller height more likely influences VTE risk through greater venous area, venous stasis, or vessel wall damage than through hypercoagulability [52]. Taller height is a risk factor for various types of cancer [53], another key risk factor for VTE. However, the non-IV OR for VTE was unchanged after excluding 112 cancer-provoked VTE cases in ARIC [OR: 1.46 (95%

13

CI: 1.18, 1.79) per 10 cm in height], and active cancer was an exclusion criterion in Mayo. Thus, it is unlikely that cancer (and its effects on hypercoagulability) acts as a mediator of the association between height and VTE. Taller height is associated with higher resting venous pressure during quiet standing in healthy people [54]. Over time, this higher pressure could lead to more damage to venous walls and thereby greater VTE risk. Also, taller people have a greater venous surface area and possibly more venous valves, allowing for a greater area in which thrombi can form. Indeed, the majority of thrombi form in valve pockets of the deep veins in lower limbs where blood flow is stagnant [55]. This is supported by our finding of a stronger association of height with DVT than with PE.

Our ability to estimate the effect of height on VTE risk depends on the validity of our IV. In the context of this study, the three assumptions of a valid IV are that it (1) is associated with height, (2) is not associated with any confounder of the height–VTE association, and (3) is conditionally independent of VTE given height and confounders. Although we found that assumption (1) was satisfied, it is not possible to explicitly confirm assumptions (2) and (3) since they involve unmeasured confounders and the function of many of the SNPs used to construct the GRS IVs is unknown. Nevertheless, we found that the GRS IVs (particularly the restricted GRS) were generally not associated (independently of height) with measured risk factors for VTE that could potentially confound the association between height and VTE, which provides support for the plausibility of assumption (2). For assumption (3), we assume that the GRS IVs are not related to VTE through any pathway other than height [i.e., there is no horizontal pleiotropy (also known as biological or type I pleiotropy)] [38,56]. Given the large number of SNPs used to construct the GRS IVs, it is difficult to know whether this assumption is reasonable. However, to our knowledge, none of the genes containing the 697 height SNPs from

14

the Wood *et al.* meta-analysis [19] has previously been reported in genome-wide association studies of VTE [32,33,57–59], and we found no association between the GRS IVs for height and a previously reported 5-SNP GRS for VTE [46]. Beyond the three IV assumptions, two additional assumptions need to be met in order to identify the effect using the generalized method of moments estimator. First, we assume there is no effect modification (on the multiplicative scale) by the IV on the association between height and VTE [60]. There was no evidence that this assumption was violated in any of the three studies (all interaction *P*-values ≥ 0.52) except in ARIC when using the full GRS IV (interaction *P*-value = 0.04). Second, we must assume the effect of height on VTE is linear on the logit scale and independent of the IV [37], which appeared reasonable from restricted cubic spline modeling.

The main GRS IV and IVW analyses are unbiased only when all IVs are valid. However, the MR-Egger regression and weighted median analyses allow for consistent effect estimation with some relaxation of IV assumptions, albeit with some loss of precision compared to IVW. Specifically, MR-Egger regression allows all IVs to violate assumption (3) provided they meet the weaker assumption that any pleiotropic effects of an IV (effects of the IV on VTE through pathways other than height) are uncorrelated with the strength of the IV–height association [38], while weighted median estimation allows for up to 50% of the weight contributed by IVs to violate assumptions (2) and (3) [39]. Summary robust IV effect estimates were slightly attenuated compared to estimates from the main IV analysis, which could have possibly resulted from using different estimation methods (ratio estimator vs generalized method of moments) and different IV weighting schemes. The fact that summary robust IV effect estimates were slightly larger when using the restricted set of SNPs may suggest reduced pleiotropy in comparison to using the full set of SNPs.

15

Although there was an association between height and risk of VTE in the main meta-analyzed results overall, in CHS the estimated effect was weak or non-existent and had wide confidence intervals. The apparent lack of association in CHS may be partially explained by the older age of this cohort, whereby the effect of height is masked in older age when there is a higher prevalence of other VTE risk factors, or by selection bias. Other observational studies have analogously reported that the effect of family history as a risk factor for VTE gets weaker with increasing age [61,62].

Strengths of this analysis include using data from three different study populations (two population-based prospective studies and one clinic-based case-control study) and having validated VTE cases in each study. However, there are some limitations of our study. We only included people of European ancestry in order to be consistent with the genome-wide meta-analysis study of height [20], which limits the generalizability of our results. Further, although we have shown that the genetic instrument was not associated with many common VTE risk factors (i.e., potential confounders), we were not able to explore the precise physiologic mechanisms that may mediate the relationship between height and VTE risk, including that taller adults may have greater venous surface area, a greater number of venous valves, and greater hydrostatic pressure. Future studies should examine the relationship of height and VTE risk in other racial/ethnic groups and further explore the mechanisms connecting taller height with VTE risk.

In summary, this analysis provides evidence that taller height is associated with increased risk of VTE in adults of European ancestry. Additionally, we supplement the results of many previous studies reporting associations between height and VTE by use of an IV technique that allows for effect estimation in a way that may be less influenced by traditional confounding

16

factors than observational analyses. Future research should continue using IV techniques to better understand the relationship of other exposures with VTE risk.

Addendum

A. R. Folsom, M. de Andrade, and M. Cushman acquired the data. N. S. Roetker, S. M. Armasu, and T. M. Palmer performed or assisted with the statistical analysis. N. S. Roetker drafted the manuscript. S. M. Armasu, J. S. Pankow, P. L. Lutsey, W. Tang, M. A. Rosenberg, T. M. Palmer, R. F. MacLehose, S. R. Heckbert, M. Cushman, M. de Andrade, and A. R. Folsom made critical comments on the manuscript. All authors read and approved the final manuscript.

Disclosures of Conflicts of Interests

The authors state that they have no conflict of interest.

References

1   Office of the Surgeon General (US), National Heart, Lung, and Blood Institute (US). The Surgeon General's Call to Action to Prevent Deep Vein Thrombosis and Pulmonary Embolism. Rockville (MD): Office of the Surgeon General (US); 2008.

2   Glynn RJ, Rosner B. Comparison of risk factors for the competing risks of coronary heart disease, stroke, and venous thromboembolism. *Am J Epidemiol* 2005; **162**: 975–82.

3   Pomp ER, le Cessie S, Rosendaal FR, Doggen CJM. Risk of venous thrombosis: obesity and its joint effect with oral contraceptive use and prothrombotic mutations. *Br J Haematol* 2007; **139**: 289–96.

4   Lutsey PL, Virnig BA, Durham SB, Steffen LM, Hirsch AT, Jacobs DR, Folsom AR.
     Correlates and consequences of venous thromboembolism: The Iowa Women's Health
     Study. *Am J Public Health* 2010; **100**: 1506–13.

5   Rosengren A, Fredén M, Hansson P-O, Wilhelmsen L, Wedel H, Eriksson H. Psychosocial
     factors and venous thromboembolism: a long-term follow-up study of Swedish men. *J
     Thromb Haemost* 2008; **6**: 558–64.

6   Braekkan SK, Borch KH, Mathiesen EB, Njølstad I, Wilsgaard T, Hansen J-B. Body height
     and risk of venous thromboembolism: The Tromsø Study. *Am J Epidemiol* 2010; **171**: 1109–
     15.

7   Severinsen MT, Johnsen SP, Tjønneland A, Overvad K, Dethlefsen C, Kristensen SR. Body
     height and sex-related differences in incidence of venous thromboembolism: a Danish
     follow-up study. *Eur J Intern Med* 2010; **21**: 268–72.

8   Borch KH, Nyegaard C, Hansen J-B, Mathiesen EB, Njølstad I, Wilsgaard T, Brækkan SK.
     Joint effects of obesity and body height on the risk of venous thromboembolism: the Tromsø
     Study. *Arterioscler Thromb Vasc Biol* 2011; **31**: 1439–44.

9   Lutsey PL, Cushman M, Heckbert SR, Tang W, Folsom AR. Longer legs are associated with
     greater risk of incident venous thromboembolism independent of total body height. The
     Longitudinal Study of Thromboembolism Etiology (LITE). *Thromb Haemost* 2011; **106**:
     113–20.

10 Emerging Risk Factors Collaboration. Adult height and the risk of cause-specific death and vascular morbidity in 1 million people: individual participant meta-analysis. *Int J Epidemiol* 2012; **41**: 1419–33.

11 Flinterman LE, van Hylckama Vlieg A, Rosendaal FR, Cannegieter SC. Body height, mobility, and risk of first and recurrent venous thrombosis. *J Thromb Haemost* 2015; **13**: 548–54.

12 Weedon MN, Lettre G, Freathy RM, Lindgren CM, Voight BF, Perry JRB, Elliott KS, Hackett R, Guiducci C, Shields B, Zeggini E, Lango H, Lyssenko V, Timpson NJ, Burtt NP, Rayner NW, Saxena R, Ardlie K, Tobias JH, Ness AR, et al. A common variant of HMGA2 is associated with adult and childhood height in the general population. *Nat Genet* 2007; **39**: 1245–50.

13 Lettre G, Jackson AU, Gieger C, Schumacher FR, Berndt SI, Sanna S, Eyheramendy S, Voight BF, Butler JL, Guiducci C, Illig T, Hackett R, Heid IM, Jacobs KB, Lyssenko V, Uda M, Diabetes Genetics Initiative, FUSION, KORA, Prostate, Lung Colorectal and Ovarian Cancer Screening Trial, et al. Identification of ten loci associated with height highlights new biological pathways in human growth. *Nat Genet* 2008; **40**: 584–91.

14 Sanna S, Jackson AU, Nagaraja R, Willer CJ, Chen W-M, Bonnycastle LL, Shen H, Timpson N, Lettre G, Usala G, Chines PS, Stringham HM, Scott LJ, Dei M, Lai S, Albai G, Crisponi L, Naitza S, Doheny KF, Pugh EW, et al. Common variants in the GDF5-UQCC region are associated with variation in human height. *Nat Genet* 2008; **40**: 198–203.

15 Weedon MN, Lango H, Lindgren CM, Wallace C, Evans DM, Mangino M, Freathy RM, Perry JRB, Stevens S, Hall AS, Samani NJ, Shields B, Prokopenko I, Farrall M, Dominiczak A, Diabetes Genetics Initiative, Wellcome Trust Case Control Consortium, Johnson T, Bergmann S, Beckmann JS, et al. Genome-wide association analysis identifies 20 loci that influence adult height. *Nat Genet* 2008; **40**: 575–83.

16 Johansson A, Marroni F, Hayward C, Franklin CS, Kirichenko AV, Jonasson I, Hicks AA, Vitart V, Isaacs A, Axenovich T, Campbell S, Dunlop MG, Floyd J, Hastie N, Hofman A, Knott S, Kolcic I, Pichler I, Polasek O, Rivadeneira F, et al. Common variants in the JAZF1 gene associated with height identified by linkage and genome-wide association analysis. *Hum Mol Genet* 2009; **18**: 373–80.

17 Lei S-F, Tan L-J, Liu X-G, Wang L, Yan H, Guo Y-F, Liu Y-Z, Xiong D-H, Li J, Yang T-L, Chen X-D, Guo Y, Deng F-Y, Zhang Y-P, Zhu X-Z, Levy S, Papasian CJ, Hamilton JJ, Recker RR, Deng H-W. Genome-wide association study identifies two novel loci containing FLNB and SBF2 genes underlying stature variation. *Hum Mol Genet* 2009; **18**: 1661–9.

18 Soranzo N, Rivadeneira F, Chinappen-Horsley U, Malkina I, Richards JB, Hammond N, Stolk L, Nica A, Inouye M, Hofman A, Stephens J, Wheeler E, Arp P, Gwilliam R, Jhamai PM, Potter S, Chaney A, Ghori MJR, Ravindrarajah R, Ermakov S, et al. Meta-analysis of genome-wide scans for human adult stature identifies novel Loci and associations with measures of skeletal frame size. *PLoS Genet* 2009; **5**: e1000445.

19 Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, Willer CJ, Jackson AU, Vedantam S, Raychaudhuri S, Ferreira T, Wood AR, Weyant RJ, Segrè AV, Speliotes EK, Wheeler E, Soranzo N, Park J-H, Yang J, Gudbjartsson D, et al. Hundreds of

variants clustered in genomic loci and biological pathways affect human height. *Nature* 2010; **467**: 832–8.

20 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, et al. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet* 2014; **46**: 1173–86.

21 The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol* 1989; **129**: 687–702.

22 Fried LP, Borhani NO, Enright P, Furberg CD, Gardin JM, Kronmal RA, Kuller LH, Manolio TA, Mittelmark MB, Newman A. The Cardiovascular Health Study: design and rationale. *Ann Epidemiol* 1991; **1**: 263–76.

23 Cushman M, Tsai AW, White RH, Heckbert SR, Rosamond WD, Enright P, Folsom AR. Deep vein thrombosis and pulmonary embolism in two cohorts: the Longitudinal Investigation of Thromboembolism Etiology. *Am J Med* 2004; **117**: 19–25.

24 Tsai AW, Cushman M, Rosamond WD, Heckbert SR, Polak JF, Folsom AR. Cardiovascular risk factors and venous thromboembolism incidence: the Longitudinal Investigation of Thromboembolism Etiology. *Arch Intern Med* 2002; **162**: 1182–9.

25 Heit JA, Cunningham JM, Petterson TM, Armasu SM, Rider DN, DE Andrade M. Genetic variation within the anticoagulant, procoagulant, fibrinolytic and innate immunity pathways as risk factors for venous thromboembolism. *J Thromb Haemost* 2011; **9**: 1133–42.

26 Inker LA, Schmid CH, Tighiouart H, Eckfeldt JH, Feldman HI, Greene T, Kusek JW, Manzi J, Van Lente F, Zhang YL, Coresh J, Levey AS, CKD-EPI Investigators. Estimating glomerular filtration rate from serum creatinine and cystatin C. *N Engl J Med* 2012; **367**: 20–9.

27 Cushman M, Cornell ES, Howard PR, Bovill EG, Tracy RP. Laboratory methods and quality assurance in the Cardiovascular Health Study. *Clin Chem* 1995; **41**: 264–70.

28 Folsom AR, Wu KK, Rosamond WD, Sharrett AR, Chambless LE. Prospective study of hemostatic factors and incidence of coronary heart disease: the Atherosclerosis Risk in Communities (ARIC) Study. *Circulation* 1997; **96**: 1102–8.

29 Folsom AR, Lutsey PL, Roetker NS, Rosamond WD, Lazo M, Heckbert SR, Basu S, Cushman M, Selvin E. Elevated hepatic enzymes and incidence of venous thromboembolism: a prospective study. *Ann Epidemiol* 2014; **24**: 817–821.e2.

30 Macy EM, Hayes TE, Tracy RP. Variability in the measurement of C-reactive protein in healthy subjects: implications for reference intervals and epidemiological applications. *Clin Chem* 1997; **43**: 52–8.

31 Psaty BM, O'Donnell CJ, Gudnason V, Lunetta KL, Folsom AR, Rotter JI, Uitterlinden AG, Harris TB, Witteman JCM, Boerwinkle E, CHARGE Consortium. Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: Design of prospective meta-analyses of genome-wide association studies from 5 cohorts. *Circ Cardiovasc Genet* 2009; **2**: 73–80.

32 Tang W, Teichert M, Chasman DI, Heit JA, Morange P-E, Li G, Pankratz N, Leebeek FW, Paré G, de Andrade M, Tzourio C, Psaty BM, Basu S, Ruiter R, Rose L, Armasu SM, Lumley T, Heckbert SR, Uitterlinden AG, Lathrop M, et al. A genome-wide association study for venous thromboembolism: the extended Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium. *Genet Epidemiol* 2013; **37**: 512–21.

33 Heit JA, Armasu SM, Asmann YW, Cunningham JM, Matsumoto ME, Petterson TM, De Andrade M. A genome-wide association study of venous thromboembolism identifies risk variants in chromosomes 1q24.2 and 9q. *J Thromb Haemost* 2012; **10**: 1521–31.

34 Li Y, Willer C, Sanna S, Abecasis G. Genotype Imputation. *Annu Rev Genomics Hum Genet* 2009; **10**: 387–406.

35 Servin B, Stephens M. Imputation-based analysis of association studies: candidate regions and quantitative traits. *PLoS Genet* 2007; **3**: e114.

36 Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. *Stat Methods Med Res* 2015; .

37 Clarke PS, Palmer TM, Windmeijer F. Estimating structural mean models with multiple instrumental variables using the generalised method of moments. *Stat Sci* 2015; **30**: 96–117.

38 Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* 2015; **44**: 512–25.

39 Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol* 2016; **40**: 304–14.

40 Palmer TM, Sterne JAC, Harbord RM, Lawlor DA, Sheehan NA, Meng S, Granell R, Smith GD, Didelez V. Instrumental variable estimation of causal risk ratios and causal odds ratios in Mendelian randomization analyses. *Am J Epidemiol* 2011; **173**: 1392–403.

41 Greco M FD, Minelli C, Sheehan NA, Thompson JR. Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. *Stat Med* 2015; **34**: 2926–40.

42 Thomas DC, Lawlor DA, Thompson JR. Re: Estimation of bias in nongenetic observational studies using "Mendelian triangulation" by Bautista et al. *Ann Epidemiol* 2007; **17**: 511–3.

43 StataCorp. Stata Statistical Software: Release 12. College Station, TX: StataCorp LP; 2011.

44 R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2015.

45 Staiger D, Stock JH. Instrumental Variables Regression with Weak Instruments. *Econometrica* 1997; **65**: 557–86.

46 de Haan HG, Bezemer ID, Doggen CJM, Le Cessie S, Reitsma PH, Arellano AR, Tong CH, Devlin JJ, Bare LA, Rosendaal FR, Vossen CY. Multiple SNP testing improves risk prediction of first venous thrombosis. *Blood* 2012; **120**: 656–63.

47   Rosenberg MA, Patton KK, Sotoodehnia N, Karas MG, Kizer JR, Zimetbaum PJ, Chang JD, Siscovick D, Gottdiener JS, Kronmal RA, Heckbert SR, Mukamal KJ. The impact of height on the risk of atrial fibrillation: the Cardiovascular Health Study. *Eur Heart J* 2012; **33**: 2709–17.

48   Tang W, Yao L, Roetker NS, Alonso A, Lutsey PL, Steenson CC, Lederle FA, Hunter DW, Bengtson LGS, Guan W, Missov E, Folsom AR. Lifetime risk and risk factors for abdominal aortic aneurysm in a 24-year prospective study: the ARIC Study (Atherosclerosis Risk in Communities). *Arterioscler Thromb Vasc Biol* 2016; **36**: 2468–77.

49   Nüesch E, Dale C, Palmer TM, White J, Keating BJ, van Iperen EP, Goel A, Padmanabhan S, Asselbergs FW, EPIC-Netherland Investigators, Verschuren WM, Wijmenga C, Van der Schouw YT, Onland-Moret NC, Lange LA, Hovingh GK, Sivapalaratnam S, Morris RW, Whincup PH, Wannamethe GS, et al. Adult height, coronary heart disease and stroke: a multi-locus Mendelian randomization meta-analysis. *Int J Epidemiol* 2015; .

50   Nelson CP, Hamby SE, Saleheen D, Hopewell JC, Zeng L, Assimes TL, Kanoni S, Willenborg C, Burgess S, Amouyel P, Anand S, Blankenberg S, Boehm BO, Clarke RJ, Collins R, Dedoussis G, Farrall M, Franks PW, Groop L, Hall AS, et al. Genetically determined height and coronary artery disease. *N Engl J Med* 2015; **372**: 1608–18.

51   Mammen EF. Pathogenesis of Venous Thrombosis. *Chest* 1992; **102**: 640S–644S.

52   Cushman M, O'Meara ES, Heckbert SR, Zakai NA, Rosamond W, Folsom AR. Body size measures, hemostatic and inflammatory markers and risk of venous thrombosis: The Longitudinal Investigation of Thromboembolism Etiology. *Thromb Res* 2016; **144**: 127–32.

53 Wirén S, Häggström C, Ulmer H, Manjer J, Bjørge T, Nagel G, Johansen D, Hallmans G, Engeland A, Concin H, Jonsson H, Selmer R, Tretli S, Stocks T, Stattin P. Pooled cohort study on height and risk of cancer and cancer death. *Cancer Causes Control* 2014; **25**: 151–9.

54 Kügler C, Strunk M, Rudofsky G. Venous pressure dynamics of the healthy human leg. Role of muscle activity, joint mobility and anthropometric factors. *J Vasc Res* 2001; **38**: 20–9.

55 Karino T, Motomiya M. Flow through a venous valve and its implication for thrombus formation. *Thromb Res* 1984; **36**: 245–57.

56 Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet* 2014; **23**: R89-98.

57 Bezemer ID, Bare LA, Doggen CJM, Arellano AR, Tong C, Rowland CM, Catanese J, Young BA, Reitsma PH, Devlin JJ, Rosendaal FR. Gene variants associated with deep vein thrombosis. *JAMA* 2008; **299**: 1306–14.

58 Trégouët D-A, Heath S, Saut N, Biron-Andreani C, Schved J-F, Pernod G, Galan P, Drouet L, Zelenika D, Juhan-Vague I, Alessi M-C, Tiret L, Lathrop M, Emmerich J, Morange P-E. Common susceptibility alleles are unlikely to contribute as strongly as the FV and ABO loci to VTE risk: results from a GWAS approach. *Blood* 2009; **113**: 5298–303.

59 Germain M, Saut N, Greliche N, Dina C, Lambert J-C, Perret C, Cohen W, Oudot-Mellakh T, Antoni G, Alessi M-C, Zelenika D, Cambien F, Tiret L, Bertrand M, Dupuy A-M, Letenneur L, Lathrop M, Emmerich J, Amouyel P, Trégouët D-A, et al. Genetics of venous thrombosis: insights from a new genome wide association study. *PloS One* 2011; **6**: e25581.

60 Hernán MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? *Epidemiology* 2006; **17**: 360–72.

61 Bezemer ID, van der Meer FJM, Eikenboom JCJ, Rosendaal FR, Doggen CJM. The value of family history as a risk indicator for venous thrombosis. *Arch Intern Med* 2009; **169**: 610–5.

62 Zöller B, Li X, Sundquist J, Sundquist K. Age- and gender-specific familial risks for venous thromboembolism: a nationwide epidemiological study based on hospitalizations in Sweden. *Circulation* 2011; **124**: 1012–20.