

Bidirectional two-sample Mendelian randomization analysis identifies protein C rather than protein S or antithrombin-III as associated with deep venous thrombosis

Liang Shu¹, Liyan Sun², Cong Yu³, Dabin Ren³, Yisong Zhang³, Ping Zheng³

¹Department of Neurology, Shanghai Ninth People's Hospital, Shanghai, China

²Department of Obstetrics and Gynecology, Shanghai Pudong New area People's Hospital, Shanghai, China

³Department of Neurosurgery, Shanghai Pudong New area People's Hospital, Shanghai, China

Submitted: 14 February 2024; **Accepted:** 1 May 2024

Online publication: 23 July 2024

Arch Med Sci

DOI: <https://doi.org/10.5114/aoms/188205>

Copyright © 2024 Termedia & Banach

Corresponding author:

Dr. Ping Zheng
Department of
Neurosurgery
Shanghai Pudong
New area People's
Hospital
490 South Chuanhuan Road
Shanghai, China
E-mail: jojo_ras@126.com

Abstract

Introduction: Observational studies have indicated significant contributions of protein C and protein S to thrombotic diseases, yet the “anticoagulation paradox” in deep venous thrombosis (DVT) remains unresolved. Therefore, we conducted an investigation to discern the causal effects of protein C, protein S and antithrombin-III on DVT risk.

Material and methods: We employed a two-sample (one to evaluate the gene-exposure relationship and the other to evaluate the gene-outcome relationship) bidirectional Mendelian randomization (MR) framework to assess the causal associations between protein C, protein S, antithrombin-III and DVT.

Results: Genetic associations with DVT were extracted from a comprehensive genome-wide association study involving 484,598 individuals. In the multivariable MR analysis, the odds ratios for DVT per standard deviation (SD) increase were 1.005 (95% CI: 1.002–1.008; $p < 0.001$) for protein C, 0.997 (95% CI: 0.992–1.001; $p = 0.146$) for protein S, and 1.001 (95% CI: 0.998–1.005; $p = 0.456$) for antithrombin-III. A two-step MR mediation analysis revealed that the association between protein C and DVT was partially mediated by body mass index, with a mediated proportion of 11.4% (95% confidence interval, 2.3% to 79.2%).

Conclusions: These findings provide insights into the genetic relationship between relative protein C rather than protein S or antithrombin-III levels and DVT, offering potential utility in identifying at-risk patients for DVT development.

Key words: protein C, deep venous thrombosis, causal effect, Mendelian randomization, mediating effect.

Introduction

Deep venous thrombosis (DVT) stands as a prominent cause of global disability, significantly contributing to the overall burden of disease. Unraveling potential causal factors for DVT and understanding their impact direction could aid in formulating effective prevention strategies during clinical interventions. Beyond established risk factors like pregnancy,

stressors, and certain sociodemographic elements [1, 2], the interplay between anticoagulation and DVT has gained considerable attention [2, 3].

While both body mass index (BMI) and obesity have been associated with DVT, lipid profiles lack a genetic link to DVT [4]. Observational studies indicate a reduced anticoagulation factor in individuals with DVT [5], yet the causal relationship with antithrombin and DVT remains contentious. This hypothesis gains support from the negative correlation between coagulation factors and DVT [6, 7]. Some studies propose that decreased anticoagulation proteins correlate with increased DVT risk, as protein C facilitates anticoagulation, resolving thrombosis and improving outcomes [8, 9]. However, data on the association between plasma protein C, protein S, antithrombin-III (AT-III), and DVT onset have yielded inconclusive and controversial results, given the connection between increased activated protein C (APC) resistance and recurrent spontaneous abortion [10]. Other observational studies and one randomized controlled trial have shown increased mortality risk when relative protein C level is increased via giving fresh frozen plasma [11]. Activated protein C can interact with thrombin, which plays a role in suppressing coagulation by deactivating clotting V and VIII factors [12]. Vitamin status is also associated with the human health situation [13, 14].

Mendelian randomization (MR) offers a valuable approach, and genetic variants may be associated with a lot of things; however, we assume that they affect the outcome on interest solely through utilizing genetic variants associated only with the risk factors of the exposure(s) of interest as instrumental variables [15]. In the context of coagulation profiles, identifying genetic variants solely linked to one component, such as protein C levels, is challenging due to the pleiotropic effects of these genes (which indicates that the genetic variation of protein C might affect other phenotypes apart from the outcome). Therefore, we adopted a mediation MR analysis to explore the potential causal association of anticoagulation factors, particularly relative protein C, with the risk factors of DVT, and further correlated these risk factors with DVT onset. Given the critical role of risk factors in DVT prevention and management [16], a two-step MR analysis was additionally conducted to investigate the mediating pathway from relative protein C, protein S and antithrombin-III against DVT through risk factor-related phenotypes.

Methods and materials

Study design and data sources

Study Design and Data Sources: Based on insights from prior research, we devised a two-sam-

ple (one to evaluate the gene-exposure relationship and the other to evaluate the gene-outcome relationship) Mendelian randomization (MR) approach to assess the causal impact of protein C, protein S and AT-III on DVT. Concurrently, we incorporated clinical risk factors associated with DVT incidence as mediating factors in the MR analysis, including body mass index (BMI), smoking, systolic blood pressure (SBP), diabetes mellitus (DM), and obesity. Single nucleotide polymorphisms (SNPs) serving as IVs were derived from previously published genome-wide association studies (GWAS) [17–19] and filtered for independence (linkage disequilibrium (LD) $r^2 < 0.001$; clumping distance of 1000 kb). An F-statistic exceeding ten was considered robust for predicting the exposure of interest [20, 21]. To mitigate selection bias, datasets focused on individuals of European ancestry.

Data for exposure

SNPs linked to protein C and protein S at genome-wide significance ($p < 5 \times 10^{-8}$) were sourced from a GWAS encompassing 10,708 participants of European ancestry [22]. As this dataset lacked SNPs for antithrombin-III (AT-III), another dataset involving 3,301 participants of European ancestry was employed [23].

Data for outcome

GWAS summary data for DVT from two studies, comprising 484,598 participants of European ancestry and encompassing 9,587,836 SNPs, were utilized. Additionally, we utilized GWAS summary data on self-reported DVT in 337,159 participants of European ancestry from the Neale Lab consortium. We used the public GWAS dataset and ethical approval was exempted by the Institutional Review Board.

Statistical analysis

The two-sample MR method assessed causal effects, presenting results as odds ratios (ORs) with 95% confidence intervals (CIs). The procedure of harmonization of SNP exposure and outcomes was described previously [24]. Five MR approaches (random-effect inverse-variance weighted [IVW], weighted median, MR-Egger, simple mode, and weighted mode) addressed potential heterogeneity and horizontal pleiotropy [25], as genetic variants may be associated with several factors; however, it is assumed that they may affect the outcome of interest only through these genetic variants as instrumental variables. Sensitivity analyses compared causal estimates from various MR methods, including MR-Egger, penalized weighted median, simple mode, IVW, and weighted mode [26]. Forest plots and leave-one-out sensitivity

analyses were employed as well [27]. Heterogeneity statistics were performed using scatter plots and Cochran's Q tests between causal estimates from multiple genetic variants [28].

Validation with MR-IVW and expression quantitative trait loci (eQTLs)

To ensure the robustness of causal estimates derived from Mendelian randomization (MR) studies, adherence to three critical assumptions is imperative: (1) the genetic variants exhibit a robust association with the exposure, (2) the genetic variants remain free from any associations with potential confounders influencing the exposure–outcome link, and (3) the variants do not autonomously impact the outcome aside from their influence on the exposure. To satisfy the requirement of instruments being associated solely with the outcome through exposure, we meticulously excluded single nucleotide polymorphisms (SNPs) strongly correlated with the outcome. Subsequent to this, we harmonized the effects of SNPs on both exposure and outcome, ensuring that β values were aligned with the same alleles. Post-harmonization, palindromic SNPs with intermediate allele frequencies (> 0.42) and potential pleiotropic outliers underwent removal via a heterogeneity test, specifically modified Q statistics.

A comprehensive two-sample bidirectional MR analysis was then conducted to explore the relationships between each of genetic exposure of relative protein C, protein S, AT-III and DVT risk. The reversal MR is used to assess the effect of DVT outcome on the genetic variable of anticoagulation factors, which is a prerequisite for the mediating MR analysis. The multivariable MR-IVW method served to validate the causal associations of protein C with DVT, independent of other anti-coagulation factors. Subsequently, multivariable MR, utilizing the inverse-variance weighted approach, estimated the direct causal effects of protein C, protein S, and AT-III on DVT. The MR Steiger test was additionally employed to gauge the potential reverse causal impact of dietary compositions on DVT. Finally, validation of our findings was undertaken through the integration of expression quantitative trait loci (eQTLs) derived from genome-wide association study (GWAS) summary datasets, with each Ensemble Gene ID considered. A significance threshold of two-sided $p < 0.05$ was established for identifying causal associations for each analysis. For the multiple variable MR analysis, we applied the p threshold as $0.05/3$ (Bonferroni correction) as three factors were included simultaneously. All statistical analyses were executed using R software version 3.6.2, with the TwoSampleMR package version 0.5.2 [29].

Exploration of potential mediators

Incorporating data from various GWAS sources, we sought potential mediators, focusing on obesity-related phenotypes with maximized sample sizes while meticulously avoiding any sample overlap (Supplementary Table S1). The selected potential mediators encompassed diverse information: body mass index (BMI) sourced from 461,460 individuals in the Genetic Investigation of Anthropometric Traits (ukb-b-19953), smoking data from 607,291 individuals in the GSCAN (ieu-b-4877), obesity-related information from 463,010 individuals in MRC-ICU (ukb-b-15541), systolic blood pressure (SBP) data from up to 757,601 individuals in the International Consortium of Blood Pressure Traits, and summary statistics for diabetes mellitus (DM) from up to 655,666 individuals (ebi-a-GCST006867).

Results

Univariable MR

Utilizing genetic variants associated with protein C, univariable MR analysis revealed a significant estimated relationship between relative protein C levels and DVT (Table I). The analysis demonstrated a 0.48% higher per one-standard deviation increment in relative protein C (IVW odds ratio (OR), 1.005; 95% confidence interval (CI): 1.002 to 1.008; $p = 9.20e-04$) (Table I). One standard deviation (SD) of relative PC has been reported to be approximately 0.530% of the protein C in UK Biobank (ukb-a-65; Supplementary Table SII). The scatter and forest plot of the protein C level and DVT relative risk correlation is shown in Figure 1. The detailed statistical information for instruments is listed in Table I. The Cochran's Q statistic showed no pleiotropy between protein C and DVT ($Q = 2.457$, $p = 0.293$). Further, no outlier heterogeneity was detected by MR-PRESSO. MR-Egger also identified no directional pleiotropy ($p = 0.084$). The estimates were similar in size in weighted median (OR = 1.005; 95% CI: 1.04 to 1.007; $p = 1.038e-13$), weighted mode (OR = 1.005; 95% CI: 1.04 to 1.007; $p = 4.775e-3$) and simple mode (OR = 1.004; 95% CI: 1.002 to 1.006; $p = 1.97e-2$), supporting a relative risk effect of relative PC against DVT. The CIs of MR-Egger were wider (OR = 1.008; 95% CI: 1.006 to 1.011; $p = 0.02$) than those of other methods. Additionally, the protein C level was almost not associated with arterial embolism genetically (finb-b-19_ ARTEMBTHR, data not shown).

Further, leave-one-out analysis showed that almost no single SNP is involved in this relationship, suggesting an overall positive relationship between protein C and DVT (Supplementary Figure S1). We also found no reverse causality by the

Table I. MR results for the relationship between anti-coagulation compositions and DVT

Method	Number of SNPs	F	OR (95% CI)	P-value
Vitamin K-dependent protein C:				
MR Egger	4	11	1.008318 1.005742 1.010901	2.394115e-02
Weighted median	4		1.005254 1.003867 1.006644	1.038157e-13
Inverse variance weighted	4		1.004755 1.001940 1.007577	9.198350e-04
Simple mode	4		1.004359 1.002484 1.006237	1.974418e-02
Weighted mode	4		1.005336 1.003953 1.006722	4.774844e-03
Vitamin K-dependent protein S:				
MR Egger	4	21	0.9994879 0.9909100 1.0081400	0.91791552
Weighted median	4		0.9975924 0.9948052 1.0003874	0.09127953
Inverse variance weighted	4		0.9965908 0.9920104 1.0011924	0.14622656
Simple mode	4		0.9918502 0.9861492 0.9975841	0.06884925
Weighted mode	4		0.9998972 0.9971954 1.0026063	0.94530664
Antithrombin-III:				
Wald ratio*	1	10	1.001473 0.9976032 1.005358	0.4561531

*No SNP for *ebi-a-GCST90019436*; we used *prot-a-2695* here.

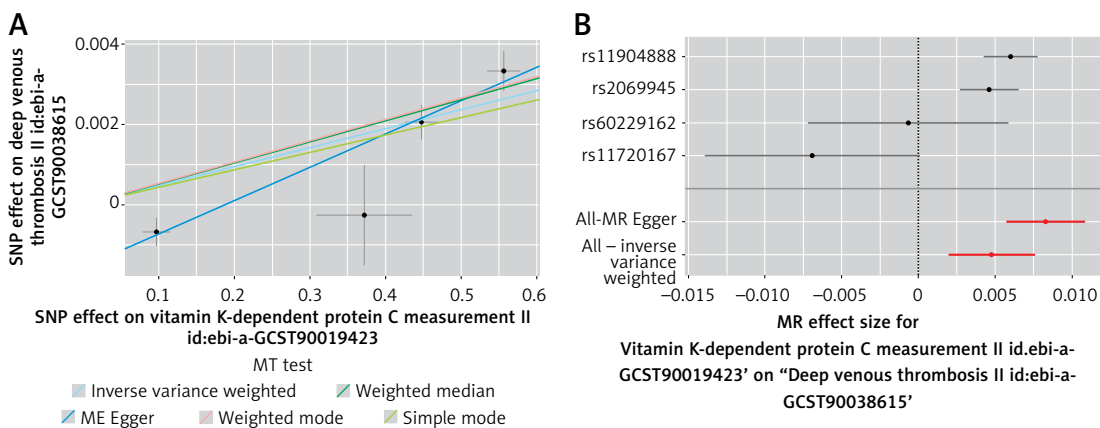


Figure 1. Mendelian randomization plots for relative protein C and DVT relationship: **A** – in this scatter plot, the effects of single nucleotide polymorphisms (SNPs) on relative protein C level are juxtaposed against their effects on deep venous thrombosis (DVT). Each line's slope corresponds to the estimated Mendelian randomization (MR) effect per method, with data presented as raw β values and accompanied by 95% confidence intervals (CIs). **B** – The forest plot shows individual and combined SNP MR-estimated effect sizes. These estimates represent the log odds for DVT per one standard deviation increase in mean relative protein C level, and error bars depict 95% CIs

MR Steiger test (Supplementary Figure S2). The univariable MR estimates were not statistically significant for the effect of relative protein S (IVW OR = 0.997; 95% CI: 0.992 to 1.001; $p = 0.146$) or relative antithrombin-III (IVW OR = 1.001; 95% CI: 0.998 to 1.005; $p = 0.456$) on DVT (Table I).

Multivariable MR

Considering the interrelated nature of risk factors, a multivariable MR analysis was conducted. The effect estimated for relative protein C on DVT remained consistent with univariable estimates. CIs for multivariable MR-Egger were comparable, and the analysis provided no conclusive evidence for horizontal pleiotropy. The effects of relative

protein S and relative AT-III on DVT were not significant (Table II). The effect estimated for relative protein C level on DVT was comparable to the univariable IVW estimate (univariable IVW OR = 1.005; 95% CI: 1.002 to 1.008; $p = 9.20e-04$; multivariable IVW OR = 1.005; 95% CI: 1.001 to 1.009; $p = 0.005$, which passed the Bonferroni correction). The CIs of multivariable MR-Egger were similar to those of multivariable IVW (Table II). The multivariable MR-Egger intercept analysis did not provide conclusive evidence for horizontal pleiotropy ($p = 0.152$). The multivariable MR estimates for relative PS level (IVW OR = 0.996; 95% CI: 0.991 to 1.001; $p = 0.088$) and relative AT-III (IVW OR = 0.999; 95% CI: 0.990 to 1.009; $p = 0.882$) were not significant (Supplementary Figure S3).

Table II. Multivariable MR analysis estimating the effect of relative PC on DVT, conditioning on other anti-coagulation compositions

Item	Std error	Estimate (95% CI)	P-value
IVW:			
Vitamin K-dependent protein C	0.002	0.005 0.001, 0.009	0.005
Vitamin K-dependent protein S	0.002	-0.004 -0.009, 0.001	0.088
Antithrombin-III	0.005	-0.001 -0.010, 0.009	0.882
MR-Egger:			
Vitamin K-dependent protein C	0.002	0.004 0.000, 0.008	0.043
Vitamin K-dependent protein S	0.003	0.000 -0.007, 0.007	0.939
Antithrombin-III	0.006	0.005 -0.007, 0.016	0.405
(Intercept)	0.001	-0.001 -0.002, 0.000	0.152

Effect of DVT on relative protein C level

In the reverse direction, bidirectional MR analyses exploring the causal effect of DVT as the genetic exposure on anti-coagulation factors as the genetic outcome found no evidence of causal relationships with relative protein C (IVW β = 153.536; 95% CI: 1.292e-05 to 1.825e+09; p = 0.545), relative protein S (IVW β = 3.552; 95% CI: 0.445 to 28.319; p = 0.231) or relative AT-III (IVW β = 2.427; 95% CI: 4.022e-03 to 1.464e+03; p = 0.786) (Table III). Further, no outlier heterogeneity was detected by MR-PRESSO. The weighted median, the weighted mode, and MR-Egger showed similar results (data not shown). Leave-one SNP-out analysis also identified no leverage points (data not shown). The relationship between pro-

tein C, protein S, antithrombin-III and DVT was lastly validated with eQTLs. Again, we found that only protein C had a positive causal association with DVT.

Mediation analysis

Given the crucial role of risk factors in venous thrombosis prevention [16], a two-step MR analysis was performed to investigate obesity-related measurements as potential mediators of the protective effect of relative protein C on DVT. The analysis identified a causal relationship between relative protein C and BMI, with increased protein C associated with decreased BMI. Subsequent analysis confirmed a causal effect of BMI on DVT, with a mediated proportion of 11.4% (Figure 2 A). First-

Table III. Bidirectional MR results for the relationship between DVT and anti-coagulation compositions

Method	No. of SNPs	F	β (95% CI)	P-value
DVT on Protein C:				
MR Egger	10	0.8360847	2.329795e-11 3.000425e+10	0.9888343
Weighted median	10	1.9124214	1.465316e-01 2.495951e+01	0.6208178
Inverse variance weighted	10	36	153.5361149 1.292036e-05 1.824511e+09	0.5447437
Simple mode	10	44.9112695	5.728143e-01 3.521250e+03	0.1215040
Weighted mode	10	2.3500054	1.796497e-01 3.074053e+01	0.5311149
DVT on protein S:				
MR Egger	10	3.2503904	0.165846811 63.70359	0.4598031
Weighted median	10	4.1160308	0.312588969 54.19804	0.2820109
Inverse variance weighted	10	14	3.5518969 0.445493981 28.31906	0.2314514
Simple mode	10	0.2704491	0.002736652 26.72708	0.5904602
Weighted mode	10	3.7894712	0.235171134 61.06231	0.3720613
DVT on antithrombin-III:				
MR Egger	10	0.1732527	1.342286e-05 2.236221e+03	0.7260170
Weighted median	10	1.1152479	4.717788e-03 2.636358e+02	0.9687976
Inverse variance weighted	10	24	2.4265302 4.021956e-03 1.463976e+03	0.7861024
Simple mode	10	9.9957232	1.102489e-07 9.062629e+08	0.8110021
Weighted mode	10	0.8984916	1.945416e-03 4.149688e+02	0.9734681

ly, exposure instruments such as protein C were employed to investigate the causal effect on the potential risk factors (mediators). Detailed information about risk factors for the SNPs with DVT is listed in Supplementary Table SIII. Among five potential mediators, we found a causally negative relationship only between protein C level and BMI (IVW $\beta = -0.024$; 95% CI: -0.043 to -0.005 ; $p = 0.014$) (Figure 2 B, Supplementary Table SIV). Secondly, we investigated the causal role of the mediators in DVT risk using genetic exposure for the risk factors. We identified a causal role of BMI (IVW OR = 1.003; 95% CI: 1.001 to 1.005;

$p = 0.0004$) in DVT (Figure 2 B and Supplementary Table SV). The results from other MR methods were consistent. We then employed a multivariable MR (MVMR) to assess the role of BMI on DVT against other obesity-related phenotypes. The MVMR effect estimated for BMI on DVT was similar to protein C (OR = 1.004; 95% CI: 1.002 to 1.006; $p = 0.0002$) (Supplementary Table SVI), which also passed the Bonferroni Correction. The multivariable MR estimates of the effects of other measures on DVT were not significant (data not shown). Finally, we assessed the indirect effect of protein C level on DVT via BMI and found that the

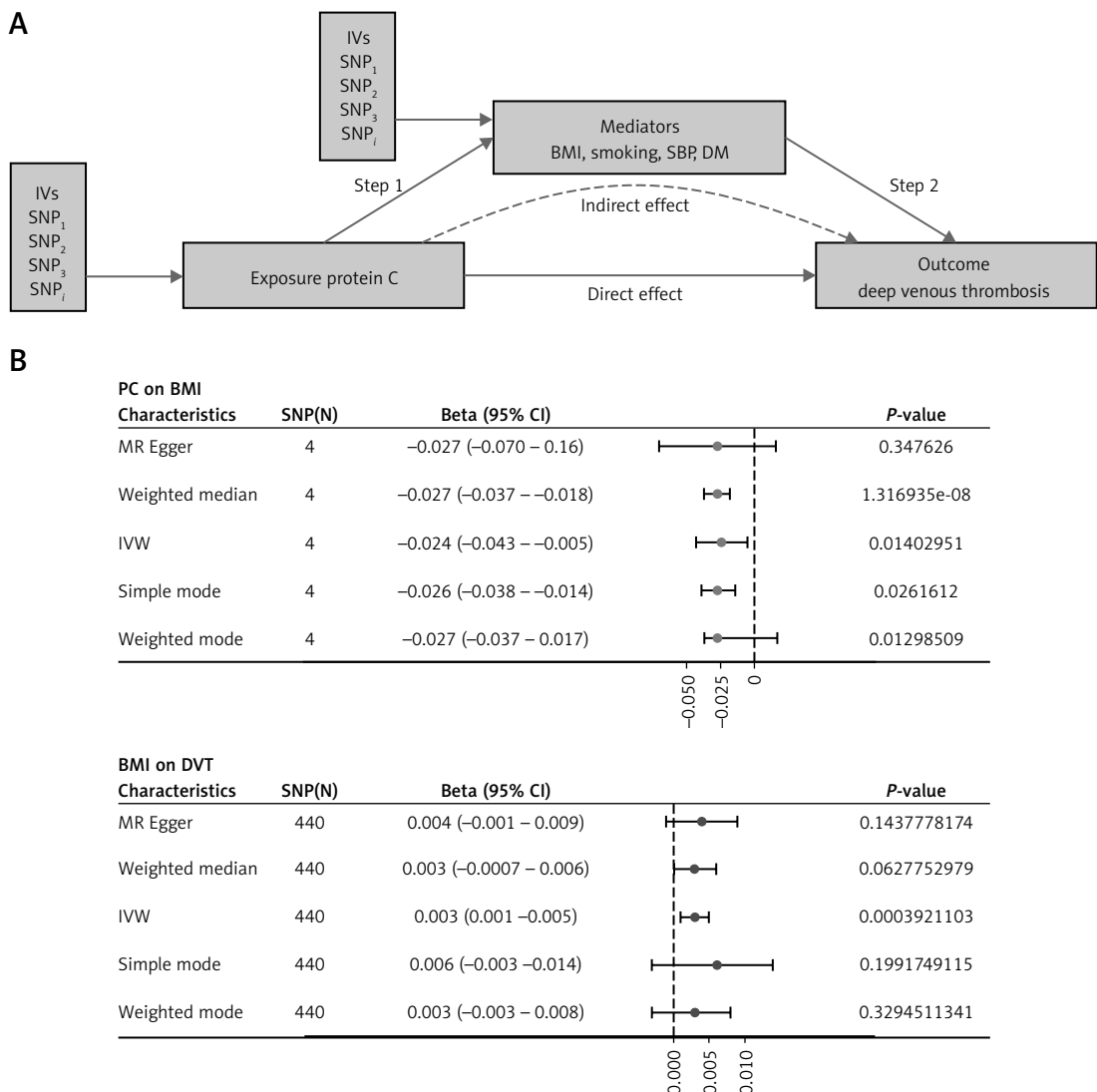


Figure 2. Mediation analysis of relative protein C effect on DVT via potential mediators: **A** – a two-step MR analysis framework is illustrated. Step 1 involves estimating the causal effect of the exposure (relative protein C) on potential mediators, while Step 2 assesses the causal effect of these mediators on DVT risk. The ‘Direct effect’ indicates the impact of relative protein C level on DVT risk after adjusting for the mediator, while the ‘Indirect effect’ signifies the influence of relative protein C level on DVT risk through the mediator. Instrumental variables (IVs) are employed in this analysis. **B** – Summary MR estimates derived from the inverse-variance weighted (IVW), weighted median, weighted mode, and MR-Egger methods are presented for the effect of relative protein C level on BMI (left) and the effect of BMI on DVT (right). The error bars denote 95% CIs, and all statistical tests were two-sided, considering $p < 0.05$ as significant.

Table IV. Mediation effect of protein C on DVT via BMI

Mediator	Total effect	Direct effect A	Direct effect B	Indirect effect C'	P	Mediated proportion (%) (95% CI)
	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)		
BMI	0.005 (0.002–0.008)	–0.024 (–0.043 – –0.005)	0.003 (0.001–0.005)	0.003 (0.001–0.004)	5.552e–05	11.417 (2.326–79.232)

mediation effect of BMI was 0.003 (95% CI: 0.001 to 0.004; $p = 5.552e-05$) with a mediated proportion of 11.4% (95% CI: 2.3% to 79.2%) (Table IV).

Validation analysis

To validate the results, we used the protein QTLs (pQTLs) for protein C, protein S and AT-III to associate them with the DVT outcome and found that only protein C was positively correlated with DVT (Supplementary Figure S4).

Discussions

This pioneering study leverages the Mendelian randomization (MR) approach to discern the causal links between protein C, protein S, antithrombin-III, and the risk of DVT. Our extensive multivariable MR (MVMR) investigation established a positive causal association between genetically predicted protein C levels and DVT risk. In contrast, coagulation traits such as protein S and antithrombin-III exhibit no discernible causal association with DVT. Additionally, a mediation analysis revealed that the influence of relative protein C on DVT is partially mediated by BMI, although the indirect effect is notably smaller than the total effect.

Factor V Leiden, marked by inadequate anticoagulant response to activated protein C, poses an elevated risk for venous thromboembolism [30]. Our findings underscore the complexity of clinical decisions for asymptomatic Factor V Leiden carriers, suggesting that long-term anticoagulation may not be universally recommended. Notably, conditions featuring low protein C or impaired function, such as protein C deficiency and antiphospholipid syndrome, increase the risk of DVT. Protein S, a natural anticoagulant akin to protein C, mirrors the relationship with DVT, with both proteins crucial in regulating blood clotting [30]. When there is a deficiency or dysfunction of protein C, the body's ability to control blood clotting is impaired. This can lead to a higher propensity for the formation of DVT, especially in situations where the risk of clot formation is already elevated (e.g., surgery, immobility, pregnancy). Conditions associated with low levels of protein C or impaired protein C function are considered hypercoagulable states. These conditions can include protein C deficiency, protein S deficiency, and antiphospholipid syndrome, among others. In these situations, the risk of

DVT and other thrombotic events is increased. Protein S is another natural anticoagulant protein in the body that, like protein C, plays a crucial role in regulating blood clotting. The relationship between protein S and DVT is similar to that of protein C.

Our robust sensitivity analyses, employing various MR methods, affirm an unconfounded effect of plasma protein C levels on DVT risk. This contradicts prior studies hinting at a link between protein C deficiency and adverse long-term post-acute DVT outcomes, which proposed that reduced initial protein C was correlated with worse long-term survival after acute DVT [1, 2, 6]. However, these studies were complicated by limitations, such as small sample sizes and statistical errors, emphasizing the importance of our large-scale, genetically informed approach. Notably, our bidirectional analysis dismisses the notion that a genetic predisposition to DVT causally influences protein C levels, challenging previous assumptions.

Various biological mechanisms are proposed to elucidate the potential detrimental role of protein C (PC) in the development of venous thrombosis. In this study, a comprehensive two-step Mendelian randomization (MR) for mediation analysis revealed that the adverse impact of elevated relative protein C on deep venous thrombosis (DVT) risk is partially mediated by body mass index (BMI), albeit with an indirect effect that is less pronounced than the total effect. In the initial MR step, univariable MR identified a causal link between relative protein C levels and BMI, indicating that increased relative protein C is associated with decreased BMI. This finding aligns with prior research demonstrating a linear increase in activated protein C (APC) resistance with rising BMI, particularly observed in men. Coagulation factor VIII (FVIII) levels were implicated in mediating the relationship between BMI and decreased APC ratio [12]. The relationship between protein C and BMI is complex and not fully understood.

The subsequent MR step furnished evidence supporting a genetically determined higher BMI as a contributing factor to increased odds of DVT. Consistent with previous MR studies, which highlighted the causal relationship between BMI and DVT or related thrombosis phenotypes [16, 31], our second-step mediation analysis aligns with these findings. Notably, Christiansen *et al.*

observed elevated DVT risk in individuals with BMI in the upper tertile compared to the lowest tertile, reinforcing the causal role of BMI in DVT pathogenesis [12]. Contrary to expectations, no significant causal associations were observed for smoking, obesity, SBP, and DM with DVT risk, underscoring the prominence of the causal effects of protein C and BMI over other factors.

Protein C is an important protein in the body that functions as an anticoagulant, which means it helps to prevent blood clots from forming. It is a vitamin K-dependent serine protease enzyme, which is activated to its active form, activated protein C (APC), by the thrombin-thrombomodulin complex on endothelial cell surfaces. APC, in turn, degrades Factor Va and Factor VIIIa, which are required for blood clotting.

For its prediction role, protein C deficiency can lead to an increased risk of venous thromboembolism, including DVT and pulmonary embolism (PE). Patients with low levels of protein C may be predisposed to these conditions. Monitoring the levels of protein C can be useful in diagnosing and predicting the progression of disseminated intravascular coagulation. Monitoring protein C levels can also help in managing anticoagulation therapy to prevent both bleeding and thrombosis.

Activation of protein C in endothelium, which co-operates with thrombin pathways, can protect against thrombotic events. However, in inflammation, infection malignancy, obesity, and even pregnancy, excess protein C might overcome its anticoagulant effect in the activated form and augment thrombosis. Therefore, it is important to note that protein C levels should be interpreted in the context of the patient's overall clinical picture and other laboratory results. The predictive value of protein C levels can vary depending on the specific disease and individual patient characteristics.

It is important to note that the relationship between protein C and BMI is still an active area of research, and the specific mechanisms and clinical implications are not fully understood. Additionally, individual responses to obesity can vary widely. People with obesity should be aware of the potential risks associated with clotting disorders and consult with healthcare providers for appropriate monitoring and management, especially if they have other risk factors or a personal or family history of clotting problems.

Admittedly, our study has limitations that warrant careful interpretation, which need to be addressed. Potential horizontal pleiotropy, introducing bias into causal effect estimates, was addressed through rigorous methods, including outlier removal using MR-PRESSO and MR-Egger approaches. Notably, neither heterogeneity nor pleiotropy influenced the causal effect in our analyses. Canalization effects, modifying the associa-

tion of genetically predicted protein C, protein S, and antithrombin-III profiles with DVT during development, were considered. Our MR analyses focused on evaluating not only heritable variations but also the causal effects of environmental exposures. Pregnancy as a DVT risk factor, gender-specific associations, and environmental factors were acknowledged as crucial components of DVT pathogenesis, albeit with inherent complexities. Although the terminology "causal relationship" was used, some findings may be just associations due to the limitation of the MR itself, and actual causal effect needs to be explored with *in-vitro* and *in-vivo* experiments in future studies. In addition, the finding regarding the positive relationship between genetic exposure of protein C and DVT onset should be interpreted carefully in a clinical session. As the effect size of this relative relationship is small, for the univariable MR analysis, it was 1.002 to 1.008 ($p = 0.20e-04$) while for the MVMR analysis, the CI was 1.001 to 1.009 ($p = 0.005$). Although both p values were smaller than the Bonferroni correction threshold, the effect size was quite small and needed to be carefully interpreted. In addition, the effect size for the mediating effect is quite wide, from 2% to 79%, and this might be due to the statistical limitation itself, which needs to be validated in a bigger population study.

In conclusion, our MVMR and mediating analyses offer genetic evidence supporting a positively potential causal association between protein C levels and increased DVT risk, with BMI playing a modest mediating role. The potential implications of our results for DVT prevention policies underscore the need for validation in well-powered randomized clinical trials.

Acknowledgments

Liang Shu, Liyan Sun, Cong Yu – co-first author.

Funding

This work is supported by Transverse Research Project of Shanghai Ninth People's Hospital, Shanghai Jiaotong University, School of Medicine (JYHX202117002), also from Shanghai Pudong Health Commission (PWZxq2022-13).

Ethics approval

Not applicable.

Conflict of interest

The authors declare no conflict of interest.

References

- Middleton P, Shepherd E, Gomersall JC. Venous thromboembolism prophylaxis for women at risk during preg-

- nancy and the early postnatal period. *Cochrane Database Syst Rev* 2021; 3: CD001689.
2. Wang X, Ma Y, Hui X, et al. Oral direct thrombin inhibitors or oral factor Xa inhibitors versus conventional anticoagulants for the treatment of deep vein thrombosis. *Cochrane Database Syst Rev* 2023; 4: CS010956.
 3. Li P, Liu Q, Huang Z, et al. miR-10b-5p regulates venous endothelial cells in deep venous thrombosis by targeting MFG-E8. *Arch Med Sci* 2023. doi: 10.5114/aoms/161674.
 4. Luo P, Yuan Q, Wan X, et al. A two-sample Mendelian randomization study of circulating lipids and deep venous thrombosis. *Sci Rep* 2023; 13: 7432.
 5. Trauscht-Van Horn JJ, Capeless EL, Easterling TR, et al. Pregnancy loss and thrombosis with protein C deficiency. *Am J Obstet Gynecol* 1992; 167: 968-72.
 6. Meng Y, Li Y, Ye YJ, et al. Associations between coagulation factor XII, coagulation factor XI, and stability of venous thromboembolism: a case-control study. *World J Clin Cases* 2022; 10: 2700-9.
 7. Spiezia L, Forestan C, Campello E, et al. Persistently high levels of coagulation factor XI as a risk factor for venous thrombosis. *J Clin Med* 2023; 12: 4890.
 8. Müller-Calleja N, Hollerbach A, Royce J, et al. Lipid presentation by the protein C receptor links coagulation with autoimmunity. *Science* 2021; 371: eabc0956.
 9. Scheiner B, Balcar L, Nussbaumer RJ, et al. Factor VIII/protein C ratio independently predicts liver-related events but does not indicate a hypercoagulable state in ACLD. *J Hepatol* 2022; 76: 1090-9.
 10. Abraitis V, Simoliūniene R, Mongirdiene A, et al. Prevalence of activated protein C resistance among women with recurrent miscarriage. *Medicina (Kaunas, Lith)* 2004; 40: 225-31.
 11. Vukovich T, Auberger K, Weil J, et al. Replacement therapy for a homozygous protein C deficiency-state using a concentrate of human protein C and S. *Br J Haematol* 1988; 70: 435-40.
 12. Christiansen SC, Lijfering WM, Næss IA, et al. The relationship between body mass index, activated protein C resistance and risk of venous thrombosis. *J Thromb Haemost* 2012; 10: 1761-7.
 13. Jeremic A, Mikovic Z, Soldatovic I, et al. Follicular and serum levels of vitamin D in women with unexplained infertility and their relationship with in vitro fertilization outcome: an observational pilot study. *Arch Med Sci* 2021; 17: 1418-22.
 14. Gunes N, Çam H. An observational study of serum vitamin d status in critically ill children admitted to the pediatric intensive care unit. *Arch Med Sci* 2022. doi: 10.5114/aoms/155137.
 15. Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003; 32: 1-22.
 16. Klovaite J, Benn M, Nordestgaard BG. Obesity as a causal risk factor for deep venous thrombosis: a Mendelian randomization study. *J Intern Med* 2015; 277: 573-84.
 17. Willer CJ, Schmidt EM, Sengupta S, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet* 2013; 45: 1274-83.
 18. Laan SW van der, Harshfield EL, Hemerich D, et al. From lipid locus to drug target through human genomics. *Cardiovasc Res* 2018; 114: 1258-70.
 19. Crone B, Krause AM, Hornsby WE, et al. Translating genetic association of lipid levels for biological and clinical application. *Cardiovasc Drug Ther* 2021; 35: 617-26.
 20. Burgess S, Thompson SG, Collaboration CCG. Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol* 2011; 40: 755-64.
 21. Burgess S, Dudbridge F, Thompson SG. Combining information on multiple instrumental variables in Mendelian randomization: comparison of allele score and summarized data methods. *Statist Med* 2015; 35: 1880-906.
 22. Pietzner M, Wheeler E, Carrasco-Zanini J, et al. Genetic architecture of host proteins involved in SARS-CoV-2 infection. *Nat Commun* 2020; 11: 6397.
 23. Sun BB, Maranville JC, Peters JE, et al. Genomic atlas of the human plasma proteome. *Nature* 2018; 558: 73-9.
 24. Hartwig FP, Davies NM, Hemani G, et al. Two-sample Mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique. *Int J Epidemiol* 2017; 45: 1717-26.
 25. Hartwig FP, Smith GD, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol* 2017; 46: 1985-98.
 26. Bowden J, Smith GD, Haycock PC, et al. Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol* 2016; 40: 304-14.
 27. Verbanck M, Chen CY, Neale B, et al. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet* 2018; 50: 693-8.
 28. Burgess S, Bowden J, Fall T, et al. Sensitivity analyses for robust causal inference from mendelian randomization analyses with multiple genetic variants. *Epidemiology* 2017; 28: 30-42.
 29. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife* 2018; 7: e34408.
 30. Kujovich JL. Factor V Leiden thrombophilia. *Genet Med* 2011; 13: 1-16.
 31. Catalano PM, Shankar K. Obesity and pregnancy: mechanisms of short term and long term adverse consequences for mother and child. *BMJ* 2017; 356: j1.