DISSERTATION FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PHD)

Genetic and Environmental Risk Factors Associated with Venous Thrombosis in the Hungarian Population

By Shewaye Fituma Natae

University of Debrecen

Doctoral School of Health Sciences

Debrecen, 2024

DISSERTATION FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PHD)

Genetic and Environmental Risk Factors Associated with Venous Thrombosis in the Hungarian Population

By Shewaye Fituma Natae

Supervisor: Dr.Szilvia Fiatal (MD, PhD)

UNIVERSITY OF DEBRECEN

DOCTORAL SCHOOL OF HEALTH SCIENCES

DEBRECEN, 2024

Table of Contents

List of abbreviations

AT-Antithrombin

ATBp3-Antithrombin Budapest

BMI-Body mass index

CADs- Coronary artery diseases

CKDs -Chronic kidney diseases

CTV- Computed tomography venography

CVDs- Cardiovascular diseases

DM- Diabetes mellitus

DNA-Deoxyribonucleic acid

DVT- Deep venous thrombosis

EU- European union

FGG- Fibrinogen gamma gene

GRS-Genetic risk score

GxE- Gene-environmental interaction

HDL-C- High density lipoprotein cholesterol

LDL-C- Low-density lipoprotein cholesterol

MAF- Mutation analysis core facility

PE-Pulmonary embolism

RCTs- Randomized control trials

SNPs- Single nucleotide polymorphisms

SERPINC1- Serpin Family C member 1

TC-Total cholesterol

TG-Triglyceride

unGRS- Unweighted genetic risk score

VT-Venous thrombosis

VTE- Venous thromboembolism

WC-Waist circumferences

wGRS- Weighted genetic risk score

1 INTRODUCTION

1.1 Definition and concept of thrombosis

Thrombosis is a severe medical condition resulting from the formation of a blood clot in the lumen of the venous or arterial blood vessels, which hinders the normal blood flow process within the circulatory system and might lead to life-threatening illnesses (1,2). It is the principal cause of cardiovascular disease (CVD) related morbidity and mortality, which shares about 25% of all deaths globally (3–6). Based on the type of blood vessels involved, thrombosis is broadly classified into arterial and venous thrombosis (5,6). Arterial thrombosis is the formation of clots within the arterial blood vessels (5), whereas venous thrombosis (VT) is the result of clots formation in the venous blood vessels (1,5,7,8). A clot formed in the venous and arterial vessels results in life-threatening complications such as pulmonary embolism (PE) (9–12), stroke (13–15), and heart attack (1,16,17), respectively. .

1.2 VT presentation

VT is usually presented as venous thromboembolism (VTE), which encompasses deep venous thrombosis (DVT) (1,7,18) and PE (1,7,9–11,18). DVT occurs when clotting blood forms in the large veins, mainly in the leg, thigh, and pelvis veins (18). Swelling, discoloration of the affected area, pain, and dilated superficial veins are the most common clinical features of DVT (12,19,20). It is mainly diagnosed using vascular Doppler ultrasound (1,18); however, computed tomography venography (CTV) is the best and most sensitive diagnostic method (7). Unresolved DVT results in PE due to the migration of the thrombus into the pulmonary arteries, which results in the occlusion of blood flow into the lung (9–12) and leads to severe medical conditions such as death (21,22).

1.3 VTE incidence

VTE is one of the three principal causes of cardiovascular disease-related mortality with a significant genetic predisposition (23–25). An overall estimated VTE incidence rate among people of European ancestry ranges from 104 to 183 per 100,000 person-years (26–31), with higher incidence (65% higher crude incidence rates per 1,000 persons-years) in African-Americans (32–36) and lower in Asians (36–38). African-Americans are more frequently diagnosed with PE than DVT (36) compared to Caucasians ancestry; 60% of VTE cases were present with DVT (29). In European ancestry, DVT is more prevalent among the young, while PE is frequent among the elderly (30).

1.4 VT burden globally and European context

VT is a multifactorial trait that contributes to the burden of CVD worldwide (29,39–41). Although the overall CVD-related morbidity is decreasing in Europe, mortality is substantially high. CVD is the principal cause of mortality in Europe and is attributed to over 3.9 million deaths annually (42–44). Furthermore, approximately 60 million premature deaths (death<70 years) occurred in Europe as a result of CVD (44). VT is a major health problem with a significant incidence (7.62/100,000) and mortality (3.70/100,000) annually (45).

A higher burden of CVD-related mortality was reported in central and eastern European countries (42,46). Hungary shares the highest proportion of this mortality, and CVD remains the prominent cause of death in Hungary (42,47). As of 2014, approximately 35,000 women and 27,000 men died of CVD, accounting for 55% and 45% of all deaths for women and men, respectively (47). The age-standardized death rate from CVD in Hungary is beyond twice the European Union (EU) average reported in 2014 (47). The availability of prophylaxis could significantly avert this burden by early addressing highly vulnerable subjects for the treatments (45,48).

1.5 VT pathogenesis

According to Rudolf Virchow's triad explanation (49), thrombosis is the result of three main factors: stasis of blood flow (50,51), endothelial injury (52–54), and hypercoagulability (55). The heritable prothrombotic factors influence the VT risk via the coagulation process (55,56), whereas the non-heritable risk factors result in VT risk either through stasis or endothelial injury (57,58). Various studies have established the impact of heritable factors on VT risk (59–61). The incidence of hospitalization due to VT was two-fold higher for persons with affected families than for the general population (23–25).

1.6 Heritable VT risk factors and their ancestry distribution

Although abundant single nucleotide polymorphisms (SNPs) provoke the susceptibility of an individual to VT (62–65), it seems that the most strongly associated SNPs: rs6025 (Leiden mutation) in the F5 gene, rs1799963 (prothrombin G20210A) in the coagulation factor two gene (F2), rs8176719 (non-O blood type) in the ABO gene, rs2036914 in the coagulation factor eleven gene (F11), rs2066865 in the fibrinogen gamma gene (FGG), and the rs121909567 (ATBp3 mutation) in the SERPINC1 gene, which play the prominent role in the determination of VT incidence and recurrence in genetically vulnerable individuals (63,66–68).

The Leiden mutation is one of the most dominant heritable VT risk factors that increase the burden of VT in genetically vulnerable individuals (69–71). The prevalence of Leiden mutation /F5 is highest in European descent populations (3-15%) (69) followed by Caucasian Americans (5.2%). However, it is rare in African-Americans (1.2%) and Asian-Americans (0.45%) (69), and is not present at all in African populations (72,73). Likewise, prothrombin gene mutation / F2, often known as the G20210A mutation , is the second most common heritable VT risk in Caucasian ancestry (3-17%), particularly in European ancestry (4%) and Caucasian Americans (2%). However, it is less frequent in African Americans, which is approximately 0.4% (1 in 250), and extremely rare in African and Asian ancestry (74).

Often, due to their joint presence and their possible gene-gene interaction, the prothrombin gene mutation (rs1799963) and Leiden mutation (rs6025) SNPs were studied together (75). Furthermore, studies have shown that the ancestry distribution of coagulation Factor 11 (rs2036914) is similar in both Caucasian ancestry and African American ancestry (76,77). Studies have indicated that O blood type individuals are at lower risk of VT disease than other non-O blood type individuals (78,79); whereas the possibility of VT disease is higher among non-O blood type subjects (80–83). In addition, Fang *et al.* reported that the risk of venous thromboembolism (VTE) is higher for the African American ancestry and non-O blood type individuals than the Caucasian ancestry and O blood type individuals (84). Moreover, studies indicated that FGG (rs2066865) is more prevalent in Caucasian ancestry than African American ancestry (85,86).

Furthermore, studies have found that antithrombin deficiency plays a vital role in the pathogenesis of VTE. Antithrombin is an essential inhibitor of blood coagulation proteases; individuals with hereditary AT deficiency have elevated thrombotic risk (87–89). Studies have revealed that the mutation profile of the AT gene (SERPINC1) is heterozygous (90–93). Formerly, it found that the prevalence of ATBp3 mutation was relatively high in the Roma population but not in the general Hungarian population (66).

1.7 Genetic risk score (GRS) and VT risk

Being a carrier for multiple variants increases the vulnerability to VT disease. Studies indicated that the combination of strongly associated VT SNPs [rs6025(F5), rs1799963 (F2), rs8176719 (ABO), rs2036914 (F11), and rs2066865 (FGG)] poses a greater risk of VT than that risk occurring by an individual SNP (63,94). The genetic risk score (GRS) of strongly associated VT variants results in the highest risk compared with several SNP combinations. Haan *et al*. showed that the VT risk prediction of the 5-SNP score is equivalent to that of the 31-SNP risk score (63).

1.8 Non-heritable VT risk factors

Different studies conducted elsewhere indicated that genetic and non-heritable VT risk factors have increased the susceptibility of an individual to VT diseases (39,63). Aging and obesity are well-known non-heritable VT risk factors that hasten the occurrence of VT (63). As a result of multiple anatomical and pathophysiological changes, elderly individuals are prone to age-related cardiovascular morbidity and mortality (95–97). Aging plays a principal role in the higher incidence of VT risk (1%) in elderly individuals (53,95–98). The diminished efficiency of the calf muscle pump due to aging could lead to reflux and stasis of peripheral blood that lead to thrombosis formation (95). Furthermore, age-related endothelial dysfunction also contributed to the higher incidence of VT in the elderly compared to younger individuals (53). Valve thickness, muscle fiber atrophy, and reduced anticoagulant properties of the endothelium are some pathophysiological changes that increase the risk of VT among elderly individuals (95,96).

Correspondingly, obese individuals were at higher risk of VT than normal-weight individuals. Formerly conducted studies showed that the risk of VT was 2-6-fold higher in obese individuals than in normal-weight individuals (BMI 20 to 24.9 kg/m^2) (99–103). A study indicated that the risk of VT was higher among aged (>50 years old) and obese individuals (102). Furthermore, the existence of other comorbidities such as cancer, diabetes mellitus (DM), chronic kidney diseases (CKD), and coronary artery diseases (CAD) increased the likelihood of VT(104–108). Furthermore, other personal and environmental risk factors such as migraine (109–111), smoking cigarettes, depression, and high levels of lipoprotein are also increasing the VT risk. VT risk is 1.3 to 1.7 times higher among current cigarette smokers than non-smokers (112–114). Harmoniously, smoking cigarettes related to an absolute risk increases of 24.3 VT cases per 100,000 person-years (114).

Studies indicated that depression is a plausible risk factor for VT (115–117); A cohort study conducted by Lee *et al*. revealed that the risk of VT was 1.38 times higher among depressive individuals than among non-depressive (115). Another systematic review and meta-analysis study showed that the risk of VT was 1.3 times higher among depressed subjects than non-depressed (116).

In addition, a high level of low-density lipoprotein cholesterol (LDL-C) contributes to the development of VTE. González-Ordóñez *et al*. reported that the risk of VTE was 2-fold higher among individuals with a high level of LDL-C compared to subjects with a normal level of LDL-C (118). A meta-analysis study of randomized control trials (RCTs) indicated that VTE risk reduced among patients who received statin treatment for a high level of LDL-C (117). Moreover, a cohort study showed a statistically significant association between lipoproteins (TG and LDL-C) and VTE risk in bivariate analysis; however, the association disappeared after controlling for possible confounders (119).

1.9 Hungarian Roma population and CVDs risk

Some populations are susceptible to cardiovascular disease due to the coexistence of genetic and environmental risk factors. The Roma population is the most marginalized ethnic group in Central-Eastern European countries, with an estimated population of 8 to 12 million. They experience social exclusion; which intensely affects their health (120). A higher burden of disease, low life expectancy, low socioeconomic status, low education, and hazardous practices are common among Roma minorities (121–125). Subsequently, cardiovascular risk factors are prevalent in the Roma population (126,127). On the other hand, due to the restriction of health-related data collection by ethnicity in the Hungarian Roma population (128), there are no data related to the incidence or prevalence of VT for this population. However, recent studies indicate that the Roma population is at higher risk of VT due to an elevated prevalence of metabolic syndrome (129) and several heritable risk factors (130).

1.10 Significance of our study

Our previous research concluded that the Roma population seems to have increased genetic susceptibility to VT and suggested further study to compare the gene-environmental interaction (GxE) for VTE risk in the general Hungarian and Roma populations (130). Understanding how genetic and environmental risk factors interact provides insight for the early identification of risk groups within populations, allowing appropriate preventive and therapeutic measures to be taken (131,132).

To our knowledge, scant evidence compares gene and environmental risk factors associated with venous thrombosis among the general Hungarian and Roma populations. Thus, the main aim of the current study was to explore the interaction of environmental risk factors with six selected prothrombotic SNPs [(rs121909567 (SERPINC1), rs1799963 (F2), rs2036914 (F11), rs2066865 (FGG), rs6025 (F5), and rs8176719 (ABO)] (63,66,130). In addition, we aimed to investigate the distribution of the rs121909567 (ATBp3) mutation (Antithrombin Budapest 3 mutation) in the SERPINC1 gene in the Hungarian population, which contributes to the vast majority of antithrombin (AT) deficiencies in the Hungarian Roma population due to its founder effect (133).

Moreover, the stratification of higher-risk individuals based on their genetic profiling for thromboprophylaxis is essential for efficient and effective utilization of available resources (63). Even though the 5-SNPs' [rs6025 (*F5 Leiden*), rs2066865 (*FGG*), rs2036914 (*F11*), rs8176719 (*ABO*), and rs1799963 (*F2*)] impact on venous thrombosis risk is high, the association and risk prediction models are seldom establish by using only those SNPs. To our knowledge, there is a lack of studies investigating the VT risk prediction ability of the combined five strongly associated prothrombotic SNPs in clinically confirmed VT subjects in the Hungarian population.

2 OBJECTIVES

2.1 General objective

The main aim of this study is: to investigate the status of selected genetic and environmental risk factors and the combined VT risk predictability of the five prothrombotic single nucleotide polymorphisms in the Hungarian population.

2.2 Specific Objectives:

- To assess VT prevalence in the Roma and general Hungarian populations
- To identify and compare the gene-environmental interactions and venous thromboembolism risk in the Roma and general Hungarian populations
- To explore the heritable VT risk in the Hungarian population
- To determine the VT risk predictability of the combined five prothrombotic single nucleotide polymorphisms and conventional VT risk factors in the Hungarian population.

3 Methods and Materials

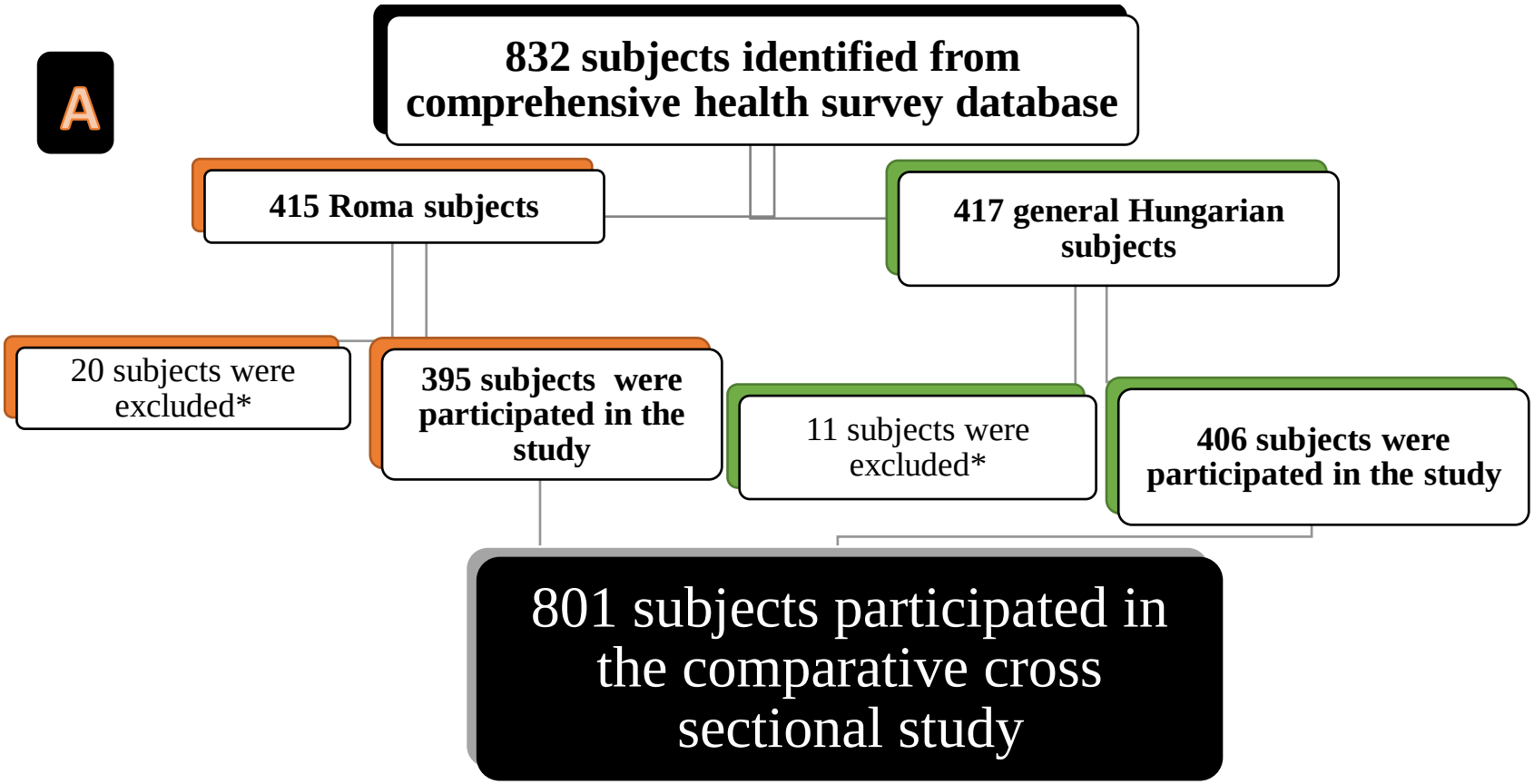
3.1 Study design, sample size, and data source

In this study, we used a comparative cross-sectional (study 1) and case-control study design (study 2). In total, 1532 subjects were identified to participate in this study (832 for comparative study and 700 for case-control study). However, due to incomplete data: 31 subjects (20 Roma and 11 subjects from the general Hungarian population) were excluded from the final analyses of the first study. Only those subjects who had complete genotype and covariate data (395 Roma and 406 general Hungarian subjects; totally 801 subjects) were involved in this study (Flowchart 1A). Data management was similar in study 2, i.e., two subjects were excluded due to missed genotype and covariate data. Finally, 698 (298 clinically confirmed VT cases and 400 healthy controls) participants were involved in the second study (Flowchart 1B).

The two counties found in northeast Hungary namely Szabolcs-Szatmár-Bereg and Hajdú-Bihar were taken as the source population. The population where the Hungarian Roma predominantly resided and the segregated Roma colonies located are taken as the study population. All the study populations involved in both studies (a comparative cross-sectional (1st study) and control group of a case-control (2nd study)/except the 298 cases were taken from a comprehensive health survey database collected for a cardio-metabolic comparative study among the Roma and general Hungarian population. A comprehensive health survey database has 1000 study participants (500 general Hungarian and 500 Roma population) who met the selection criteria (Age 20-64 years old and residence of the indicated location). The study subjects who met the selection criteria were randomly selected from each household of the identified colony. In total, 25 colonies were randomly selected and 20 households from each colony were randomly identified for the study. One eligible individual was involved in the study from each household.

Data collection was made using structured questionnaires, physical examinations, and laboratory-based investigations from each study unit. The questionnaire-based data were collected through face-to-face interviews with data collectors who previously had data collection experience. Data collection related to physical examination and blood samples was done by general practitioners (GPs), and the whole data collection process was supervised by public health coordinators daily. The details of data collection and subject recruitment process were published by Ádány et al (2020) (134).

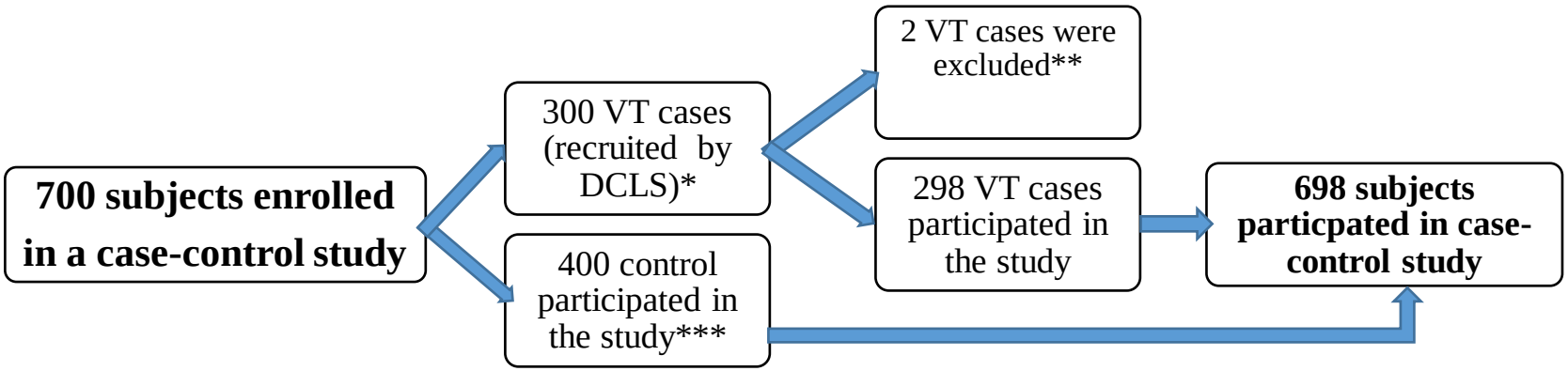
The case-control study design was conducted among 400 healthy control and 298 clinically confirmed VT cases to explore the heritable VT risk and to estimate the VT risk predictability of the combined strongly associated five SNPs (rs6025 (*F5 Leiden*), rs2066865 (*FGG*), rs2036914 (*F11*), rs8176719 (*ABO*), and rs1799963 (*F2*) and well-known non-heritable VT risk factors (age, obesity, and sex) in the Hungarian population. The control subjects were recruited from a comprehensive health survey database (134), whereas, the VT cases were recruited by the Division of Clinical Laboratory Science (DCLS), Department of Laboratory Medicine, Faculty of Medicine, University of Debrecen for one year. Color Doppler ultrasound or phlebography was used to diagnose the VT cases at the Department of Internal Medicine, Faculty of Medicine, University of Debrecen.



*Subjects were excluded from the study due to incomplete genotype and covariate data

Flowchart 1 A: Sample size and sampling procedure of comparative cross sectional study in general Hungarian and Roma population.

B



Flowchart 1 B: Sample size and sampling procedure of case-control study in the Hungarian population.

*Clinically confirmed VT cases were recruited by Division of Clinical Laboratory Science, Department of Laboratory Medicine, Faculty of Medicine, University of Debrecen, Hungary. **VT cases were excluded due to incomplete genotype data.

*** Extracted from Comprehensive health survey database

3.2 Personal and environmental risk factors for comparative study in the general Hungarian and Roma population

For a comparative cross sectional study, the environmental, and personal risk factors data that affect the VTE risk of an individual (104,105,108,116,135–137) were extracted from a comprehensive health survey database (134). Generally, the following data were included:

i) Chronic diseases such as cancer, DM, CAD, and CKD
ii) Lipid parameters (TC, LDL-C, HDL-C, and TG),
iii) Lifestyle factors and anthropometric data such as smoking status, BMI, and WC were collected from the comprehensive health survey database (134).

Data regarding the VT status of an individual and chronic disease (cancer, DM, CAD, and CKD) that has proven as potential VTE risk factors (104,108,138) were collected via a self-reported survey. As a result, those subjects who replied “Yes” to relevant survey questions were considered positive for that specific disease and those who replied “No” were considered negative for that specific disease. The smoking status also assesses their current smoking status; so, if they replied as currently smoking, consider them as smokers if not non-smokers.

Likewise, based on the following three questions, the VTE status of the study participants was also assessed: (i) Did you have a thrombosis in the last 12 months? (ii) Have you been diagnosed with thrombosis? And (iii) Have you received hospital treatment for thrombosis? Consequently, if the respondents replied "Yes" to either of those questions we considered them as "VTE cases" while others who said, "No" as "non-cases.

3.3 Personal and environmental risk factors for case-control study

For a case-control study survey, we only considered the genotypic data and the well-known conventional VT risk factors such as age, obesity, and sex for comparing the VT risk distribution among VT cases and control subjects in the Hungarian population. To verify the risk of VT among elderly and obese subjects, we used age $\geq$60 years, and BMI $\geq$30 kg/m^2 as cut-off values during the data analysis and we used ages below 60 years and BMI below 30 kg/m^2 as reference.

3.4 Genotyping Process

3.4.1 DNA Extraction

DNA was extracted from peripheral blood samples using the MagNA Pure LC system (Roche Diagnostics, Basel, Switzerland) and the MagNA Pure LC DNA Isolation Kit–Large Volume. The extraction process was carried out following the manufacturer's instructions, ensuring high yield and purity. The extracted DNA was then eluted in 200 µl of the elution buffer provided with the MagNA Pure LC DNA Isolation Kit–Large Volume.

3.4.2 Assay Design and Genotyping

The design of the genotyping assays and the actual genotyping were carried out at the Mutation Analysis Core Facility (MAF) of the Karolinska University Hospital, Stockholm, Sweden. The genotyping was performed using the MassARRAY platform (Sequenom, CA, USA) with iPLEX Gold chemistry.

3.4.2.1 MassARRAY Platform and iPLEX Gold Chemistry

The MassARRAY platform is a highly sensitive and accurate system for detecting genetic variations, including single nucleotide polymorphisms (SNPs) and insertions/deletions (indels). The iPLEX Gold chemistry utilized in this platform involves several key steps:

3.4.2.1.1 PCR Amplification

Target regions containing the genetic variants of interest are amplified using polymerase chain reaction (PCR). This step ensures that there is sufficient DNA to work with in subsequent steps.

3.4.2.1.2 Shrimp Alkaline Phosphatase (SAP) Treatment

The amplified PCR products are treated with SAP to dephosphorylate any unincorporated nucleotides. This step is crucial to prevent interference in the subsequent single base extension reaction.

3.4.2.1.3 Single Base Extension (SBE) Reaction

In this step, a primer that anneals just upstream of the SNP site is extended by a single base using a mixture of mass-modified terminators. Each terminator has a distinct mass, allowing for the identification of the specific nucleotide present at the variant site.

3.4.2.1.4 MALDI-TOF Mass Spectrometry

The extended primers are then analyzed using Matrix-Assisted Laser Desorption/Ionization Time-of-Flight (MALDI-TOF) mass spectrometry. The mass of the extended primer

indicates which nucleotide was incorporated, thereby identifying the genotype at each variant site.

3.4.3 Quality Control and Validation

Quality control, validation, and concordance analysis were rigorously conducted by the MAF to ensure the accuracy and reliability of the genotyping results.

The quality control process includes several measures:-

3.4.3.1 Replicate Testing

A subset of samples is genotyped in duplicate to assess reproducibility and consistency of the results.

3.4.3.2 Known Controls

Samples with known genotypes are included in each batch to serve as controls, ensuring that the genotyping process is functioning correctly.

3.4.3.3 Concordance Analysis

The genotyping results are compared with previously obtained data (if available) or with results from an alternative genotyping method to confirm their accuracy.

3.4.3.4 Call Rate and Quality Scores

The platform provides call rates and quality scores for each SNP, and only those with high-quality scores and high call rates are included in the final analysis.

This comprehensive approach to genotyping, combined with stringent quality control measures, ensures that the genetic data used in this study are both accurate and reliable, forming a solid foundation for subsequent analyses of genetic factors and their effects.

3.5 SNPs selection

For both studies, the five strongly associated and influential prothrombotic SNPs: rs1799963 (*F2*), rs6025 (*F5*, Leiden), rs2066865 (*FGG*), rs2036914 (*F11*), and rs8176719 (*ABO*) were considered due to their large effect size on the GWA (64,139,140) studies and strong relationships with the higher incidence and recurrence rate of VT risk in the formerly conducted studies (63,130). In addition, due to the indicative result of previous studies (66,133), which identified *SERPINC1* (rs121909567)/ Antithrombin Budapest3 (ATBp3 mutation ; p.Leu131Phe) as the common cause of the antithrombin (AT) deficiency in Hungary; particularly the Roma population we consider it in our comparative cross-sectional study.

3.6 Genetic risk score (GRS) computations

To establish the pooled effect of the included prothrombotic SNPs: rs6025 (*F5* Leiden), rs2066865 (*FGG*), rs2036914 (*F11*), rs8176719 (*ABO*), and rs1799963 (*F2*) on VT risk we also computed weighted (wGRS) and unweighted (unGRS) genetic risk score. In the genetic risk score (GRS), the individuals were assigned a score based on the number of risk alleles they carried. Accordingly, "0", "1", and "2" codes were given for the absence of risk alleles, heterozygous, and homozygous for risk alleles, respectively. When the risk allele was found protective, the coding for the homozygous risk allele becomes "0", while "2" for the other homozygous allele (130). Unweighted-GRS (unGRS) was simply computed by adding all risk alleles in the loci by assuming that all alleles have the same effect. Whereas, weighted genetic risk score (wGRS) was computed with the assumption that, SNPs with larger effects contributed more to the GRS. Weights were derived from the risk coefficient for each allele depending on the odds ratio (OR) reported in the former genetics associations study (63). For this study, only those SNPs who had an effect size in the previously conducted study were included for computing wGRS. Median values of wGRS and unGRS were used to compare the association between genetics risk score and VTE risk factors in the study populations.

3.7 Statistical Analysis used

3.7.1 Comparative cross sectional study (study 1)

Statistical tests were computed using IBM SPSS Version 26 statistical software. The Shapiro-Wilk normality test was used to test the distribution of quantitative variables. Non-normally distributed variables were transformed using a two-step transforming approach proposed by Templeton (141). The presence of Hardy-Weinberg Equilibrium (HWE) and allele frequency differences of all included SNPs between the two study populations was evaluated by using PLINK statistical software Version 1.9. Logistic regression analysis was used to determine the associations between individual SNPs, environmental risk factors, and VTE. A multivariable linear regression analysis with 95% CI, was used to test the impact of GxE on VTE risk after interaction terms between 6 SNPs and environmental risk factors were generated. An interaction term was created between each SNP and environmental risk factor to assess the combined effect of them on VTE risk (142,143).

Three categories of odds ratios (ORs) were defined based on the GxE assumption models: subjects who were unexposed to environmental risk factors and free from risk variants (also known as wild type) (E=G=0), were used as reference groups (OR_{00}); OR_{11} represented subjects

with both genetic and environmental risk factor exposure (E=G=1); OR_{10} = subjects were exposed to environmental risk factors but not with genetic risk (E=1, G=0); and OR_{01} = subjects were with only genetic exposure but not environmental risk exposure (E=0, G=1). Thus, under the multiplicative interaction model, if $OR_{11}= OR_{01}\times OR_{10}$, there was no interaction between genetic and environmental risk factors; however, if $OR_{11}\neq OR_{01}\times OR_{10}$, there was a multiplicative interaction between the given environmental risk factors and genetic risk factors, which was either a synergistic interaction ($OR_{11}> OR_{01}\times OR_{10}$) or an antagonistic interaction, if $OR_{11}< OR_{01}\times OR_{10}$ (143). For the additive model, no interaction was concluded based on the null hypothesis (H_0) of $OR_{11}= OR_{01}+ OR_{10}-1$; however, the interaction would be considered synergistic when $OR_{11}> OR_{01}+ OR_{10}-1$ or antagonistic if $OR_{11}< OR_{01}+ OR_{10}-1$ (143).

During the analysis, the Hungarian general and Roma population samples were combined and ethnicity was included as a variable in the model to differentiate the effect on VTE risk. For this study, the Hungarian general population was used as a reference population to make a comparison of GxE interaction and VTE risk in the two populations. All analyses were adjusted for age and reported p-values were two-sided, and α level of 0.05 was used to define statistical significance.

3.7.2 Case-control study (study 2)

Statistical tests were computed using PLINK (version 1.9) and IBM SPSS (version 26.0) statistical software. The Mann–Whitney U test was used to compare the age, BMI, and GRS distribution in the study populations. The Shapiro–Wilk normality test was used to test the distribution of quantitative variables. The presence of Hardy-Weinberg equilibrium (HWE) and risk allele frequency differences between VT cases and controls were estimated using the X^2 test. The association between the five SNPs and VT disease risk was assessed by the odds ratio (ORs) with their respective 95% confidence interval (CI) under all genetic models: the multiplicative model, the additive model, the dominant model, and the recessive, and genotypic model. Likewise, logistic regression analysis was also used to compute the OR with 95% of individual SNPs, genetic, non-heritable, and combined VT risk factors.

In addition, the receiver-operating characteristic (ROC) curve was determined to assess how well its score classifies venous thrombosis patients and control subjects. In general, the AUC ranges from 0.5 (no discrimination between VT patients and control subjects) to 1.0 (perfect discrimination). To determine which SNP is more influential in their discriminatory accuracy of the area under the receiver operating characteristics curve (AUC), we added each SNP one-by-one into a model. Therefore, we started with the SNP with the highest odds ratio (OR), i.e.,

the Leiden mutation (rs6025) in the F5 gene, and assessed whether including more additional SNPs into a model could improve the AUC value. We continued adding all other SNPs into a model until we verified that the addition of more SNPs into a model could not reveal any more significant change in discriminatory accuracy.

We compared the AUCs of genetic, non-heritable, and combined risk models. We included each non-heritable risk factor and their combination with genetic VT risk factors into a model to verify the difference in AUC value and their VT risk predictability in the study population. A logistic regression model was used to generate a combined risk score of genetic and non-heritable VT risk factors. SPSS version 26.0 was used to calculate ROC curves and AUCs. Bonferroni multiple testing was employed to prevent the problem associated with multiple comparisons ($0.05/5= p<0.01$). Statistical significance was accepted at 5% level.

3.8 Ethical approval

The committee of the Hungarian Scientific Council on Health Research approved the protocol (61327-2017/EKU). Written informed consent was obtained from all study subjects.

4 Results

4.1 Characteristics of the study participants

In total, 1499 subjects (406= general Hungarian, 395=Roma population, 400= Healthy control, and 298=clinically confirmed VT cases) who had complete genotype and covariate data were included in both studies. In the first study, the female proportion was higher and statistically significant in the Roma population than the general Hungarian (73.9% vs. 55.4%; p<0.001) (Table 1).

Table 1: Demographic and Anthropometric characteristics of the study participants (study1)

Variables	**General Hungarian (N=406) (Mean, 95% CI)**	**Roma (N=395) (Mean, 95% CI)**	***p*-value**
Age (years)	44.3 (43.1 - 45.5)	43.5 (42.2 - 44.7)	0.35
Female (%)	**55.4**	**73.9**	**<0.001**
BMI (kg/m^2)	27.2 (26.7-27.7)	27.5 (26.8-28.2)	0.48
WC (cm)	96.0 (94.5-97.5)	95.0 (93.3-96.7)	0.38

BMI =Body mass index, WC= waist circumference

Table 2: Mean value distribution of lipoproteins and fasting blood glucose of the study participants (study 1)

Variables	**general Hungarian (N=406) (Mean, 95% CI)**	**Roma (N=395) (Mean, 95% CI)**	***p*-value**
Cholesterol (mmol/l)	5.0 (4.9-5.1)	4.9 (4.8-5.0)	0.45
LDL-C (mmol/l)	3.1 (3.0-3.2)	3.1 (3.0 -3.2)	0.31
HDL-C (mmol/l)	**1.4 (1.3-1.4)**	**1.3 (1.2-1.3)**	**<0.001**
TG (mmol/l)	1.6 (1.5-1.7)	1.7 (1.6-1.9)	0.29
FBG (mmol/l)	5.3 (5.1-5.5)	5.2 (5.0-5.4)	0.51

LDL-C= Low-density lipoprotein cholesterol, HDL-C = High density lipoprotein cholesterol, TG= Triglycerides, FBG= Fasting blood glucose

However, in study 2; the elderly (Figure 1) and male (Figure 2) subjects were highly prevalent in the cases group than in the control group. The observed differences were statistically significant. The mean age of the VT cases subjects was higher than that of the control group (63.4±16.4 years vs. 43.8±12.6 years, p<0.001), respectively.

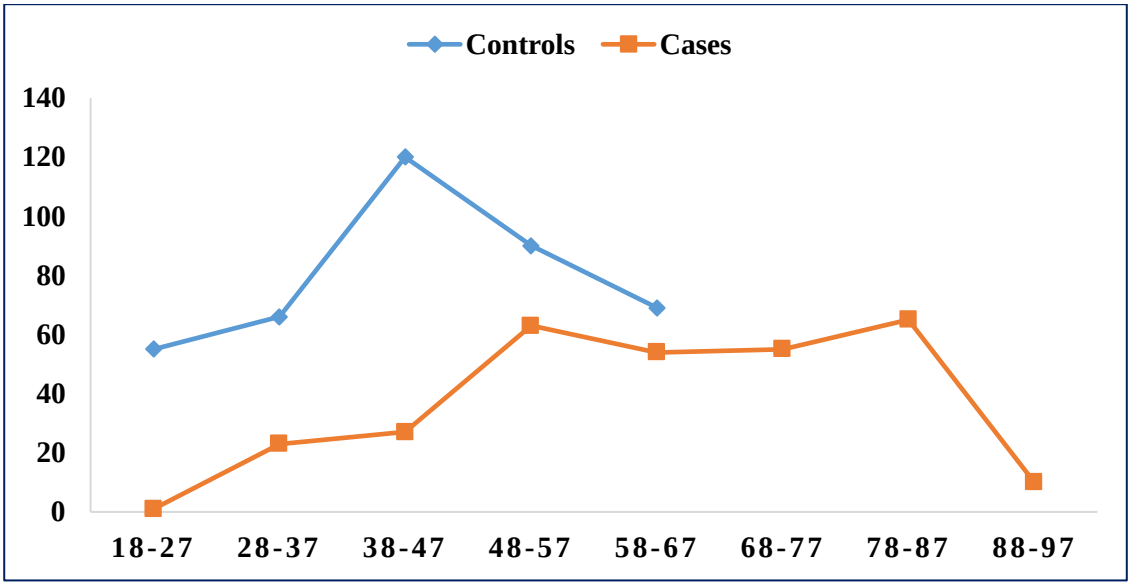


Figure 1: Age distribution of VT patients and the control group of the Hungarian population

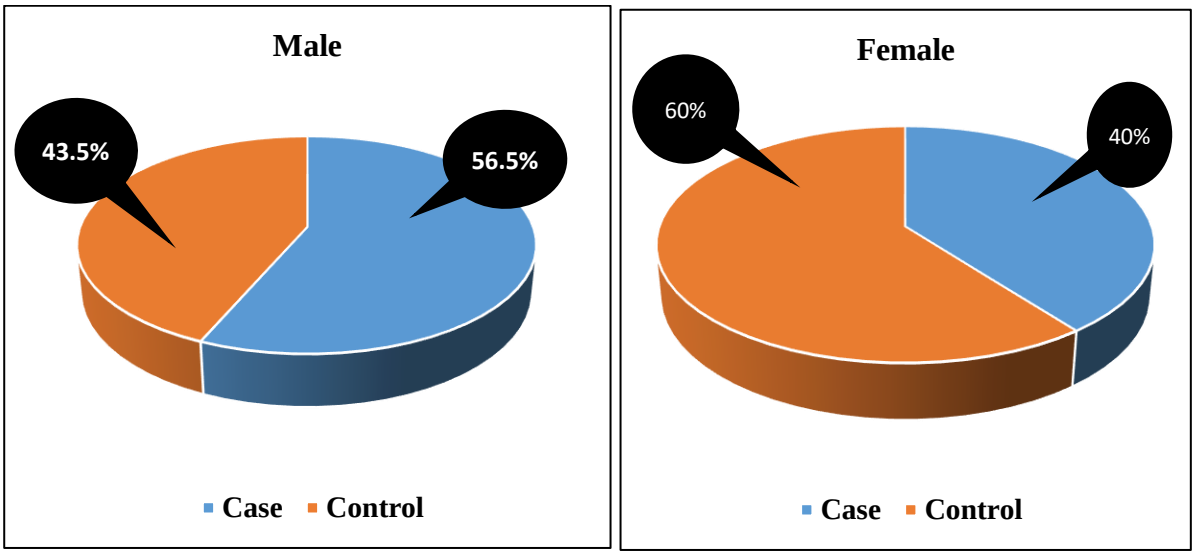


Figure 2: VT case distribution by sex in the study populations.

4.2 VTE morbidity in the study populations

In study 1, only 6 (1.5%) and 12 (3.0%) of the general Hungarian and Roma populations reported VTE during the survey, respectively. The proportion of VTE cases was higher among the Roma population; in particular, the VTE risk was higher among female subjects (Table 3). However, in the case-control study, the VT cases were higher among the male participants than the female participants (Figure 2).

Table 3: Distribution of VTE cases by population and sex among the study subjects

Sex	general Hungarian (N=406)		Roma (N=395)	
	Yes (%)	**No (%)**	**Yes (%)**	**No (%)**
Male	4 (2.2)	177 (97.8)	4 (3.9)	99 (96.1)
Female	2 (0.9)	223 (99.1)	8 (2.7)	284 (97.3)
Total	6 (1.5)	400 (98.5)	12 (3.0)	383 (97.0)

4.3 Risk allele frequency comparison in the study population

Table 4 presents the risk allele frequencies comparison of the six prothrombotic SNPs: rs6025 (*F5*), rs2066865 *(FGG*), rs2036914 (*F11*), rs1799963 (*F2*), rs8176719 (*ABO*), and rs121909567 (*SERPINC1* included in both studies (comparative cross section study and case-control study). Genotype results were available for 801 subjects (406 general Hungarian and 395 Roma population in the comparative study and 698 (298VT cases and 400 healthy control) for the case-control study. All included SNPs were tested to determine whether the observed allele frequencies were in accordance with HWE; no significant deviation from HWE was detected.

In study 1, only the allele frequency of the *SERPINC1* SNP (rs121909567) remains significantly distinct in the study populations after multiple correction testing. The *SERPINC1* (rs121909567) SNP is only present in the Roma population and does not occur in the general population (Table 4). In the case-control study (study 2), the risk allele frequencies of the rs6025 (Leiden mutation), rs2036914 (*F11*), and the rs8176719 (*ABO*) were higher in the case group than in the control group, and the differences remained statistically significant after multiple correction testing ($p<0.01$) (Table 4). Nevertheless, before the multiple correction testing except for *F2* or Prothrombin mutation (rs1799963) the allele frequencies of the remaining four SNPs: rs6025 (*F5*), rs2066865 (*FGG*), rs2036914 (*F11*), rs8176719 (*ABO)* were significantly distinct ($p<0.05$) among cases and control group (Table 4).

Table 4: Risk allele frequency distribution comparison in the comparative cross-sectional and case-control studies.

Gene	SNPs	Risk allele	Comparative cross-sectional study (N=801)			Case-control study (N=698)		
			general Hungarian (N=406), Frequency (%)	**Roma (N=395) Frequency (%)**	***p-value***	**VT cases (N=298) Frequency (%)**	**Control (N=400) Frequency (%)**	***p-value***
F5	**rs6025**	**T**	0.07 (7%)	0.09 (9%)	0.2	**0.20 (20.3%)**	**0. 07 (6.75%)**	**<0.001***
FGG	rs2066865	A	0.23 (23%)	0.28 (28%)	0.05	0.28 (28.36%)	0.23 (23.12%)	0.03
F11	**rs2036914**	**C**	0.54 (54%)	0.51 (51%)	0.2	**0. 62 (61.91%)**	**0.54 (54.12%)**	**0.003***
ABO	**rs8176719**	**G**	0.47 (47%)	0.48 (48%)	0.6	**0.40 (40.44%)**	**0.47 (47.38%)**	**0.001***
F2	rs1799963	A	0.02 (2%)	0.01 (1%)	0.04	0.03 (3.02%)	0.02 (2%)	0.22
SERPINC1	**rs121909567**	**A**	**0 (0%)**	**0.01 (1%)**	**<0.001***	NA	NA	NA

p<0.01 considered significant after multiple correction testing, NA= not applicable (this SNP is not included in the case-control study)

4.4 Associations between SNPs and VT risk in the study populations

4.4.1Comparative cross-sectional study (study 1)

4.4.1.1 Associations of individual SNPs with VTE

As shown in **Table 5,** only the Leiden mutation (rs6025) and the Fibrinogen gamma gene (*FGG*) (rs2066865) were significantly associated with VTE risk and found to be nominally significant for the Roma population but not for the general population. Study subjects who have homozygous risk alleles for the *FGG* gene were 5.9 times more at risk for VTE than subjects without a risk allele (OR=5.9; 95% CI: [1.23,28.4]). Furthermore, people with the heterozygous risk allele of the Leiden mutation (rs6025) were 3.8 times more likely to develop VTE than individuals without the risk allele, and the risk was higher for the Hungarian Roma population.

However, this significance disappeared ($p>0.0083$) after adjustment to Bonferroni multiple correction test. Furthermore, our study reveals that some individuals who had VTE were carriers for homozygous risk variants of multiple SNPs. In our study, three Roma subjects who had VTE were carriers of the homozygous risk variant rs2036914 (*F11*), and another 3 individuals carried the homozygous variant of rs2066865 (*FGG*), as depicted in Table5. However, 2 out of the six Roma subjects who had VTE and were carriers for homozygous risk variants of rs2036914 (*F11*) and rs2066865 (*FGG*) also carried homozygous risk variants of both SNPs.

Table 5: Comparisons of associations between individual SNPs and VTE in general Hungarian and Roma populations

Gene(SNPs)	**Genotype**	**VTE Cases**					
		general Hungarian			**Roma**		
		Yes	**No**	**OR (95%CI)**	**Yes**	**No**	**OR (95%CI)**
SERPENC1 (rs121909567)	G G	6 (1.5)	400 (98.5)	NA‡	12 (3.1)	372 (96.9)	1.00
	G A	0 (0.0)	0 (0.0)		0 (0.0)	11 (100)	2.2E-8
F2 (rs1799963)	G G	6 (1.5)	384 (98.5)	1.00	12 (3.1)	377 (96.9)	1.00
	G A	0 (0.0)	16 (100)	4.4E-8	0 (0.0)	6 (100.0)	2.9E-8
F11 (rs2036914)	T T	1 (1.2)	83 (98.8)	1.00	3 (3.0)	97 (97.0)	1.00
	C T	4 (2.0)	201 (98.0)	1.4 (0.13,13.9)	6 (3.2)	184 (96.8)	0.96 (0.22,4.2)
	C C	1 (0.9)	116 (99.1)	0.49 (0.03,9.5)	3 (2.9)	102 (97.1)	1.15 (0.21,6.4)
***FGG* (rs2066865)**	G G	2 (0.9)	233 (99.1)	1.00	5 (2.4)	201 (97.6)	1.00
	G A	3 (2.0)	149 (98.0)	2.5 (0.39,16.2)	4 (2.5)	156 (97.5)	0.95 (0.24,3.8)
	A A	1 (5.3)	18 (94.7)	5.6 (0.43,73.9)	3 (10.3)	26 (89.7)	**5.9 (1.23,28.4)***
***F5* (rs6025)**	C C	4 (1.1)	347 (98.9)	1.00	8 (2.4)	323 (97.6)	1.00
	C T	2 (3.7)	52 (96.3)	3.5 (0.57,21.3)	4 (6.7)	56 (93.3)	**3.8 (1.01,14.2)***
	T T	0 (0.0)	1 (100)	1.1E-7	0 (0.0)	4 (100.0)	1.1E-8
ABO (rs8176719)	C C	5 (4.2)	114 (95.8)	1.00	5 (4.7)	101 (95.3)	1.00
	C DEL	1 (0.5)	193 (99.5)	0.11 (0.01,0.9)	5 (2.5)	192 (97.5)	0.49 (0.133,1.9)
	DEL DEL	0 (0.0)	93 (100.0)	1.6E-8	2 (2.2)	90 (97.8)	0.43 (0.08,2.5)

‡=no risk allele found in the Hungarian general for rs121909567 SNP, *=p<0.05. Significant differences are highlighted

4.4.1.2 Gene-environmental interaction (GxE) and VTE risk

The multivariate linear regression analysis was used to calculate the standardized beta value of GxE and VTE risk. We found several statistically significant multiplicative interactions on VTE risk for the majority of the included SNPs, such as rs2036914 (*F11*), rs2066865 (*FGG*), rs6025 (*F5*, Leiden), and rs8176719 (*ABO*). Since the ATBp3 mutation was not present in the general population, regression analysis was only performed for the Roma population. Although the trend of their relationships indicated the possibility of higher VTE risk, the multiplicative interaction was not statistically significant for the *SERPINC1* (rs121909567) and F2 (rs1799963) genes (p-value for interaction >0.05). The regression coefficient of the multivariate linear regression analysis indicates that the likelihood of VTE was higher among the study subjects who had cancer, DM, CAD, and CKD, experienced migraine and depression in addition to their genetic susceptibility to VTE. Furthermore, cigarette smoking, a high level of LDL-C, and obesity increased the VTE risk for the study subjects.

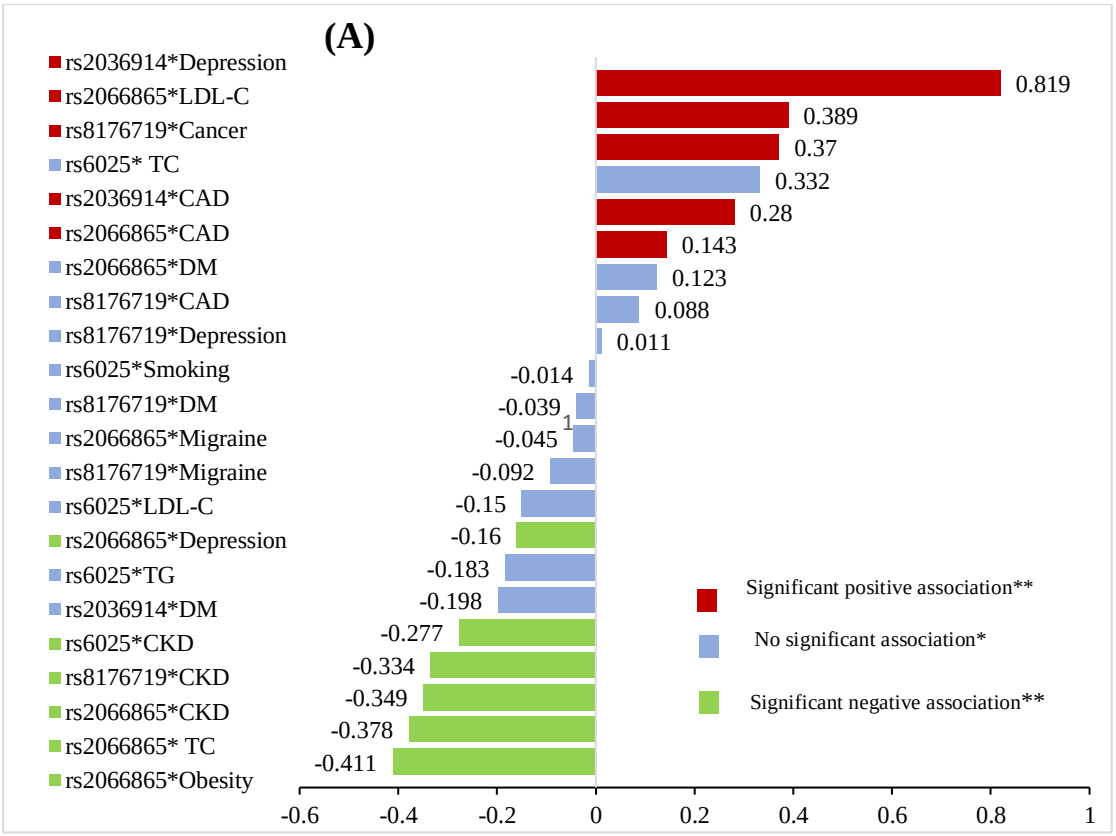
Figure 3 shows that the standardized beta (β) values of the GxE and VTE risk from the linear regression analyses were significant in both study populations. The observed multiplicative interaction and VTE risk were bidirectional: a positive beta (β) value indicated a VTE risk increment as a result of GxE (red color), whereas a negative beta (β) value indicated the reverse (risk decrements) (green color).

The risk of VTE was significantly higher (β= 0.819, p= 0.02) among depressive Roma individuals with the rs2036914 risk variant, but not for general Hungarian individuals (β= 0.343, p= 0.33). The presence of high levels of LDL-C and the rs2066865 (*FGG*) risk variant makes Roma subjects to be at higher risk of VTE (β= 0.389, p= 0.002); however, the joint presence of those risk factors did not increase the VTE risk in the general subjects (β=0.048, p=0.70). The existence of a multiplicative interaction between CAD and rs2036914 (*F11*) increases the VTE risk among both populations: for the Roma population β= 0.280 (p= 0.001) and for the general Hungarian population β= 0.423 (p= 0.001)

As a result of the multiplicative interaction between rs2066865 (*FGG*) and CAD, VTE risk was higher for the Roma population (β= 0.143, p= 0.046) but not for the general Hungarian population (β= -0.329, p <0.001). Nonetheless, the interaction between this particular SNP and depression was not positively related to VTE risk (β= -0.160, p= 0.046) for the Roma or general Hungarian population (β= -0.119, p= 0.11). The interaction between rs6025 (*F5, Leiden*) and smoking (β= 0.172, p= 0.008) as well as Leiden and LDL-C (β= 0.368, p= 0.001) increased the

risk of VTE for the general population only but not for the Roma population (β= -0.014, p= 0.86 and β= -0.150, p= 0.55, respectively).

Our study also identifies the higher risk of VTE as a result of a multiplicative interaction between rs8176719 (*ABO*) and cancer, and the risk was higher for the Roma population (β=0.370, p<0.001) than for the general Hungarian population (β=-0.042, p=0.6). Nevertheless, the interaction of rs8176719 (*ABO*) with CAD, (β= 0.197, p= 0.009), migraine (β= 0.287, p= 0.001), and depression (β= 0.342, p <0.001) significantly increased VTE risk only for the general population. The risk of VTE was higher for general Hungarian subjects (β= 0.194, p <0.01) who had diabetes mellitus and non O blood type but not for the Roma subjects (β=-0.039, p=0.63) (Figure 3).



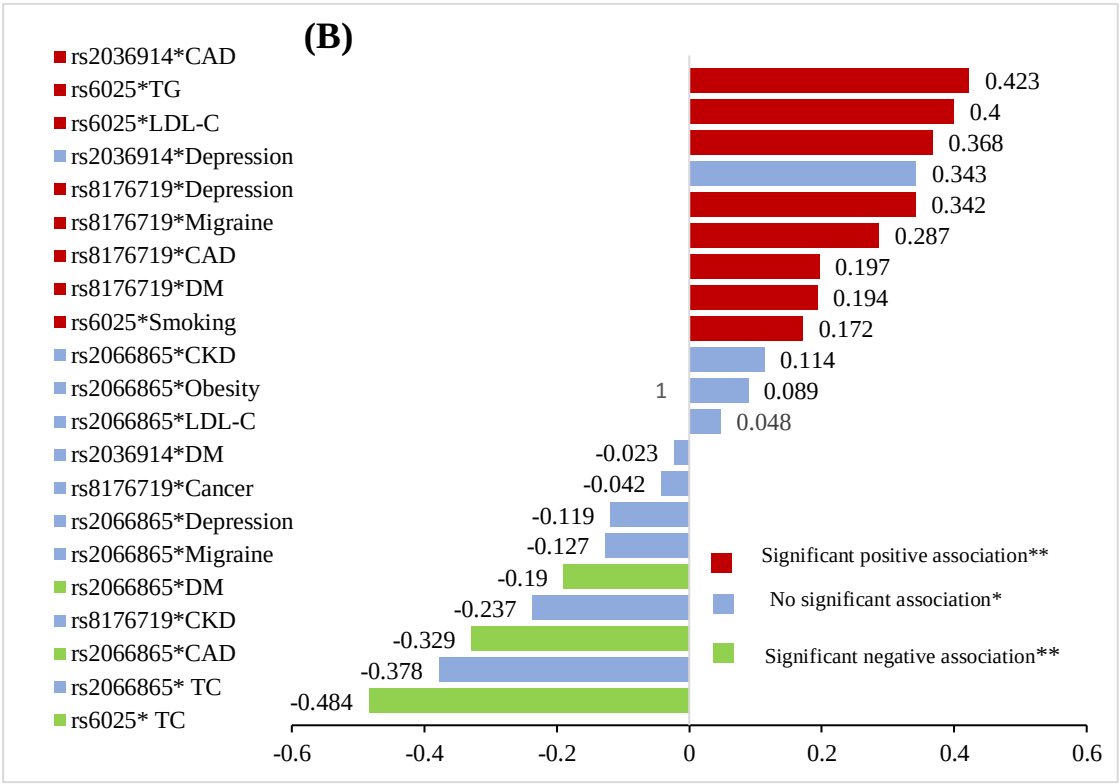


Figure 3: Comparison of G×E on the VTE risk among Roma population (A) and general Hungarian (B) based on standardized linear regression coefficients from multivariate linear regression analysis after interaction terms included between genes and VTE risk factors. **p-value <0.05, *p-value>0.05, LDL-C = Low-density lipoprotein cholesterol, TC=Total cholesterol, TG= Triglyceride, CAD=Coronary Artery Disease, DM=Diabetes Mellitus, CKD= Chronic kidney disease

4.4.2 Case-control study (study 2)

4.4.2.1 Associations between SNPs and VT risk in the study population using genetic association models

The strengths of association regarding VT risk using complete genetic association models [multiplicative, additive, dominant, recessive, and genotypic model] were estimated. Only the Leiden mutation (rs6025) and *F11* (rs2036914) remained significant after multiple testing adjustments ($p<0.01$). In particular, the Leiden mutation variant strongly influenced the VT disease risk in the Hungarian population ($p<0.001$): the odds of VT disease risk among VT cases due to the Leiden mutation ranged from 3.25 (heterozygous genotypic for risk variant (OR=3.25, 95% CI: 2.22, 4.76) to 19.67 (OR=19.67, 95% CI: 2.57, 150.4) in the recessive model/homozygous for risk variant.

The rs8176719 (*ABO*) remained statistically significant only in the multiplicative (OR=1.33, 95% CI: 1.07, 1.64**)** and genotypic models (OR=1.77, 95% CI: 1.14, 2.73); nevertheless, it lost its significance in other models after multiple testing adjustments. Similarly, the protective variant of rs8176719 (*ABO*) remained statistically significant only in multiplicative (OR=0.75, 95% CI: 0.61, 0.93**).** Additionally, the *FGG* (rs2066865) expressed a significant association with VT risk in the multiplicative, additive, and dominant models before multiple testing; however, it lost its significance after adjustment was performed. Conversely, the *F2* (rs1799963) did not show any statistically significant association with VT directly with any of the used models (Table 6).

Table 6: Genetic association test results in the VT cases and control groups of the study population: implication to determine the heritable venous thrombosis disease risk factors in the Hungarian population.

Model	**Gene**	**F5**	**FGG**	**F11**	**ABO**	**F2**
	SNP	***rs6025***	***rs2066865***	***rs2036914***	***rs8176719***	***rs1799963***
Multiplicative	p	**<0.001**	0.026	**0.004**	**0.001**	0.221
	X^2	57.21	4.94	8.47	6.66	1.50
	OR (95% CI)	3.52 (2.50,4.95)	1.32 (1.03,1.68)	1.38 (1.11,1.71)	**1.33 (1.07,1.64)**	1.53 (0.77,3.02)
Additive	p	**<0.001**	0.02302	**0.003**	0.011	0.216
	X^2	54.35	5.17	8.59	6.50	1.53
	OR (95% CI)	3.52 (2.50,4.95)	1.32 (1.03,1.68)	1.38 (1.11,1.71)	1.33 (1.07,1.64)	1.53 (0.77,3.02)
Dominant	p	**<0.001**	0.038	**0.017**	0.03	0.215
	X^2	49.61	4.32	5.71	4.88	1.53
	OR (95% CI)	3.67 (2.52,5.33)	1.38 (1.02,1.86)	1.64 (1.09;,2.47)	1.54 (1.04,2.26)	1.543
Recessive	p	**<0.001**	0.147	**0.017**	0.047	NA
	X^2	16.07	2.1	5.72	3.96	
	OR (95% CI)	19.67 (2.57,150.4)	1.61 (0.84,3.08)	1.47 (1.07,2.03)	1.39 (1.00,1.91)	
Genotypic	p	**<0.001**	0.075	**0.013**	**0.01**	0.215
	X^2	54.35	5.18	8.64	6.60	1.534
	OR (95% CI)	3.25 (2.22,4.76)*	1.81 (0.94,3.51)	1.96 (1.24,3.08)	1.77 (1.14,2.73)	1.54 (0.77,3.07)

**CT (heterozygous for a risk variant), NA= the value of one cell is 0 i.e. <5 hence, the X^2 test is not applicable*

4.5 Comparison of genetic risk scores and risk of VT in study populations

4.5.1 Comparative cross-sectional study

In this study, wGRS was computed only using 5 SNPs (Table 7) which showed a strong association with VTE from previously conducted studies (63,130). Due to the absence of a published external weight (144) for rs121909567 (ATBp3 mutation), the wGRS computation did not consider this particular SNP. The wGRS ranges from 0.0 to 4.7 and 0.0 to 4.6 for the general Hungarian and Roma populations, respectively. The mean wGRS was 1.8 ± 0.8; 95% CI [1.7, 1.9] for the general Hungarian and 1.9 ± 0.76; 95% CI [1.8, 1.9] for the Roma population (Figure 4). The unweighted GRS was calculated for 6 SNPs, and ranges from 0.0 to 7.0 for both study populations, with a mean of 2.7 ± 1.2; 95% CI [2.6, 2.8] for the general Hungarian and 2.8 ± 1.2; 95% CI [2.7, 2.9] for the Roma population (Figure 5).

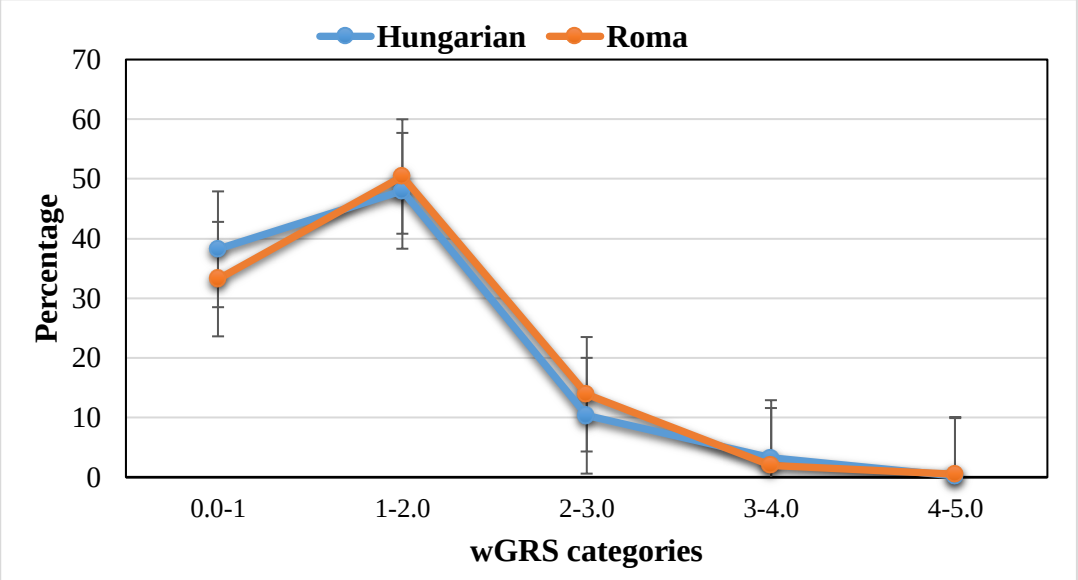


Figure 4: Distribution of wGRS in the Hungarian general and Roma population (The error bars indicate the standard error of the mean)

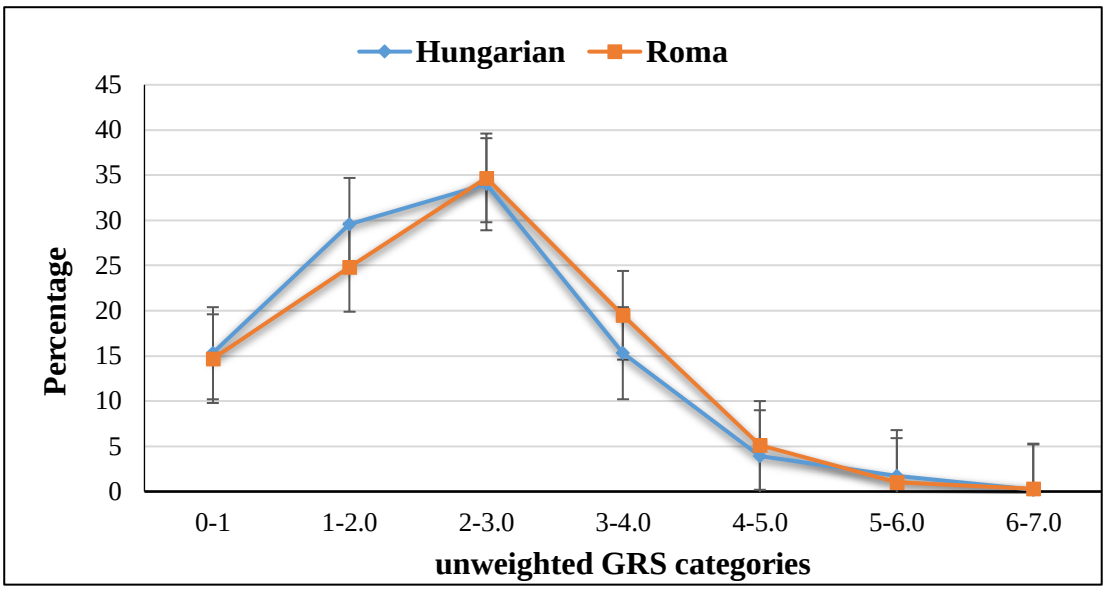


Figure 5: Distribution of unweighted genetics risk scores in the Hungarian general and Roma population (the error bars indicate the standard error of the mean)

Although, the risk of VTE was not significant based on the joint effect of a prothrombotic risk allele, it was higher only for Roma subjects who had ≥3 wGRS (OR=4.74; 95% CI [0.45, 50.3]) compared to individuals with 0-1 wGRS value, but not for the general Hungarian (OR=3.1×10^{-8} $p > 0.05$). Additionally, the risk of VTE was 6.6 times higher in Roma population who had ≥3 risk allele compared to individuals with 0-1 risk allele, and the risk was much higher for the Roma population (OR=6.6; $p > 0.05$) than the Hungarian general population (OR=1.5; $p > 0.05$).

Table 7: SNPs and related genes used in the weighted GRS computation, and weighting numbers with the original publication they were adopted from.

Gene	SNPs	Risk Allele	OR	Weighting Number (lnOR)	Reference
F2	rs1799963	A	2.78	1.02	
F11	rs2036914	C	1.32	0.28	
FGG	rs2066865	A	1.56	0.44	
F5	rs6025	T	3.79	1.33	
ABO	rs8176719	G	1.85	0.62	(63)

4.5.2 Case-control study (study 2)

The unGRS and wGRS for 5 SNPs were computed for 298 VT patients and 400 healthy controls. The unGRS ranged from 0 to 6 (3.46±1.31) and 0 to 7(2.77±1.28) for VT patients and healthy controls, respectively **(Figure 6A).** The wGRS ranged from 0 to 4.6 (1.93±0.97) for VT patients and 0 to 4.7 (1.37±0.78) for healthy controls **(Figure 6B).**

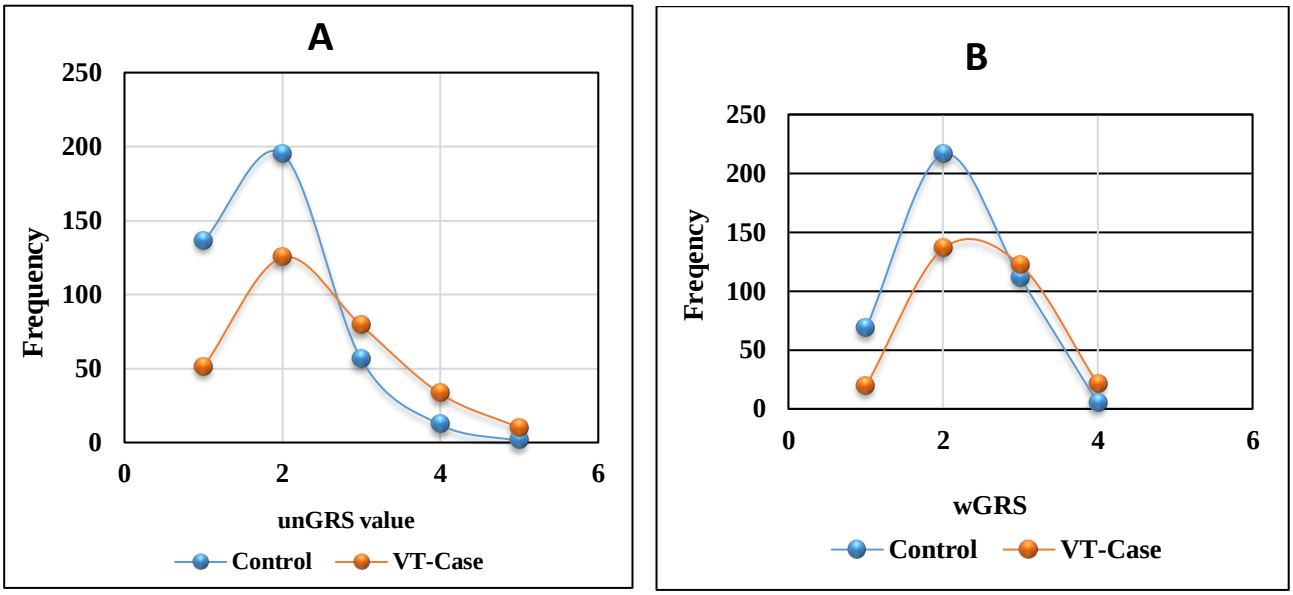


Figure 6: Unweighted (A) and weighted (B) GRS distribution comparison among VT cases and the control group of the Hungarian population

4.5.2.1 Associations of GRS with VT risk

The distributions of other covariates including GRS were significantly distinct ($p<0.001$) between study groups **(Table 8).** The test revealed significant differences (cases vs. control) in age (median=65; 44, $p<0.001$), BMI (median=28.72; 26.75, $p<0.001$), unGRS (median=3; 3, $p<0.001$), and wGRS (median=1.79; 1.34, $p<0.001$). Although the median values of unGRS for the VT patients and healthy controls were the same, a higher value of unGRS (3) was more frequent in the cases than in the healthy controls.

Table 8: Mann–Whitney U test of non-normally distributed covariate variables of the study population.

Variables	Median		p value
	Cases (n=298)	Controls (n=400)	
Age	65	44	<0.001
BMI	28.72	26.75	<0.001
unGRS[¶]	3	3	<0.001
wGRS	1.79	1.34	<0.001

[¶]Skewed data with high proportion of the same median value (3) for both groups; however, the distribution is still higher for cases (35.5%) than control group (29.3%).

Table 9 shows the results of multivariate logistic regression analysis of covariate variables adjusted for sex and age. Among the well-known non-heritable risk factors for VT, age and obesity were found to be significantly associated with VT risk in the study population, with a higher risk observed in the case group compared to the control group (Table 9). The risk of VT among elderly subjects aged ≥60 years old was 12.8 times higher than that of subjects aged below 60 years (AOR=12.83, 95%CI: 8.38, 19.63). Likewise, the VT risk was 2.3 times higher for obese subjects (BMI>30 kg/m^2) than for normal-weight subjects (AOR=2.28, 95%CI: 1.51**,** 3.42). Furthermore, the wGRS remained statistically significant after adjusting for both sex and age [AOR=2.69, 95%CI: 1.74, 4.19 and AOR=2.24, 95% CI: 1.51, 3.32], respectively. However, the unGRS lost its statistical significance (AOR=0.88, 95% CI: 0.65, 1.18) in the multivariate regression analyses model (Table 9).

Table 9: Association of covariates with VT risk in the Hungarian population (study 2).

Variables	VT risk[a]			VT risk[b]	
	ß	p value	COR (95% CI)	AOR (95% CI)[£]	AOR (95% CI)[¶]
Sex (Male)***	0.271	0.077	1.31 (0.97,1.77)	_____	1.16 (0.84,1.61)
Age≥60 yrs	2.468	<0.001	11.79 (7.96,17.49)	**12.83 (8.38,19.63)****	_____
BMI (≥30 kg/m^2)	0.850	<0.001	2.34 (1.59,3.45)	1.41 (0.88,2.26)	**2.28 (1.51,3.42)****
unGRS	0.412	<0.001	1.51 (1.34,1.71)	0.88 (0.65,1.18)	0.94 (0.72,1.21)
wGRS	**0.729**	**<0.001**	**2.07 (1.72,2.50)**	**2.69 (1.74,4.19)****	**2.24 (1.51,3.32)****

****** Female is a reference, a=crude odds ratio (COR), b=adjusted odds ratio (AOR), [£]adjusted for sex, [¶]adjusted for age, **p value<0.0001.***

4.5.2.2 Venous thrombosis risk prediction in a study population

For a case-control study, we calculated the receiver operating characteristic curve (ROC) to assess how well the score classified venous thrombosis in patients and control subjects. The AUC value of the included SNPs ranges from 0.51 [95%CI: 0.47, 0.55, p-value 0.64] for rs1799963 in *F2* up to 0.62 [95%CI: 0.57, 0.66, p value<0.001] for rs6025 in F5. The discriminative accuracy of the model improved with the addition of each SNP (Figure 7**Figure**). We started addition with the Leiden mutation (rs6025), which had the highest effect size, and ended with rs2036914 (F11), which had the lowest effect size among the included 5 SNPs. The addition of each SNP increases the AUC value after *F2* (rs1799963).

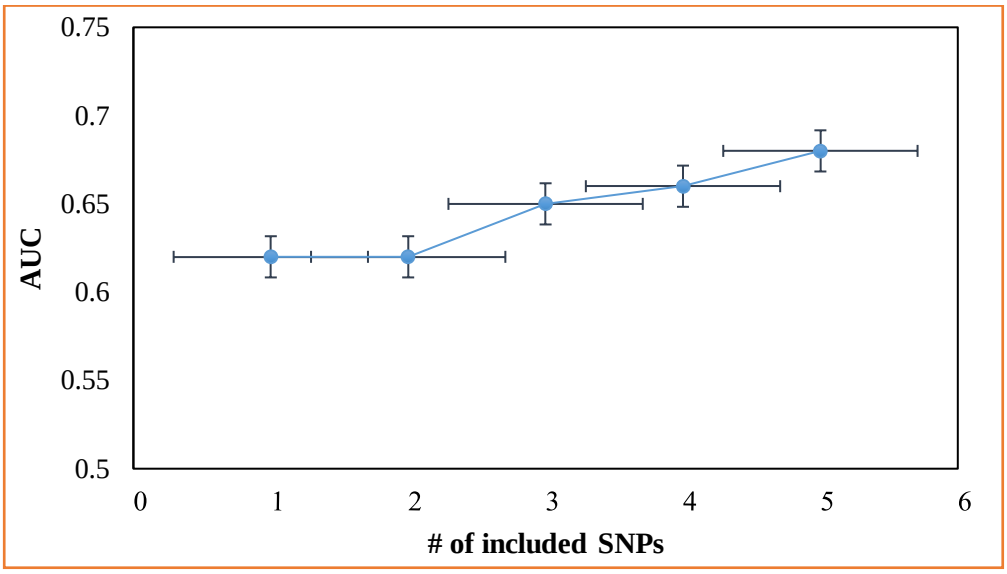


Figure 7: The change in the discriminatory accuracy of the AUC of the genetic risk score after the addition of each SNP into a model.

The AUC for the weighted 5-SNP risk score had an AUC of 0.68 [95% CI: 0.64; 0.72], i.e., there is a 68% probability that a randomly selected patient will have a higher score than a randomly chosen control subject. The wGRS was a better predictor than the unGRS [AUC= 0.65, 95% CI: 0.60; 0.69].

The proportions of variability explained by the Leiden mutation (rs6025) was higher than that of the 5-SNP risk score (8% vs. 7%). Furthermore, approximately 39% of the observed variability is related to the combination of genetic and non-heritable risk factors, which is higher than those factors independently (Table 10).

Table 10: Venous thrombosis risk predictability of the Leiden mutation, genetic risk, non-heritable risk, and combined model in the Hungarian population.

Variables	**r^2**	**N=698**	
		AUC (95% CI)	**p value**
Leiden mutation risk model[*]	0.08	0.62 (0.57,0.66)	<0.0001
Genetic risk model[¶]	0.09	0.68 (0.64, 0.72)	<0.0001
Non-heritable risk model[$]	0.31	0.85 (0.82, 0.88)	<0.0001
Combined model[§]	0.39	0.89 (0.86,0.91)	<0.0001
Difference[**]	-	0.039 (0.02, 0.06)	<0.0001

Note: CI: Confidence interval; AUC: Area under curve, r^2: variability explained by each variable
[*]Leiden mutation; the most prevalent heritable VT risk variant in the study population
[¶]Genetic risk model: weighted GRS computed from 5 SNPs (rs6025, rs2066865, rs2036914, rs8176719, and rs1799963)
$Non-heritable risk model: age (5-year interval, sex, and BMI (<25kg/m^2, 25-30kg/m^2, and ≥30kg/m^2)
[§]Combined risk model: genetic risk model plus non-heritable risk model
[**]difference between the combined and non-heritable risk models

There was a difference between the discriminative accuracy of the 5-SNP risk score in men [AUC=0.68, 95% CI: 0.62; 0.74, p-value <0.001] and women [AUC=0.61: 95% CI: 0.55; 0.67, p-value <0.001]. Moreover, the AUC of wGRS of 5-SNPs was significantly higher for men [AUC=0.71, 95% CI: 0.65; 0.76, p-value <0.001] than women [AUC=0.63, 95% CI: 0.57; 0.69, p-value<0.001].

4.5.2.2.1 Risk prediction based on a combination of genetic and non-heritable risk factors

We also evaluated the discriminative accuracy of well-known non-heritable VT risk factors such as age, sex, and obesity to explore their independent and combined VT risk predictability in the study subjects. The independent AUCs of age and obesity were 0.84 (p value<0.0001) and 0.59 (p value<0.001), respectively. The combination of all well-known VT risk factors changes the discriminative accuracy of the AUC in to 0.85, p value<0.0001. Likewise, when we added the non-heritable risk factors into the genetic risk factors, the AUC significantly projected to 0.89 [95% CI:0.86;0.91] compared to genetic (AUC=0.68) or non-heritable risk factor predictability (AUC=0.85; p<0.0001) **(**Figure 8**)**. The AUC difference of the combined and non-heritable risk factors was significant (AUC=0.039, 95% CI: 0.02; 0.059, p value<0.0001). There was no significant difference in the AUC value between men and women in the non-heritable [men: AUC=0.81, 95% CI: 0.76; 0.86 vs. women: AUC=0.82, 95% CI: 0.78; 0.87] and combined risk score [men: AUC=0.87, 95% CI: 0.83; 0.91 vs. women: AUC=0.86, 95% CI: 0.82; 0.90] models.

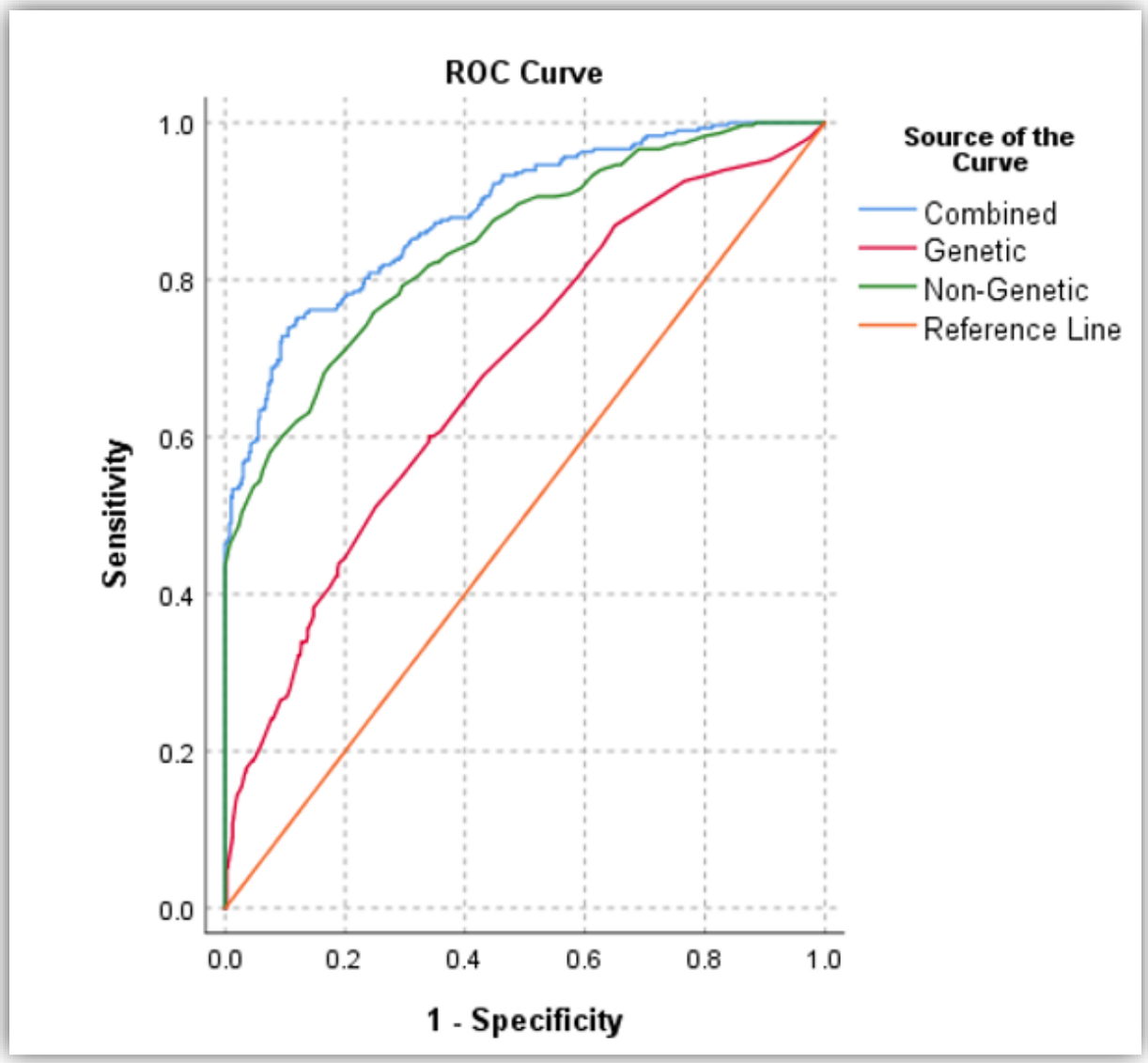


Figure 8: Illustration of the ROC curves value comparison for the independent VT risk factors [genetics (5 prothrombotic SNPs) and non-heritable s (age, sex, BMI)] and their combination to verify their VT risk predictability in the Hungarian population.

5 Discussion

Overview of the studies

In our present study, a total of 1499 subjects have participated in a comparative cross-sectional (801) and a case-control study (698) to investigate the genetic and environmental VT risk in the Hungarian population. In a comparative study, 406 general Hungarian and 395 Roma indviduals have participated. We assessed the gene-environmental interaction and VTE risk in the general Hungarian and Roma populations using six prothrombotic SNP: rs6025 (*F5 Leiden*), rs2066865 (*FGG*), rs2036914 (*F11*), rs8176719 (*ABO*), rs1799963 (*F2*), and rs121909567 (*SERPINC1*) and proven environmental VTE risk factors such as smoking, DM, CAD, CKD, obesity, depression, and high level of LDL-C. The former research highlighted Roma population is highly vulnerable to CVD due to the prevalent of both heritable and non-heritable VT risk factors. Consequently, we investigate and compare GxE and VTE risk in the general Hungarian and Roma population to establish the VTE risk in the Hungarian population. Our study is the first to explore and compare gene-environmental interaction (GxE) and VTE risk in this population. The present study revealed evidence that the Roma population was at higher risk of VTE due to dual exposure and the possibility of the synergistic interaction of heritable and non-heritable VTE risk factors.

Multiplicative interaction and VTE risk

Our study explored the interaction between the genetic and environmental risk factors, which are divergent in both populations. A significant multiplicative interaction was observed between the rs8176719 (*ABO*) gene and diabetes mellitus, migraine, depression, and CAD for the general Hungarian population but only with cancer for the Roma population. The regression coefficient of multiplicative interactions of DM and rs8176719 (*ABO*) was synergistic and significant. The finding demonstrated that non-O blood group diabetes subjects in the general Hungarian were more likely to have higher VTE risk than O blood group subjects. Our study is consistent with a systematic review and meta-analysis of cohort study results revealed the risk of VTE is 1.4 times higher among diabetes mellitus individuals compared to non-diabetic individuals (145). However, another systematic review of the case-control studies results shows the absence of significant differences in VTE risk among diabetes mellitus and non-diabetes mellitus subjects (146). This inconsistency might be attributable to the study design and sample size they considered for analysis.

We verified that subjects experiencing migraine and non-O blood types were at higher risk of developing VTE. This finding corresponds to studies results which designated that the presence of migraine increases VTE risk by 1.3 (110),1.5 (109,111), and 2.5 (147) fold. The present study also revealed the risk of VTE is higher for depressive Roma subjects who carry a risk variant of coagulation factors eleven (rs2036914) and non-O-blood group (rs8176719) carrier general Hungarian subjects. The result is compliant with a former investigation which indicates depression raises VTE risks (115,116). Likewise, the risk of developing VTE was 6 to 7 times higher in cancer patients (29,148,149). This result is along with our study findings that reveal the VTE risk is higher for the Roma subjects who had cancer.

The risk of VTE is higher among Roma subjects who had cancer and non-O blood group. Our present work findings also support former related studies' results which demonstrated the existence of cancer increases the likelihood of VTE (106,150). Further studies on cancer and prothrombotic genotypes highlight VTE risk increased by 11-12 fold due to the concurrent presence of cancer and rs8176719 (ABO) risk variant (151), (152). The authors also found that 39% (151) and 30% (152) of VTE risk was attributed to the joint presence of the non-O-blood group and cancer.

Though the association between the joint presence of Leiden mutation and cancer in our present study is not significant; our finding was in accordance with an earlier study (148) which found the existence of cancer and Leiden mutation (rs6025) variant increases the VTE risk by 2-fold among dual exposed subjects. The absence of significance might be due to the small number of individuals with cancer and VTE, the study design, and the relatively small sample size in our cases.

Furthermore, our study revealed a synergistic and statistically significant association between cigarette smoking and the Leiden mutation (rs6025) variant carrier which increases the risk of VTE in those subjects. The risk of VTE is higher among general Hungarian subjects who carry the risk variant for Leiden mutation (rs6025) and smokers. Prior studies also showed that the combined effect of rs6025 (*F5*) and smoking increased the risk of VTE (153,154). A large population-based case-control study also supported our finding, where the joint effect of rs6025 (*F5*) and current smoking resulted in a 5 fold increased VTE risk (155). Another cohort study revealed that the concurrent existence of smoking other than rs6025 (*F5*) increased the VTE risk by 51% and 10% at 10 years for homozygous and heterozygous risk variants, respectively (156). Crous-Bou et al. also found an additive interaction between prothrombotic SNPs and smoking that increased VTE risk (136).

Likewise, individuals with coronary artery diseases and rs2036914 (*F11*) or rs8176719 (*ABO*) were at higher risk of VTE than individuals without CAD and those variants. The existence of an interaction between rs2036914 (*F11*) and CAD increased the VTE risk among general Hungarian and Roma populations who carry the risk variant for coagulation factors eleven (rs2036914) and rs8176719 (*ABO*). Our finding was also consistent with a study that found that myocardial infarction patients with ≥1 risk allele at rs2036914 (*F11*) had a 1.8-fold higher risk of pulmonary embolism than non-carriers (157). Besides, the risk of VTE was 1.5-fold higher among individuals with non-O blood type and myocardial infarction (157).

Additive interaction and VTE risk

Our study also explores an additive interaction between a high level of LDL-C, migraine, current cigarette smoking, and ≥3 wGRS values. The risk of VTE increased by 3.2-fold for the Roma subjects with a high level of LDL-C and ≥3 risk alleles. This result was consistent with a GWAS (64), which revealed one standard deviation (SD) of elevated LDL-C was associated with an increased risk of VTE. The finding of current cigarette smoking was in agreement with the study of Crous-Bou et al., in which the relationships between current smoking and VTE genetic factors were additive (136). In our study, individuals who had experience migraine in addition to a wGRS value ≥3 had a 3.9 times higher risk for VTE than individuals with either a wGRS ≥3 or migraine. This finding was in harmony with a study by Peng et al., which revealed that migraine headaches increased the risk of VTE (147).

The SERPINC1 and prothrombin mutation synergistic interaction, and VTE risk

Even if the multiplicative interaction between the SERPINC1 (rs121909567), prothrombin mutation / coagulation factors two (rs1799963) (F2), and VTE risk factors were not statistically significant, their regression coefficient designated the likelihood of higher VTE risk among individuals which had dual exposures to those variants and comorbid diseases such as CAD, CKD, cancer, DM, depression, migraine, and obesity. Interestingly, rs121909567 (*SERPINC1*) (ATBp3 mutation) multiplicatively interacted with CAD, CKD, cancer, DM, depression, migraine, and obesity. Even though their relationships were not significant, the trend of interaction showed the probability of VTE risk increment among the Roma population. The lack of statistical significance between GxE and VTE risk for rs121909567 (*SERPINC1*) and rs1799963 (*F2*) might be due to the very small number of VTE cases. Our finding is in line with the formerly conducted studies which showed the prevalence of cardiovascular risk factors was

higher in the Roma population (126,127), and we found that in addition to environmental factors, genetic susceptibility contributes to high cardiovascular mortality/morbidity (158).

In a relatively isolated population such as the Roma, the consanguinity rate is high (159); consequently, it was assumed that a founder mutation had an impact on the development of thrombotic diseases (66). A recent study also identified a high prevalence (2.74%) of ATBp3 mutation in the Roma population; however, no ATBp3 mutation was found in the general Hungarian population (133). This finding was in-line with our study results, which revealed that ATBp3 mutation s were found only in the Roma population but not in the general Hungarian population.

Altogether, our comparative study among general Hungarian and Roma populations demonstrated the simultaneous presence of comorbidity such as CAD, DM, cancer, migraine, depression and cigarette smoking and genetic risk factors increased the VTE risk among individuals with either heterozygous or homozygous variants of rs6025 (F5, Leiden), rs2066865 (*FGG*), rs2036914 (*F11*), and rs8176719 (*ABO*). For general Hungarian population the rs8176719 (*ABO*), rs2036914 (*F11*), and rs6025 (*F5, Leiden*), played a great role in increasing the risk of VTE among those subjects who had CAD, DM, migraine, depression, and cigarette smoking; whereas, for the Roma population, rs121909567 (*SERPINC1*), rs1799963 (*F2*), rs2066865 (*FGG*), rs2036914 (*F11*), and rs8176719 (*ABO*) interact with VTE risk factors which increased the burden of VTE for individuals who had genetics as well as environmental risk factors.

Furthermore, our study confirmed that the rs121909567 (*SERPINC1*, ATBp3) is a founder mutation among the Roma population but not for the general Hungarian population. Although the findings were subjected to selection and observation biases due to the small number of cases and the study design, our study reveals some clues about the burden of the joint presence of genetic and environmental risk factors on VTE risk. Due to higher genetic load and gene-environmental interactions, this minority Roma population is at higher risk of VTE than the general Hungarian population.

Thus, our results suggest that an intensive search for the rs121909567 (*SERPINC1*; ATBp3) founder mutation might be an important factor for the assessment of thrombotic disease susceptibility among the Roma population. In addition, we strongly recommend further studies among a large number of VTE cases to explore the more precise impact of genetic and environmental risk factors on VTE in the study populations.

Results of strongly associated SNPs and VT risk determination

Based on the findings of our first study, we further conduct a case-control study using 298 clinically confirmed VT cases and 400 healthy control to explore the genetic background of VT risk in the Hungarian population and to determine the combined VT risk predictability of strongly associated SNPs: rs6025 (*F5 Leiden*), rs2066865 (*FGG*), rs2036914 (*F11*), rs8176719 (*ABO*), and rs1799963 (*F2*) and well-known conventional VT risk factors (Age, Sex, and Obesity) in the population.

Though numerous SNPs provoke VT events in genetically vulnerable individuals, the contributions of the five strongly associated SNPs to VT risk are enormously high (63,67,68). Besides, previously conducted studies demonstrated the importance of those five prothrombotic SNPs in the reoccurrence of heritable VT diseases (63,67,68). Consequently, we aimed to explore the combined genetic risk predictability of strongly associated VT SNPs and well-known non-heritable VT risk factors in the Hungarian population. In the present study, only three SNPs: rs6025 (*F5*), rs2036914 (*F11*), and rs8176719 (*ABO*) remained statistically significant after adjustment for multiple testing ($p<0.01$). The highest VT risk was detected among Leiden mutation carriers/rs6025 (OR=3.52, 95% CI: 2.50; 4.95). Its allele frequency was approximately 3-fold higher in VT cases (20%) than in control (6.8%) groups. Our findings are consistent with previously conducted studies that indicated the odds of VT risk are 3.5 (62) and 4.38 times higher for rs6025(F5) variant carriers than for noncarriers (160).

Moreover, several studies suggest that the Leiden mutation is vastly prevalent in Caucasian ancestry, particularly in European descent. It is one of the most influential heritable VT risk factors that increase the burden of VT in genetically vulnerable individuals (69–71). Our finding is consistent with these findings where F5 is highly prevalent among cases (20%) in the present study. Furthermore, it was strongly associated with the trait in all genetic association models. This highlights the fact that the Leiden mutation is an independent predictor of venous thrombosis risk (62,70,161), and its contribution to the burden of VT disease is remarkable (161), particularly for genetically susceptible individuals and Caucasian ancestry (69).

Likewise, the risk allele frequency and VT risk for rs2036914 (*F11*) and rs8176719 (*ABO*) were higher for cases even after adjustment for multiple testing. A previously conducted study revealed that *F11* (rs2036914) is an independent predictor of VT (162), which is compatible with our finding where the VT risk was 1.38 times higher for cases that of controls. Different

studies have shown that VT risk distribution due to F11 (rs2036914) is similar in Caucasians and African Americans (76,77). The allele frequency of the ABO SNP (rs8176719) was more prevalent in the control group (47.4% vs. 40.4%). Additionally, the risk of VT was lower (OR= 0.75, 95% CI: 0.61; 0.93) among subjects with the rs8176719 variant. Furthermore, the VT risk was 1.33 times higher for risk variant carriers. Several studies have indicated that individuals with the O-blood type are at the lowest risk of VT compared with individuals without the O-blood type (78,79).

On the other hand, numerous studies have shown that non-O blood type subjects (A, AB, and B) were at higher risk of developing venous thrombosis compared to O-blood type subjects (80–83). Our study findings are also consistent with those of previously conducted studies. Fang *et al.* reported that the risk of VTE is higher for the African American race and non-O blood type individuals than the Caucasian race and O blood type individuals (84).

Although the risk allele frequencies and VT disease risk were not distinct in the case of prothrombin mutation (rs1799963), it was the second most prevalent risk variant in the Hungarian population. The lack of statistical significance despite the large odds ratio (OR) might attributed to the limited number of VT cases in our present study compared to previously conducted studies. Our study finding is consistent with previous studies' findings, which revealed that rs1799963 is more prevalent in European ancestry than other ancestries. (American, African American, Asian, and African) (57). Several studies have shown that approximately 2-4-fold venous thrombosis risk was attributed to a hypercoagulability state that resulted from a mutation in the prothrombin gene/rs1799963 (70,163,164).

Some studies have indicated that pooled variants have more impact on VT risk determination than a single variant (63,94); consequently, we computed the weighted and unweighted GRS to determine the VT risk in the study. Our findings showed that the wGRS is an independent predictor of VT risk in the study population, and 2.37 times higher among VT cases than controls. This finding is also supported by studies conducted by Kujovich *et al*. (2011) and Skille *et al*. (2020)(63,94).

Non-heritable VT risk factors and their impact

The impact of non-heritable risk factors on VT risk is also considerable. Our study showed that VT was more prevalent among elderly (≥ 60 years) subjects than their counterparts (58.1% vs. 10.5%; p<0.0001). Likewise, the odds of VT risk for elderly subjects were 12-fold higher than those of aged <60 years. The risk of VT increases with age due to different factors, such as anatomical (95), pathophysiological (95–98), and hormonal derangement (165). Consequently, it hastens and increases the vulnerability of elderly individuals to VT risk and other CVD (53,96–98). Furthermore, our findings show that the risk of VT is 2.28 times higher among obese subjects than normal-weight subjects. This finding is similar to the results of previous conducted studies elsewhere where obesity shows an independent predictor of VT risk (99–103).

VT risk predictability: individual SNP, non-heritable, and their combination in the Hungarian population

The ability to predict the risk of a certain event before its occurrence is important in clinical epidemiology. Precise risk prediction helps to control an event at as early stage as possible (166–169) and offers to use the available resources effectively and efficiently (166–169). We used ROC curves to establish individual and combined VT risk predictability of the SNPs and nonheritable VT risk factors to develop a risk stratification tool.

In our study, the highest AUC result was obtained for the Leiden mutation (AUC=0.62), whereas the lowest AUC was obtained for prothrombin mutation /*F2* (0.52). Generally, the AUC value increases as more SNPs are added to the model in addition to *F5*. Our finding is also consistent with studies done elsewhere, where the addition of more SNPs into the model showed an increase in the AUC value to a certain extent (63,170).

We also found that the wGRS is a better predictor of VT risk than individual SNPs (0.68 vs. 0.62), and their combination with non-heritable risk factors yields a larger AUC value with higher discriminatory accuracy (AUC=0.89). This finding is consistent with previously conducted studies, which reveal the combination of clinical and genetic risk factors increases the risk of VT 8 times more than the genetic or clinical model independently (171,172). Overall, the Leiden mutation, *F11*, and *ABO* risk alleles are highly prevalent and strongly determine venous thrombosis disease risk in the Hungarian population. The pooled genetic risk variants are more influential than a single variant alone.

The combined model is the best predictor of VT risk, so the stratification of highly vulnerable individuals based on their genetic profiling and well-known nonmodifiable VT risk factors is important for the effective and efficient utilization of preventive and control measures of VT risk.

6 Strengths and limitations of our study

In our comparative study, we tried to verify the GxE and VTE risk in the highly vulnerable Roma and the general Hungarian populations; in this study, we explore the fact that dual exposure (genetic and other comorbidities) increases the likelihood of VTE in a given population like the Roma population. It provides insight into the stratification of the subjects based on their risk burden for VTE controlling measures. This study also confirmed that the origin of the SERPNC1 mutation is the Roma population. Our present study is in line with the findings of similar studies reported by other scholars.

The main limitation of this study was the smaller sample size, especially the VTE cases, and the data collected were self-reported by respondents (not confirmed clinically); hence, it might expose to social desirability and recall bias. Therefore, we further suggest research with adequate sample size and clinically confirmed cases with better design to have strong evidence of the VT risk background in the Hungarian population. Consequently, we used a case-control study to distinguish the VT risk predictability of the five strongly associated VT SNPs and well-known non-heritable VT risk factors in the Hungarian population. Our study explore the possibility of efficiently and effectively utilizing the available resource for risk prediction in the given population. However, our study lacked the comparison of formerly identified strongly associated VT SNPs with the novel loci, which are strongly associated with VT risk as well (140,173–175). As a result, we recommend further study that considers the comparison of VT risk predictability in the Hungarian population using the novel and strongly associated VT SNPs with the formerly identified VT SNPs.

7 Conclusions and Recommendations

Generally, both conducted studies provides an insight into the heritable and non-heritable VT risk factors in the Hungarian and Roma populations. The Leiden mutation (rs6025), coagulation factors 11 (rs2036914), and the *ABO* gene (rs8176719) of non-O-blood type determine the heritable VTE risk in the Hungarian population.

However, in the Roma population the SERPINC1 mutation (rs121909567) and *FGG* (rs2066865), coagulation factors 11 (rs2036914), and the *ABO* gene (rs8176719) determine the heritable VT risk in this particular study subject. The coexistence of heritable and co-morbidity such as DM, CAD, and smoking affects the Hungarian population at large. However, cancer, CAD, obesity, and LDL-C affect the Roma population at large.

In conclusion, the Hungarian population is at higher risk of VT due to aging and the presence of heritable VT risks such as Leiden mutation, *F11*, and *ABO*. Particularly, the Leiden mutation influences VT in the Hungarian population with the largest risk variants frequency. The Leiden mutation is the most prevalent genetic VT risk in the Hungarian population. The pooled variant predicts the VT risk in the Hungarian population. Moreover, the combination of non-modifiable VT risk and the heritable risk increases the VT risk predictability accuracy of the model. The difference between the heritable and non-heritable VT risk factors is statistically significant.

The stratification of highly vulnerable individuals based on their genetic profiling and well-known non-modifiable VT risk factors is important for the effective and efficient utilization of preventive and control measures for VT risk. Furthermore, we recommend further study that considers the novel and strongly associated VT SNPs and the formerly identified SNPs and their comparison on the VT risk predictability in the Hungarian population.

8 Summary

Background: Venous thrombosis (VT) is one of the three principal causes of cardiovascular disease (CVD) related mortality with a significant genetic predisposition. In Europe, although overall CVD-related morbidity is decreasing, mortality is still high. Hungary shares the largest proportion of this mortality, and CVD remains the prominent cause of death in Hungary. The coexistence of heritable and non-heritable risk factors increases the burden of VT in dual-exposed individuals. Formerly conducted studies revealed that the Roma population is at higher risk of CVD due to heritable and non-heritable VT risk factors. Thus, we aimed to explore and compare the gene-environmental interactions and VTE risk in general Hungarian and Roma subjects. We further aimed to investigate the combined VT risk predictability of the five strongly associated SNPs and well-known conventional VT risk factors in the Hungarian population.

Methods: A comparative cross-sectional study design was employed among 406 general Hungarian and 395 Roma subjects. Moreover, a case-control study was conducted among 298 clinically confirmed VT cases and 400 health controls to investigate the heritable background of VT in the Hungarian population. Except for the 298 subjects data; which is collected by the DCLS, Department of Laboratory Medicine, Faculty of Medicine, University of Debrecen, all other included data were extracted from the comprehensive database. Plink 1.9 version and IMB SPSS version 26 were used for the analysis.

Furthermore, the AUCs for genetic and non-heritable risk factors were estimated to explore their VT risk predictability in the case-control study. An odds ratio (OR) with their respective 95% CI at 0.05 alpha value was used to declare any association between VT risk and its factors.

Results: In both studies, the risk allele frequencies of *F5* (rs6025), *F11* (rs2036914), and *ABO* (rs8176719) are higher in the Hungarian population. However, in the comparative study, the *SERPINC1* (rs121909567) SNP risk allele is only present in the Roma population but not all in the general Hungarian population, which confirmed that the Roma population is the origin of this particular mutation. The coexistence of genetic and non-heritable risk factors increases the likelihood of VT in double-exposed individuals. Due to the multiplicative interaction between the genetic and non-heritable risk factors such as depression and rs2036914 risk variant (β=0.819, p=0.02), high levels of LDL-C and rs2066865 (*FGG*) (β= 0.389, p= 0.002), rs8176719 (*ABO*/non-O blood type and cancer (β=0.370, p<0.001), rs2066865 (*FGG*) and CAD

(β=0.143, *p*= 0.046) the risk of VTE is higher for the Roma subjects than the general Hungarian subjects.

However, as a result of the synergistic interaction between cigarette smoking and Leiden mutation (β= 0.172, p= 0.008), diabetes mellitus and rs8176719 (*ABO*) (β= 0.194, p <0.01), rs8176719 (*ABO*) and CAD, (β= 0.197, p= 0.009), migraine (β= 0.287, p= 0.001), and depression (β= 0.342, p <0.001) the risk of VTE significantly higher only for the general Hungarian population but not for the Roma populations. Likewise, in a case-control study, the risk allele frequencies of *F5, F11*, and *ABO* are higher among the VT cases than in the control group. Specifically, the Leiden mutation risk allele is highly prevalent among VT cases. Its risk allele frequency was 3.52-fold higher in the VT group than in the control group (AOR =3.52, 95% CI: 2.50; 4.95). The combined (genetic and non-heritable) (AUC=0.89, p<0.001) show good discrimination between VT cases and control.

Conclusions: In general, our studies provide insight into VT background (heritable and non-heritable VT risk factors) in the Hungarian population. The Leiden mutation (rs6025), *F11* (rs2036914), and *ABO* (rs8176719) determine the genetic VT risk factors in the Hungarian populations. The presence of other non-heritable VT risk factors such as aging, CAD, DM, cancer, smoking, and obesity, increases the likelihood of VT risk in dual-exposed individuals compared to their counterparts. The combined model predicts the risk of VT risk in the Hungarian population.

Recommendations

Stratification of highly vulnerable individuals based on their genetic profiling and comorbidities are of paramount importance for the efficient and effective utilization of scarcely available VT risk controlling and preventive measures in the Hungarian population. Furthermore, we suggest further study that considers the novel and strongly associated VT SNPs and the formerly identified SNPs and their comparison on the VT risk predictability in the Hungarian population.

9 Reference

1. North American Thrombosis forum. Preventing, Managing, & Thriving with Thrombosis [Internet]. 2018. 730-4120. Available from: https://thrombosis.org/wp-content/uploads/2018/11/Preventing-Managing-Thriving-with-Thrombosis.pdf

2. Jon C. Aster, H. Franklin Bunn. Pathophysiology of Blood Disorders, second Edition. McGraw Hill Professional, Jan 6, 2017

3. ISTH Steering Committee for World Thrombosis Day. Thrombosis : a major contributor to the global disease burden. J Thromb Haemost. 2014;12(July):1580–90.

4. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012;380(9859):2095–128.

5. Thrombosis | Johns Hopkins Medicine. What is thrombosis? In. Available from: https://www.hopkinsmedicine.org/health/conditions-and-diseases/thrombosis

6. https://my.clevelandclinic.org/health/diseases/22242-thrombosis. Thrombosis: Types, Causes, Symptoms & Treatment. In 2021. Available from: https://my.clevelandclinic.org/health/diseases/22242-thrombosis

7. Kakkos SK, Gohel M, Baekgaard N, Bauersachs R, Bellmunt-Montoya S, Black SA, et al. Clinical Practice Guidelines on the Management of Venous Thrombosis. Eur J Vasc Endovasc Surg [Internet]. 2021;61(1):9–82. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1078588420308686

8. Wendelboe AM, Raskob GE. Global Burden of Thrombosis: Epidemiologic Aspects. Circ Res. 2016;118(9):1340–7.

9. Konstantinides S V., Barco S, Lankeit M, Meyer G. Management of Pulmonary Embolism: An Update. J Am Coll Cardiol. 2016;67(8):976–90.

10. Duffett L, Castellucci LA, Forgie MA. Pulmonary embolism: Update on management and controversies. BMJ. 2020;370.

11. Karia S, Screaton N. Pulmonary embolism. ERS Monogr. 2015;2015(9781849840668):60–79.

12. Di Nisio M, van Es N, Büller HR. Deep vein thrombosis and pulmonary embolism. Lancet. 2016;388(10063):3060–73.

13. Rinde LB, Småbrekke B, Mathiesen EB, Løchen ML, Njølstad I, Hald EM, et al. Ischemic Stroke and Risk of Venous Thromboembolism in the General Population: The Tromsø Study. J Am Heart Assoc. 2016;5(11):1–9.

14. Huang JC, Bhaskar SMM. Clot Morphology in Acute Ischemic Stroke Decision Making. Int J Mol Sci. 2022;23(20):1–16.

15. Ntaios G, Hart RG. Embolic Stroke. Circulation. 2017;136(25):2403–5.

16. Palasubramaniam J, Wang X, Peter K. Myocardial Infarction - From Atherosclerosis to Thrombosis: Uncovering New Diagnostic and Therapeutic Approaches. Arterioscler Thromb Vasc Biol. 2019;39(8):E176–85.

17. May JE, Moll S. Unexplained arte rial throm bo sis : approach to diag no sis and treat ment. 2021;76–84.

18. Albricker ACL, Freire CMV, Dos Santos SN, De Alcantara ML, Saleh MH, Cantisano AL, et al. Joint Guideline on Venous Thromboembolism, 2022. Arq Bras Cardiol. 2022;118(4):797–857.

19. Kruger PC, Eikelboom JW, Douketis JD, Hankey GJ. Deep vein thrombosis: update on diagnosis and management. Med J Aust. 2019;210(11):516–24.

20. Mazzolai L, Aboyans V, Ageno W, Agnelli G, Alatri A, Bauersachs R, et al. Diagnosis and management of acute deep vein thrombosis: A joint consensus document from the European Society of Cardiology working groups of aorta and peripheral vascular diseases and pulmonary circulation and right ventricular function. Eur Heart J. 2018;39(47):4208–18.

21. Hobohm L, Lankeit M. Pulmonary Embolism. Pneumologie. 2021;75(10):800–18.

22. Klok FA, van der Hulle T, den Exter PL, Lankeit M, Huisman M V., Konstantinides S. The post-PE syndrome: A new concept for chronic complications of pulmonary embolism. Blood Rev [Internet]. 2014;28(6):221–6. Available from: http://dx.doi.org/10.1016/j.blre.2014.07.003

23. Zöller B, Pirouzifard M, Svensson PJ, Holmquist B, Stenman E, Elston RC, et al. Familial Segregation of Venous Thromboembolism in Sweden: A Nationwide Family Study of Heritability and Complex Segregation Analysis. J Am Heart Assoc. 2021;10(24).

24. Zöller B, Li X, Ohlsson H, Ji J, Sundquist J, Sundquist K. Family history of venous thromboembolism as a risk factor and genetic research tool. Thromb Haemost. 2015;114(5):890–900.

25. Zöller B, Ohlsson H, Sundquist J, Sundquist K. Family history of venous thromboembolism (VTE) and risk of recurrent hospitalization for VTE: A nationwide family study in Sweden. J Thromb Haemost. 2014;12(3):306–12.

26. Beckman MG, Hooper WC, Critchley SE, Ortel TL. Venous Thromboembolism. A Public Health Concern. Am J Prev Med [Internet]. 2010;38(4 SUPPL.):S495–501. Available from: http://dx.doi.org/10.1016/j.amepre.2009.12.017

27. Raskob GE, Silverstein R, Bratzler DW, Heit JA, White RH. Surveillance for Deep Vein Thrombosis and Pulmonary Embolism. Recommendations from a National Workshop. Am J Prev Med [Internet]. 2010;38(4 SUPPL.):S502–9. Available from: http://dx.doi.org/10.1016/j.amepre.2010.01.010

28. Heit JA. The epidemiology of venous thromboembolism in the community. Arterioscler Thromb Vasc Biol. 2008;28(3):370–2.

29. Heit JA, Spencer FA, White RH. The epidemiology of venous thromboembolism. J Thromb Thrombolysis. 2016;41(1):3–14.

30. Cohen AT, Katholing A, Rietbrock S, Bamber L, Martinez C. Epidemiology of first and recurrent venous thromboembolism in patients with active cancer: A population-based cohort study. Thromb Haemost. 2017;117(1):57–65.

31. Lind MM, Johansson M, Själander A, Johansson L. Incidence and risk factors of

venous thromboembolism in men and women. Thromb Res. 2022;214(February):82–6.

32. Zakai NA, Mcclure LA. Racial differences in venous thromboembolism. J Thromb Haemost. 2011;9(10):1877–82.

33. Folsom AR, Basu S, Hong CP, Heckbert SR, Lutsey PL, Rosamond WD, et al. Reasons for Differences in the Incidence of Venous Thromboembolism in Black Versus White Americans. Am J Med. 2019;132(8):970–6.

34. Folsom AR, Lutsey PL, Cushman M. Racial and Regional Differences in Venous Thromboembolism in the United States in Three Cohorts. Circulation. 2014;129(14):1502–9.

35. Schneider D, Lilienfeld DE, Im W. The Epidemiology of Pulmonary Embolism : Racial Contrasts in Incidence and In-Hospital Case Fatality. 2006;98(12):1967–72.

36. White RH, Keenan CR. Effects of race and ethnicity on the incidence of venous thromboembolism. Thromb Res [Internet]. 2009;123(SUPPL. 4):S11–7. Available from: http://dx.doi.org/10.1016/S0049-3848(09)70136-7

37. Law Y, Chan YC, Cheng SWK. Epidemiological updates of venous thromboembolism in a Chinese population. Asian J Surg. 2018;41:176–82.

38. Lee LH, Gallus A. Incidence of Venous Thromboembolism in Asian Populations : A Systematic Review. 2017;

39. Natae SF, Kósa Z, Sándor J, Merzah MA, Bereczky Z, Pikó P, et al. The Higher Prevalence of Venous Thromboembolism in the Hungarian Roma Population Could Be Due to Elevated Genetic Risk and Stronger Gene-Environmental Interactions. Front Cardiovasc Med. 2021;8(October):647416.

40. Morange PE, Suchon P, Trégouët DA. Genetics of venous thrombosis: Update in 2015. Thromb Haemost. 2015;114(5):910–9.

41. Raskob GE, Angchaisuksiri P, Blanco AN, Buller H, Gallus A, Hunt BJ, et al. Thrombosis: A major contributor to global disease burden. Arterioscler Thromb Vasc Biol. 2014;34(11):2363–71.

42. Nichols M, Townsend N, Scarborough P, Rayner M. Cardiovascular disease in Europe: Epidemiological update. J Clin Oncol. 2014;32(21).

43. Tadayon S, Wickramasinghe K, Townsend N. Examining trends in cardiovascular disease mortality across Europe: How does the introduction of a new European Standard Population affect the description of the relative burden of cardiovascular disease? Popul Health Metr. 2019;17(1):1–17.

44. Townsend N, Kazakiewicz D, Lucy Wright F, Timmis A, Huculeci R, Torbica A, et al. Epidemiology of cardiovascular disease in Europe. Nat Rev Cardiol. 2022;19(2):133–43.

45. Cohen AT, Agnelli G, Anderson FA, Arcelus JI, Bergqvist D, Brecht JG, et al. Venous thromboembolism (VTE) in Europe - The number of VTE events and associated morbidity and mortality. Thromb Haemost. 2007;98(4):756–64.

46. Movsisyan NK, Vinciguerra M, Medina-Inojosa JR, Lopez-Jimenez F. Cardiovascular diseases in central and Eastern Europe: A call for more surveillance and evidence-based health promotion. Ann Glob Heal. 2020;86(1):1–10.

47. European Commission. State of Health in the EU Hungary Country Health Profile. 2017.

48. Monreal M, Agnelli G, Chuang LH, Cohen AT, Gumbs PD, Bauersachs R, et al. Deep Vein Thrombosis in Europe—Health-Related Quality of Life and Mortality. Clin Appl Thromb. 2019;25.

49. Kumar DR, Hanlin ER, Glurich I, Mazza JJ, Yale SH. Virchow's contribution to the understanding of thrombosis and cellular biology. Clin Med Res. 2010;8(3–4):168–72.

50. Byrnes JR, Wolberg AS. New findings on venous thrombogenesis. Hamostaseologie. 2017;37(1):25–35.

51. Bovill EG, Van Der Vliet A. Venous valvular stasis-associated hypoxia and thrombosis: What is the link? Annu Rev Physiol. 2011;73:527–45.

52. Hosseinzadegan H, Tafti DK. Prediction of Thrombus Growth: Effect of Stenosis and Reynolds Number. Cardiovasc Eng Technol. 2017;8(2):164–81.

53. Poredos P, Jezovnik MK. Endothelial Dysfunction and Venous Thrombosis. Angiology [Internet]. 2018;69(7):564–7. Available from: https://doi.org/10.1177/0003319717732238

54. David Motto. Endothelial cells and thrombotic microangiopathy. NIH Public Access [Internet]. 2012;32(2):208–14. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3624763/pdf/nihms412728.pdf

55. Monie DD, DeLoughery EP. Pathogenesis of thrombosis: Cellular and pharmacogenetic contributions. Cardiovasc Diagn Ther. 2017;7(Suppl 3):S291–8.

56. Previtali E, Bucciarelli P, Passamonti SM, Martinelli I. Risk factors for venous and arterial thrombosis. Blood Transfus. 2011;9(2):120–38.

57. Huang SS, Liu Y, Jing ZC, Wang XJ, Mao YM. Common genetic risk factors of venous thromboembolism in Western and Asian populations. Genet Mol Res. 2016;15(1):1–11.

58. Ashrani AA, Silverstein MD, Lahr BD, Petterson TM, Bailey KR, Melton LJ, et al. Risk factors and underlying mechanisms for venous stasis syndrome: A population-based case-control study. Vasc Med. 2009;14(4):339–49.

59. Crous-Bou. M, B. Harrington L KC. Environmental and genetic risk factors associated with venous thromboembolism. Semin Thromb Hemost. 2016;42(8):808–20.

60. Jaworek T, Xu H, Gaynor BJ, Cole JW, Rannikmae K, Stanne TM, et al. Contribution of Common Genetic Variants to Risk of Early-Onset Ischemic Stroke. Neurology. 2022;99(16):E1738–54.

61. Zöller B, Svensson PJ, Dahlbäck B, Lind-Hallden C, Hallden C, Elf J. Genetic risk factors for venous thromboembolism. Expert Rev Hematol [Internet]. 2020;13(9):971–81. Available from: https://doi.org/10.1080/17474086.2020.1804354

62. Klarin D, Emdin CA, Natarajan P, Conrad MF, Kathiresan S. Genetic Analysis of Venous Thromboembolism in UK Biobank Identifies the ZFPM2 Locus and Implicates Obesity as a Causal Risk Factor. Circ Cardiovasc Genet. 2017;10(2):1–17.

63. De Haan HG, Bezemer ID, Doggen CJM, Cessie S Le, Reitsma PH, Arellano AR, et al.

Multiple SNP testing improves risk prediction of first venous thrombosis. Blood. 2012;120(3):656–63.

64. Klarin D, Busenkell E, Judy R, Lynch J, Levin M, Haessler J, et al. Genome-wide association analysis of venous thromboembolism identifies new risk loci and genetic overlap with arterial vascular disease. Nat Genet [Internet]. 2019;51(11):1574–9. Available from: http://dx.doi.org/10.1038/s41588-019-0519-3

65. Rinde LB, Morelli VM, Småbrekke B, Mathiesen EB, Løchen ML, Njølstad I, et al. Effect of prothrombotic genotypes on the risk of venous thromboembolism in patients with and without ischemic stroke. The Tromsø Study. J Thromb Haemost. 2019;17(5):749–58.

66. Gindele R, Oláh Z, Ilonczai P, Speker M, Udvari A, Selmeczi A, et al. Founder effect is responsible for the p.Leu131Phe heparin-binding-site antithrombin mutation common in Hungary: Phenotype analysis in a large cohort. J Thromb Haemost. 2016;14(4):704–15.

67. Brotman DJ, Necochea AJ, Wilson LM, Crim MT, Bass EB. Predictive Value of Factor V Leiden and Prothrombin G20210A in Adults With Venous Thromboembolism and in Family Members of Those With a Mutation . Jama. 2009;301(23):2472–85.

68. Van Hylckama Vlieg A, Flinterman LE, Bare LA, Cannegieter SC, Reitsma PH, Arellano AR, et al. Genetic variations associated with recurrent venous thrombosis. Circ Cardiovasc Genet. 2014;7(6):806–13.

69. Kujovich JL. Factor v Leiden thrombophilia. Genet Med [Internet]. 2011;13(1):1–16. Available from: https://doi.org/10.1097/GIM.0b013e3181faa0f2

70. B Simone, V De Stefano, E Leoncini, J Zacho, I Martinelli JE. Risk of venous thromboembolism associated with single and combined effects of Factor V Leiden, Prothrombin 20210A and Methylenetethraydrofolate reductase C677T: a meta-analysis involving over 11,000 cases and 21,000 controls. Eur J Epidemiol [Internet]. 2013;28(8):621–47. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3624763/pdf/nihms412728.pdf

71. Hosseini S, Kalantar E, Hosseini MS, Tabibian S, Shamsizadeh M, Dorgalaleh A. Genetic risk factors in patients with deep venous thrombosis, a retrospective case control study on Iranian population. Thromb J [Internet]. 2015;13(1):4–9. Available from: http://dx.doi.org/10.1186/s12959-015-0064-y

72. Fall AOT, Proulle V, Sall A, Mbaye A, Ba PS, Diao M, et al. Risk factors for thrombosis in an african population. Clin Med Insights Blood Disord. 2014;7:1–6.

73. Abdi AA, Osman A. Prevalence of common hereditary risk factors for thrombophilia in Somalia and identification of a novel Gln544Arg mutation in coagulation factor V. J Thromb Thrombolysis. 2017;44(4):536–43.

74. Jadaon MM. Epidemiology of prothrombin G20210A mutation in the mediterranean region. Mediterr J Hematol Infect Dis. 2011;3(1).

75. Shafia S, Zargar MH, Khan N, Ahmad R, Shah ZA, Asimi R. High prevalence of factor V Leiden and prothrombin G20101A mutation s in Kashmiri patients with venous thromboembolism. Gene. 2018;654(February):1–9.

76. H. AUSTIN, C. DE STAERCKE, C. LALLY, I . D. BEZEMER F. RR and WC. New

gene variants associated with venous thrombosis : a replication study in White and Black Americans. J Thromb Haemost. 2011;9(December 2010):489–95.

77. Cushman M, Pankow JS. Prospective study of circulating factor XI and incident venous thromboembolism: the Longitudinal Investigation of Thromboembolism Etiology (LITE). Am J Hematol. 2015;90(11):1047–51.

78. Senthil K. Vasan, MD, PhD; Klaus Rostgaard, MSc; Ammar Majeed, MD, PhD; Henrik Ullum, MD, PhD; Kjell-Einar Titlestad, MD, PhD; Ole B.V. Pedersen, MD, PhD; Christian Erikstrup, MD, PhD; Kaspar Rene Nielsen, MD, PhD; Mads Melbye, MD, DrMedSci; Olof Nyrén, M P. ABO Blood Group and Risk of Thromboembolic and Arterial Disease A Study of 1.5 Million Blood Donors. Circulation. 2016;1449–57.

79. Dentali F, Sironi A, Ageno W, Turato S, Bonfanti C, Frattini F, et al. Non-O blood type is the commonest genetic risk factor for VTE: Results from a meta-analysis of the literature. Semin Thromb Hemost. 2012;38(5):535–47.

80. Sun X, Feng J, Wu W, Peng M, Shi J. ABO blood types associated with the risk of venous thromboembolism in Han Chinese people : A hospital-based study of. Nat Publ Gr. 2017;(January):1–6.

81. Yu M, Wang C, Chen T, Hu S, Yi K, Tan X. ABO blood groups and risk of deep venous thromboembolism in Chinese Han population from Chaoshan region in South China. 2017;38(4):396–9.

82. Massimo Franchini MM. Non-O blood group: an important genetic risk factor for venous thromboembolism. Blood Transfus. 2013;11(2):2012–3.

83. Miñano A, Ordóñez A, España F, González-porras JR, Lecumberri R, Fontcuberta J, et al. AB0 blood group and risk of venous or arterial thrombosis in carriers of factor V Leiden or prothrombin G20210A polymorphisms. 2008;93(5):729–34.

84. Kinsella M, Monk C. Race, ABO Blood Type and VTE Risk: Not Black and White. Transfusion. 2012;23(1):1–7.

85. Folsom AR, Tang W, Weng LC, Roetker NS, Cushman M, Basu S, et al. Replication of a genetic risk score for venous thromboembolism in whites but not in African Americans. J Thromb Haemost [Internet]. 2016;14(1):83–8. Available from: https://doi.org/10.1111/jth.13193

86. Jiang J, Liu K, Zou J, Ma H, Yang H, Zhang X, et al. Associations between polymorphisms in coagulation-related genes and venous thromboembolism: A meta-analysis with trial sequential analysis. Med (United States). 2017;96(13).

87. Luxembourg B, Pavlova A, Geisen C, Spannagl M, Bergmann F, Krause M, et al. Impact of the type of serpinc1 mutation and subtype of antithrombin deficiency on the thrombotic phenotype in hereditary antithrombin deficiency. Thromb Haemost. 2013;111(2):249–57.

88. Polyak ME, Zaklyazminskaya E V. New genetic variant in the SERPINC1 gene: Hereditary Antithrombin deficiency case report, familial thrombosis and considerations on genetic counseling. BMC Med Genet [Internet]. 2020;21. Available from: https://doi.org/10.1186/s12881-020-01001-5

89. Selmeczi A, Gindele R, Ilonczai P, Fekete A, Komáromi I, Schlammadinger Á, et al. Antithrombin Debrecen (p.Leu205Pro) – Clinical and molecular characterization of a

novel mutation associated with severe thrombotic tendency. Thromb Res. 2017;158:1–7.

90. de la Morena-Barrio ME, Antón AI, Martínez-Martínez I, Padilla J, Miñano A, Navarro-Fernández J, et al. Regulatory regions of SERPINC1 gene: Identification of the first mutation associated with antithrombin deficiency. Thromb Haemost. 2012;107(3):430–7.

91. Mulder R, Croles FN, Mulder AB, Huntington JA, Meijer K, Lukens M V. SERPINC1 gene mutation s in antithrombin deficiency. Br J Haematol. 2017;178(2):279–85.

92. Corral J, de la Morena-Barrio ME, Vicente V. The genetics of antithrombin. Thromb Res [Internet]. 2018;169:23–9. Available from: https://doi.org/10.1016/j.thromres.2018.07.008

93. Tamura S, Hashimoto E, Suzuki N, Kakihara M, Odaira K, Hattori Y, et al. Molecular basis of SERPINC1 mutation s in Japanese patients with antithrombin deficiency. Thromb Res [Internet]. 2019;178:159–70. Available from: https://doi.org/10.1016/j.thromres.2019.04.004

94. Skille H, Paulsen B, Hveem K, Gabrielsen ME, Brumpton B, Hindberg K, et al. Combined effects of five prothrombotic genotypes and cancer on the risk of a first venous thromboembolic event. J Thromb Haemost. 2020;18(11):2861–9.

95. Van Langevelde K, Šrámek A, Rosendaal FR. The Effect of aging on venous valves. Arterioscler Thromb Vasc Biol. 2010;30(10):2075–80.

96. Engbers MJ, van Hylckama Vlieg A, Rosendaal FR. Venous thrombosis in the elderly: Incidence, risk factors and risk groups. J Thromb Haemost. 2010;8(10):2105–12.

97. Donato AJ, Machin DR, Lesniewski LA. Mechanisms of dysfunction in the aging vasculature and role in age-related disease. Circ Res. 2018;123(7):825–48.

98. Karasu A, Šrámek A, Rosendaal FR, van der Geest RJ, van Hylckama Vlieg A. Aging of the venous valves as a new risk factor for venous thrombosis in the elderly: the BATAVIA study. J Thromb Haemost. 2018;16(1):96–103.

99. Genyan Yang, Christine De Staercke and WCH. The effects of obesity on venous thromboembolism: A review. Open J Prev Med. 2012;2(4):499–509.

100. Morgan ES, Wilson E, Watkins T, Gao F, Hunt BJ. Maternal obesity and venous thromboembolism. Int J Obstet Anesth [Internet]. 2012;21(3):253–63. Available from: http://dx.doi.org/10.1016/j.ijoa.2012.01.002

101. Horvei LD, Brækkan SK, Mathiesen EB, Njølstad I, Wilsgaard T, Hansen JB. Obesity measures and risk of venous thromboembolism and myocardial infarction. Eur J Epidemiol. 2014;29:821–30.

102. Hotoleanu C. Association between obesity and venous thromboembolism. Med Pharm Reports. 2020;93(2):162–8.

103. Kabrhel C, Varraso R, Goldhaber SZ, Rimm EB, Camargo CA. Prospective study of BMI and the risk of pulmonary embolism in women. Obesity [Internet]. 2009;17(11):2040–6. Available from: http://dx.doi.org/10.1038/oby.2009.92

104. Cushman M. Epidemiology and Risk Factors for Venous Thrombosis. Semin Hematol. 2007;44(2):62–9.

105. Ageno W, Di Minno MND, Ay C, Jang MJ, Hansen JB, Steffen LM, et al. Association between the metabolic syndrome, its individual components, and unprovoked venous thromboembolism: Results of a patient-level meta-analysis. Arterioscler Thromb Vasc Biol. 2014;34(11):2478–85.

106. Heit JA. Epidemiology of venous thromboembolism. Nat Rev Cardiol [Internet]. 2015;12(8):464–74. Available from: http://www.nature.com/nrcardio/journal/v12/n8/pdf/nrcardio.2015.83.pdf

107. Crous-Bou M, Harrington LB, Kabrhel C. Environmental and Genetic Risk Factors Associated with Venous Thromboembolism. Semin Thromb Hemost. 2016;42(8):808–20.

108. Derek Klarin, Connor A. Emdin, Pradeep Natarajan, Mark F. Conrad SK. Genetic Analysis of Venous Thromboembolism in UK Biobank Identifies the ZFPM2 Locus and Implicates Obesity as a Causal Risk Factor. Circ Cardiovasc Genet. 2017;10(2).

109. Adelborg K, Szépligeti SK, Holland-Bill L, Ehrenstein V, Horváth-Puhó E, Henderson VW, et al. Migraine and risk of cardiovascular diseases: Danish population based matched cohort study. BMJ. 2018;360.

110. Folsom AR, Lutsey PL, Misialek JR, Cushman M. A prospective study of migraine history and venous thromboembolism in older adults. Res Pract Thromb Haemost. 2019;3(3):357–63.

111. Elgendy IY, Nadeau SE, Bairey Merz CN, Pepine CJ, Lindley K, Lindley K, et al. Migraine Headache: An Under-Appreciated Risk Factor for Cardiovascular Disease in Women. J Am Heart Assoc. 2019;8(22).

112. Zhang G, Xu X, Su W, Xu Q. Smoking and risk of venous thromboembolism: a systematic review. Vol. 45, The Southeast Asian journal of tropical medicine and public health. 2014. p. 736–45.

113. Enga KF, Brækkan SK, Hansen-Krone IJ, Le Cessie S, Rosendaal FR, Hansen JB. Cigarette smoking and the risk of venous thromboembolism: The Tromsø Study. J Thromb Haemost. 2012;10(10):2068–74.

114. Cheng YJ, Liu ZH, Yao FJ, Zeng WT, Zheng DD, Dong YG, et al. Current and Former Smoking and Risk for Venous Thromboembolism: A Systematic Review and Meta-Analysis. PLoS Med. 2013;10(9).

115. Lee CWS, Liao CH, Lin CL, Liang JA, Sung FC, Kao CH. Depression and risk of venous thromboembolism: A population-based retrospective cohort study. Psychosom Med. 2015;77(5):591–8.

116. Kunutsor SK, Seidu S, Khunti K. Depression, antidepressant use, and risk of venous thromboembolism: systematic review and meta-analysis of published observational evidence. Ann Med. 2018;50(6):529–37.

117. Karampitsakos T, Koukaki E, Tzouvelekis A, Tzilas V, Bouros E, Dassiou M, et al. Depression, antidepressants and pulmonary embolism. Pneumon. 2018;31(3):147–50.

118. González-Ordóñez AJ, Fernández-Carreira JM, Fernández-Alvarez CR, Obaya RV, Macías-Robles MD, González-Franco A A-GM. The concentrations of soluble vascular cell adhesion molecule-1 and lipids are independently associated with venous thromboembolism. J Hematol. 2003;88(09):1035–43.

119. van Schouwenburg IM, Mahmoodi BK, Gansevoort RT, H. Muntinghe FL, F. Dullaart RP, Kluin-Nelemans HC, Veeger Nic J. G. M MK. Lipid levels do not influence the risk of venous thromboembolism. Results of a population-based cohort study. Thromb Haemost. 2012;108(5):923–9.

120. Kósa Z, Széles G, Kardos L, Kósa K, Németh R, Országh S, et al. A comparative health survey of the inhabitants of Roma settlements in Hungary. Am J Public Health. 2007;97(5):853–9.

121. Rosicova K, Geckova AM, Dijk JP, Rosic M, Zezula I, Groothoff JW. Socioeconomic indicators and ethnicity as determinants of regional mortality rates in Slovakia. Int J Public Health. 2009;54(4):274–82.

122. Rosicova K, Reijneveld SA, Madarasova Geckova A, Stewart RE, Rosic M, Groothoff JW, et al. Inequalities in mortality by socioeconomic factors and Roma ethnicity in the two biggest cities in Slovakia: A multilevel analysis. Int J Equity Health [Internet]. 2015;14(1):1–10. Available from: http://dx.doi.org/10.1186/s12939-015-0262-z

123. Babinská I, Gecková AM, Jarčuška P, Pella D, Mareková M, Štefková G, et al. Does the population living in Roma settlements differ in physical activity, smoking and alcohol consumption from the majority population in Slovakia? Cent Eur J Public Health [Internet]. 2014;22(88):S22–7. Available from: https://doi.org/10.21101/cejph.a3897

124. Škarić-Jurić T, Klarić IM, Narančić NS, Drmić S, Salihović MP, Lauc LB, et al. Trapped between tradition and transition - Anthropological and epidemiological cross-sectional study of Bayash Roma in Croatia. Croat Med J. 2007;48(5):708–19.

125. Paulik Edit, Nagymajtényi Lászlo, Douglas Easterling RTR. Smoking behaviour and attitudes of Hungarian Roma and non- Roma population towards tobacco control policies. Int J Public Heal. 2011;56(5).

126. Vozarova De Courten B, De Courten M, Hanson RL, Zahorakova A, Egyenes HP, Tataranni PA, et al. Higher prevalence of type 2 diabetes, metabolic syndrome and cardiovascular diseases in gypsies than in non-gypsies in Slovakia. Diabetes Res Clin Pract. 2003;62(2):95–103.

127. Bogdanović D, Nikić D, Petrović B, Kocić B, Jovanović J, Nikolić M, et al. Mortality of Roma population in Serbia, 2002-2005. Croat Med J. 2007;48(5):720–6.

128. Ádány R. Roma health is global ill health. Eur J Public Health. 2013;24(5):702–3.

129. Kósa Z, Moravcsik-Kornyicki Á, Diószegi J, Roberts B, Szabó Z, Sándor J, et al. Prevalence of metabolic syndrome among Roma: A comparative health examination survey in Hungary. Eur J Public Health. 2015;25(2):299–304.

130. Fiatal S, Pikó P, Kósa Z, Sándor J, Ádány R. Genetic profiling revealed an increased risk of venous thrombosis in the Hungarian Roma population. Thromb Res [Internet]. 2019;179:37–44. Available from: https://doi.org/10.1016/j.thromres.2019.04.031

131. Kraft P, Hunter D. Integrating epidemiology and genetic association: The challenge of gene-environment interaction. Philos Trans R Soc L B Biol Sci. 2005;360(1460):1609–16.

132. Kraft P, Yen YC, Stram DO, Morrison J, Gauderman WJ. Exploiting gene-environment interaction to detect genetic associations. Hum Hered. 2007;63(2):111–9.

133. Bereczky Z, Gindele R, Fiatal S, Speker M, Miklós T, Balogh L, et al. Age and Origin of the Founder Antithrombin Budapest 3 (p.Leu131Phe) Mutation ; Its High Prevalence in the Roma Population and Its Association With Cardiovascular Diseases. Front Cardiovasc Med. 2021;7(February):1–15.

134. Ádány R, Pikó P, Fiatal S, Kósa Z, Sándor J, Bíró É, et al. Prevalence of insulin resistance in the Hungarian general and roma populations as defined by using data generated in a complex health (Interview and examination) survey. Int J Environ Res Public Health. 2020;17(13).

135. Bergendal A, Bremme K, Hedenmalm K, Lärfars G, Odeberg J, Persson I, et al. Risk factors for venous thromboembolism in pre-and postmenopausal women. Thromb Res. 2012;130(4):596–601.

136. Crous-Bou M, De Vivo I, Camargo CA, Varraso R, Grodstein F, Jensen MK, et al. Interactions of established risk factors and a GWAS-based genetic risk score on the risk of venous thromboembolism. Thromb Haemost. 2016;116(4):705–13.

137. Mi Y, Yan S, Lu Y, Liang Y, Li C. Venous thromboembolism has the same risk factors as atherosclerosis. Medicine (Baltimore). 2016;95(32).

138. Goldhaber SZ. Risk Factors for Venous Thromboembolism. J Am Coll Cardiol. 2010;56(1):1–7.

139. Zhang Z, Li H, Weng H, Zhou G, Chen H, Yang G, et al. Genome-wide association analyses identified novel susceptibility loci for pulmonary embolism among Han Chinese population. BMC Med. 2023;21(1):1–13.

140. Ghouse, J., Tragante, V., Ahlberg G et al. Genome-wide meta-analysis identifies 93 risk loci and enables risk prediction equivalent to monogenic forms of venous thromboembolism,". Nat Genet. 2023;55:399–409.

141. Templeton GF. A two-step approach for transforming continuous variables to normal: Implications and recommendations for IS research. Commun Assoc Inf Syst [Internet]. 2011;28:41–58. Available from: https://aisel.aisnet.org/cais/vol28/iss1/4%0A

142. Ottman R. Theoretical epidemiology: Gene-environment interaction: definitions and study designs. Prev Med (Baltim). 1996;25(6):764–70.

143. Liu CY, Maity A, Lin X, Wright RO, Christiani DC. Design and analysis issues in gene and environment studies. Environ Heal [Internet]. 2012;11. Available from: http://www.ehjournal.net/content/11/1/93

144. Hüls A, Krämer U, Carlsten C, Schikowski T, Ickstadt K, Schwender H. Comparison of weighting approaches for genetic risk scores in gene-environment interaction studies. BMC Genet. 2017;18.

145. Bai J, Ding X, Du X, Zhao X, Wang Z, Ma Z. Diabetes is associated with increased risk of venous thromboembolism: A systematic review and meta-analysis. Thromb Res [Internet]. 2014;135(1):90–5. Available from: http://dx.doi.org/10.1016/j.thromres.2014.11.003

146. Gariani K, Mavrakanas T, Combescure C, Perrier A, Marti C. Is diabetes mellitus a risk factor for venous thromboembolism? A systematic review and meta-analysis of case-control and cohort studies. Eur J Intern Med [Internet]. 2015;28:52–8. Available from: http://dx.doi.org/10.1016/j.ejim.2015.10.001

147. Peng KP, Chen YT, Fuh JL, Tang CH, Wang SJ. Association between migraine and risk of venous thromboembolism: A nationwide cohort study. Headache. 2016;56(8):1290–9.

148. Blom JW, Doggen CJM, Osanto S, Rosendaal FR. Malignancies, Prothrombotic Mutation s, and the Risk of Venous Thrombosis. JAMA [Internet]. 2005;293(6):715–22. Available from: http://jama.jamanetwork.com

149. Buller HR, Van Doormaal FF, Van Sluis GL, Kamphuisen PW. Cancer and thrombosis: From molecular mechanisms to clinical presentations. J Thromb Haemost. 2007;5:246–54.

150. Falanga A, Russo L, Milesi V, Vignoli A. Mechanisms and risk factors of thrombosis in cancer. Crit Rev Oncol Hematol [Internet]. 2017;118:79–83. Available from: http://dx.doi.org/10.1016/j.critrevonc.2017.08.003

151. Skille H, Paulsen B, Hveem K, E. Gabrielsen M, Brumpton B, Hindberg K, et al. Combined effects of five prothrombotic genotypes and cancer on the risk of a first venous thromboembolic event. J Thromb Haemost. 2020;

152. Gran O V., Brækkan SK, Hansen JB. Prothrombotic genotypes and risk of venous thromboembolism in cancer. Thromb Res [Internet]. 2018;164:S12–8. Available from: https://doi.org/10.1016/j.thromres.2017.12.025

153. Auro K, Alanne M, Kristiansson K, Silander K, Kuulasmaa K, Salomaa V, et al. Combined effects of thrombosis pathway gene variants predict cardiovascular events. PLoS Genet. 2007;3(7):1244–53.

154. McDaid A, Logette E, Buchillier V, Muriset M, Suchon P, Pache TD, et al. Risk prediction of developing venous thrombosis in combined oral contraceptive users. PLoS One [Internet]. 2017;12(7). Available from: https://doi.org/10.1371/journal.pone.0182041

155. Pomp ER, Rosendaal FR DCJ. Smoking increases the risk of venous thrombosis and acts synergistically with oral contraceptive use. Am J Hematol. 2008;83(2):97–102.

156. Juul K, Tybjaerg-Hansen A, Schnohr P NB. Factor V Leiden and the Risk for Venous Thromboembolism in the Adult Danish Population. Ann Intern Med. 2004;140(5):330–7.

157. Sejrup JK, Morelli VM, Løchen M-L, Njølstad I, Mathiesen EB, Wilsgaard T, et al. Myocardial infarction, prothrombotic genotypes, and venous thrombosis risk: The Tromsø Study. Res Pract Thromb Haemost. 2020;4(2):247–54.

158. Pikó P, Fiatal S, Kósa Z, Sándor J, Ádány R. Data to genetic risk assessment on high-density cholesterol level associated polymorphisms in Hungarian general and Roma populations. Data Br [Internet]. 2017;14:354–9. Available from: http://dx.doi.org/10.1016/j.dib.2017.07.053

159. Assal S, Susanszky E CA. High consanguinity rate in Hungarian gipsy communities. Acta Paediatr Hung. 1991;31(3):299–304.

160. Eppenberger D, Nilius H, Anagnostelis B, Huber CA, Nagler M. Current Knowledge on Factor V Leiden Mutation as a Risk Factor for Recurrent Venous Thromboembolism: A Systematic Review and Meta-Analysis. Frontiers in Cardiovascular Medicine. 2022; 9:1-13.

161. Eppenberger D, Nilius H, Anagnostelis B, Huber CA, Nagler M. Current Knowledge on Factor V Leiden Mutation as a Risk Factor for Recurrent Venous Thromboembolism: A Systematic Review and Meta-Analysis. Front Cardiovasc Med. 2022;9(April):1–13.

162. Y. L I, ID. BEZEMER, CM. ROWLAND, CH. TONG, AR. ARELLANO JJC, JJ. DEVLIN, PH. REITSMA LB and FR. Genetic variants associated with deep vein thrombosis : the F11 locus. J Thromb Haemost. 2009;7(June):1802–8.

163. Bucciarelli P, Stefano V De, Passamonti SM, Tormene D, Legnani C, Rossi E, et al. In fl uence of proband ' s characteristics on the risk for venous thromboembolism in relatives with factor V Leiden or prothrombin. 2016;122(15):2555–62.

164. Bezgin T, Kaymaz C, Tokgo HC. Thrombophilic Gene Mutation s in Relation to Different Manifestations of Venous Thromboembolism : A Single Tertiary Center Study. Clin Appl Thromb. 2018;

165. Karppinen JE, Törmäkangas T, Kujala UM, Sipilä S, Laukkanen J, Aukee P, et al. Menopause modulates the circulating metabolome: evidence from a prospective cohort study. Eur J Prev Cardiol. 2022;29(10):1448–59.

166. Pate A, Emsley R, Ashcroft DM, Brown B, Van Staa T. Correction: The uncertainty with using risk prediction models for individual decision making: An exemplar cohort study examining the prediction of cardiovascular disease in English primary care (BMC Medicine (2019) 17 (134) DOI: 10.1186/s12916-019-1368-8). BMC Med. 2019;17(1):1–16.

167. Kamal SA, Yin C, Qian B, Zhang P. An interpretable risk prediction model for healthcare with pattern attention. BMC Med Inform Decis Mak [Internet]. 2020;20(11):1–11. Available from: https://doi.org/10.1186/s12911-020-01331-7

168. Shantanam S, MUELLER. Health Risk Prediction Models Incorporating Personality Data: Motivation, Challenges, and Illustration. Pers Disord. 2018;10(1):46–58.

169. Sharma V, Ali I, Van Der Veer S, Martin G, Ainsworth J, Augustine T. Adoption of clinical risk prediction tools is limited by a lack of integration with electronic health records. BMJ Heal Care Informatics. 2021;28(1):2020–2.

170. Ahmad A, Sundquist K, Palmér K, Svensson PJ, Sundquist J, Memon AA. Risk prediction of recurrent venous thromboembolism: a multiple genetic risk model. J Thromb Thrombolysis [Internet]. 2019;47(2):216–26. Available from: http://dx.doi.org/10.1007/s11239-018-1762-7

171. Kolin DA, Kulm S, Elemento O. Prediction of primary venous thromboembolism based on clinical and genetic factors within the U.K. Biobank. Sci Rep [Internet]. 2021;11(1):1–9. Available from: https://doi.org/10.1038/s41598-021-00796-4

172. Hodeib H, Youssef A, Allam AA, Selim A, Tawfik MA, Abosamak MF, et al. Genetic risk profiling associated with recurrent unprovoked venous thromboembolism. Genes (Basel). 2021;12(6):1–11.

173. Thibord F, Klarin D, Brody JA, Chen MH, Levin MG, Chasman DI, et al. Cross-Ancestry Investigation of Venous Thromboembolism Genomic Predictors. Circulation. 2022;146(16):1225–42.

174. Lindström S, Wang L, Smith EN, Gordon W, Van Hylckama Vlieg A, De Andrade M,

et al. Genomic and transcriptomic association studies identify 16 novel susceptibility loci for venous thromboembolism. Blood. 2019;134(19):1645–57.

175. Langley J, Huddleston DE, Merritt M, Chen X, Silver M, Factor SA, et al. Genome-wide association analysis of venous thromboembolism identifies new risk loci and genetic overlap with arterial vascular disease. Hum Brain Mapp. 2017;37(7):2547–56.

10 Keywords

GxE interaction, prothrombotic SNPs, risk prediction, SERPINC1 mutation, venous thrombosis (VT), Hungarian, Roma population.

11 Acknowledgements

First and foremost, I would like to thank my Lord and His Mother St. Virgin Mary for their support and unexplained gift during my entry life and in my PhD journey.

I would also like to extend my thanks and appreciation to my supervisor Dr. Szilvia Fiatal for her unreserved support, guidance, and encouragement to reach at this stage.

I would also like to extend my sincere gratitude to the Dean of the Doctoral School of Health Sciences, Professor Mariann Harangi, the secretarial, Dr.Mónika Katkó and the head of the Department of Public Health and Epidemiology, Professor János Sándor, and all staff of the Department of Public Health and Epidemiology who support me in a different perspective for the fruitfulness of my study.

My sincere thanks to my special and strong mother Aregash Guddisa, my brother Jinenus Fituma, and my children Meskerem Tarressa, Soreti Dereje, and Bona Ayele for their patience and unexplained encouragement for the success of my PhD Journey.

Lastly, I dedicate this dissertation to the Soul of my Hero Father Fituma Natae Murti; who has always been my energy, confidence, value, and motivator to be a risk taker, influential, and successful woman in any journey in addition to my study. I always pray and ask my GOD to rest his Soul in heaven. Amen, Amen, Amen.

12 Funding

This work was supported by the Stipendium Hungaricum Scholarship Programme of the Tempus Public Foundation, the European Union under the European Regional Development Fund (GINOP-2.3.2-15-2016-00005), and the National Research, Development and Innovation Office, Hungarian Ministry of Innovation and Technology (Grant No. OTKA K139293).

13 Publications

UNIVERSITY of DEBRECEN

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

Registry number: DEENK/507/2023.PL
Subject: PhD Publication List

Candidate: Shewaye Natae
Doctoral School: Doctoral School of Health Sciences
MTMT ID: 10086242

List of publications related to the dissertation

1. **Natae, S.**, Merzah, M., Sándor, J., Ádány, R., Bereczky, Z., Fiatal, S.: A combination of strongly associated prothrombotic single nucleotide polymorphisms could efficiently predict venous thrombosis risk.
   Front. Cardiovasc. Med. 10, 1-11, 2023.
   DOI: http://dx.doi.org/10.3389/fcvm.2023.1224462
   IF: 3.6 (2022)
2. **Natae, S.**, Kósa, Z., Sándor, J., Merzah, M., Bereczky, Z., Pikó, P., Ádány, R., Fiatal, S.: The Higher Prevalence of Venous Thromboembolism in the Hungarian Roma Population Could Be Due to Elevated Genetic Risk and Stronger Gene-Environmental Interactions.
   Front. Cardiovasc. Med. 8, 1-13, 2021.
   DOI: http://dx.doi.org/10.3389/fcvm.2021.647416
   IF: 5.846

List of other publications

3. Merzah, M., Póliska, S., Balogh, L., Sándor, J., Szász, I., **Natae, S.**, Fiatal, S.: A Transcriptomic Analysis of Smoking-Induced Gene Expression Alterations in Coronary Artery Disease Patients.
   Int. J. Mol. Sci. 24 (18), 1-14, 2023.
   DOI: http://dx.doi.org/10.3390/ijms241813920
   IF: 5.6 (2022)
4. **Natae, S.**, Nigatu, D. T., Negawo, M. K., Mengesha, W. W.: Cervical cancer screening uptake and determinant factors among women in Ambo town, Western Oromia, Ethiopia: community-based cross-sectional study.
   Cancer Medicine. 10 (23), 8651-8661, 2021.
   DOI: http://dx.doi.org/10.1002/cam4.4369
   IF: 4.711

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

5. Merzah, M., Kósa, Z., Sándor, J., **Natae, S.**, Pikó, P., Ádány, R., Fiatal, S.: Roma Socioeconomic Status Has a Higher Impact on Smoking Behaviour than Genetic Susceptibility.
   Int. J. Environ. Res. Public Health. 18 (6), 1-12, 2021.
   DOI: http://dx.doi.org/10.3390/ijerph18063206
   IF: 4.614

6. Nigatu, D. T., Seme, A., **Natae, S.**, Segni, M. T.: Sexual initiation and associated factors among young women in West Shoa, Ambo Town, Ethiopia: a community-based cross-sectional study.
   BMC Womens Health. 18 (1), 1-6, 2018.
   DOI: http://dx.doi.org/10.1186/s12905-018-0563-7
   IF: 1.592

Total IF of journals (all publications): 25,963
Total IF of journals (publications related to the dissertation): 9,446

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of the Journal Citation Report (Impact Factor) database.

13 November, 2023

DEBRECENI EGYETEM
Egyetemi és Nemzeti Könyvtár