Universidade de Lisboa

Faculdade de Farmácia

PREVALENCE OF THROMBOGENIC RISK FACTORS IN THE PORTUGUESE POPULATION AND IDENTIFICATION OF HIGH-RISK INDIVIDUALS IN PORTUGAL

Micaela Patrícia Gomes dos Santos

Dissertation supervised by Doctor Ana Catarina Alves
and co-supervised by Professor Isabel Antolin Rivera

Master in Biopharmaceutical Sciences

2020

Universidade de Lisboa

Faculdade de Farmácia

PREVALENCE OF THROMBOGENIC RISK FACTORS IN THE PORTUGUESE POPULATION AND IDENTIFICATION OF HIGH-RISK INDIVIDUALS IN PORTUGAL

Micaela Patrícia Gomes dos Santos

Dissertation supervised by Doctor Ana Catarina Alves
and co-supervised by Professor Isabel Antolin Rivera

Master in Biopharmaceutical Sciences

2020

Acknowledgements

This thesis is dedicated to my parents: I am grateful for the trust you placed in me and for all the sacrifices and investments in me and in my education. Without your support this journey would have been impossible and so I will always be thankful.

I would like to express my sincere gratitude to Dr. Ana Catarina Alves for all her guidance, encouragement and time spent during the course and improvement of this thesis. I also leave a special appreciation to Dr. Mafalda Bourbon for receiving me in her research group, giving me the opportunity to develop such project at the INSA. Finally, but not least, I am thankful to Professor Isabel Antolin Riviera who was always prompted to help during this journey.

To my friends, who were always present and believed in me whenever I didn't: it was a rough year, but you constantly find a way to make it easier, pulling me up whenever I needed an extra strength.

Resumo

Introdução: A trombose é uma patologia comum e complexa, referente à formação de um coágulo, que dificulta o normal fluxo sanguíneo. Sob condições normais, os mecanismos procoagulantes, anticoagulantes e fibrinolíticos regulam a hemostasia para evitar a formação patológica destes coágulos. No entanto, alterações nos genes de proteínas envolvidas direta ou indiretamente na regulação hemostática podem levar a um aumento anómalo da coagulação e a uma maior predisposição à formação de coágulos obstrutivos. Esta condição é conhecida como trombofília e descreve um estado de hipercoagulabilidade responsável pelo desenvolvimento de eventos trombóticos que comprometem a vida (como o enfarte do miocárdio, acidente vascular cerebral e embolia pulmonar) e que geralmente se desenvolvem na presença de um ou mais fatores de risco herdados (variantes genéticas) e ambientais (adquiridos). A trombose tem, portanto, uma origem multifatorial causada pela interação destes fatores de risco herdados e/ou outros adquiridos (tabagismo, gravidez, uso de contracetivos orais e terapia hormonal), que podem coexistir no mesmo indivíduo e conferir um risco contínuo ao longo da vida. Devido ao seu estado adquirido de hipercoagulabilidade, as grávidas apresentam ainda um risco acrescido à manifestação de tromboses, traduzido pela ocorrência de abortos espontâneos recorrentes. Através da eliminação de fatores de risco adquiridos é possível prevenir (até certo ponto) o desenvolvimento de trombose. É assim importante conhecer o perfil genético da população de forma a reunir dados que permitam a identificação de indivíduos de alto risco e antecipar a ocorrência destes eventos, bem como desenvolver medidas preventivas dirigidas aos mesmos. Os objetivos deste projeto são então estimar a prevalência das variantes trombogénicas *FV G1691A* (*FV Leiden*), *FII G20210A*, *AT Cambridge II*, *PAI-1 4G/5G*, *MTHFR C677T* e *MTHFR A1298C* na população portuguesa, estabelecer associações entre a presença destas variantes genéticas e outros fatores de risco adquiridos no desenvolvimento de eventos trombóticos e identificar grupos de indivíduos de alto risco em Portugal.

Metodologia: 774 participantes do estudo e_COR, um estudo epidemiológico transversal, previamente desenvolvido no Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA) com participantes das 5 regiões de Portugal Continental (Norte, Centro, Lisboa, Alentejo e Algarve), foram aleatoriamente selecionados para o presente estudo de forma a serem representativos da população portuguesa. A todos os participantes foi realizado um exame físico, bem como um questionário com informação clínica e pessoal necessária à realização do estudo. Após recolha das amostras de sangue, os

parâmetros bioquímicos foram determinados e o ADN extraído, procedendo-se à genotipagem das variantes *FV Leiden*, *FII* G20210A, *AT* Cambridge II, *PAI-1* 4G/5G, *MTHFR* C677T e *MTHFR* A1298C pela tecnologia *TaqMan® OpenArray™.* As frequências alélicas e genotípicas das variantes foram determinadas de forma a estimar a sua prevalência na população portuguesa. Realizaram-se correlações entre os genótipos e a informação pessoal, bioquímica e clínica dos participantes, tendo sido também determinadas as frequências de subgrupos em risco acrescido de trombose.

Resultados: Dos indivíduos estudados, apenas 2,2%, 4,1% e 0,6% apresentavam, respetivamente, as variantes genéticas *FVL*, *FII* G20210A e *AT* Cambridge II em heterozigotia, não sendo identificados casos de homozigotia para estas alterações genéticas. Por outro lado, as variantes *PAI-1* 4G/5G e *MTHFR* C677T e A1298C revelaram estar amplamente distribuídas pela população: 50,4% e 21,8% dos participantes eram respetivamente heterozigotos e homozigotos para o *PAI-1* 4G/5G; na mesma ordem, 43.2% e 12,8% eram heterozigotos e homozigotos para a *MTHFR* C677T; e 39,5% e 9,8% exibiam heterozigotia e homozigotia para a *MTHFR* A1298C. Apenas 3,6% dos participantes não apresentaram qualquer alelo de risco trombogénico para as variantes em estudo. A maioria dos Portugueses (69% da população) eram portadores de pelo menos 2 variantes genéticas – a combinação mais prevalente foi a variante PAI-1 4G/5G com as variantes MTHFR C677T (24%) ou A1298C (19,8%), englobando ambos os genótipos hetero- e homozigótico. No que diz respeito aos fatores de risco trombogénicos adquiridos, observou-se que: 64,7% dos indivíduos apresentavam excesso de peso, 19,4% hiperhomocisteinemia e 24,3% hábitos tabágicos. Estes fatores revelaram ser mais prevalentes entre o sexo masculino. As mulheres podem apresentar um risco acrescido pelo uso de contracetivos orais (29,5% da população feminina em estudo) ou tratamentos hormonais (12,3%). Foi também analisada a ocorrência de enfarte do miocárdio e acidente vascular cerebral na população, bem como de abortos espontâneos entre as mulheres. Comparando os indivíduos que não desenvolveram nenhum evento trombótico, com o grupo de indivíduos que sofreu de enfarte do miocárdio, verificou-se que estes englobavam, com maior frequência, portadores das variantes *FVL* (5,3% *vs.* 2,1%) e *FII* G20210A (15,8% *vs.* 3,7%), assim como de homozigotos para a variante *PAI-1* 4G/5G (31,6% *vs.* 21,6%). O grupo com diagnóstico de AVC apresentava maior frequência de heterozigotos para o *FVL* (13,0% *vs.* 1,9%), de heterozigotos para a *MTHFR* C677T (65,2% *vs.* 42,4%), bem como, de homozigotos para o *PAI-1* 4G/5G (30,4% *vs.* 21,6%). Por fim, entre as mulheres que sofreram de aborto espontâneo, observou-se uma maior frequência dos genótipos heterozigóticos para o *FVL* (4,5% *vs.* 2%) e *PAI-1* 4G/5G (56,8 % *vs.* 47%),

assim como, para o genótipo homozigótico da variante *MTHFR* A1298C (15,9% *vs.* 9,2%, p=0,037). Por outro lado, as mulheres fumadoras que tiveram pelo menos um aborto espontâneo apresentavam maior frequência do genótipo homozigótico da *MTHFR* C677T (18,2% *vs.* 3%).

Discussão: A variante genética mais prevalente neste estudo foi a *PAI-1* 4G/5G, seguida pelas variantes *MTHFR* C677T e A1298C. A prevalência estimada para as variantes *FVL* e *FII* G20210A é claramente menor que as anteriores e a *AT* Cambridge II revelou-se rara na população Portuguesa. Embora as prevalências estimadas para Portugal se encontrem de acordo com as frequências esperadas para populações Caucasianas, as frequências determinadas para o *FVL* e *FII* G20210A diferem entre países Europeus. As variantes *FVL*, *FII* G20210A e *MTHFR* C677T parecem influenciar a ocorrência de trombose arterial. Pela análise realizada neste estudo, 0,3% da população pode estar em risco uma vez que apresenta as três variantes em simultâneo. Apresentar o genótipo *PAI-1* 4G/4G pode também influenciar estas ocorrências, colocando 21,8% da população portuguesa em maior risco trombótico, particularmente aqueles que apresentam simultaneamente hipercoagulabilidade e hipofibrinólise: 1,9% e 2,8% da população que é também portadora das variantes *FVL* ou *FII* G20210A, respetivamente. 43,8% da população Portuguesa apresenta simultaneamente a variante genética *PAI-1* 4G/5G com uma das variantes da *MTHFR*, o que sugere um perfil trombofílico caraterizado por fibrinólise anormal, maior procoagulação, possível aumento da inflamação, bem como, disfunção endotelial. As mulheres podem ter ainda acrescido risco pelo uso de contracetivos orais e por serem portadoras de pelo menos um alelo de risco das variantes: *PAI-1* 4G/5G (presente em cerca de 20% da população feminina), *MTHFR* C677T (cerca de 15%) e *MTHFR* A1298C (cerca de 15%). Têm também um risco aumentado 6 a 10% da população feminina portadora de pelo menos um dos alelos de risco para cada uma dessas variantes referidas e sob tratamento hormonal. Verificou-se que as mulheres que tomam contracetivos orais e apresentam as variantes *FVL* (0,2%) e *FII* G20210A (0,7%), assim como, as que fazem terapia hormonal e são portadoras do *FII* G20210A (0,7%) estão em maior risco, especialmente se forem obesas e/ou fumadoras. As grávidas que apresentam a variante *MTHFR* A1298C correm maior risco de aborto espontâneo, apesar das variantes *FVL* e *PAI-1* 4G/5G poderem também influenciar o risco. As variantes *FII* G20210A e *MTHFR* C677T (isoladamente) não parecem influenciar a ocorrência de aborto; no entanto, a homozigotia para a *MTHFR* C677T juntamente com o tabagismo aparentam ter influência nessas incidências. A presença de pelo menos

um alelo de risco de qualquer uma das variantes da *MTHFR* em fumadoras pode também favorecer a ocorrência de aborto.

Conclusão: A população portuguesa apresenta vários fatores de risco genéticos e adquiridos que podem interagir e culminar na incidência de eventos trombóticos. Consciencializar para a trombofilia é crítico para a prevenção, quer seja pela redução de fatores de risco adquiridos (mudanças no estilo de vida) ou pela otimização da terapêutica, que pode melhorar o prognóstico destes indivíduos. Desta forma, os resultados obtidos no presente estudo são importantes para a elaboração de estratégias terapêuticas e diretrizes de saúde para a identificação dos grupos de maior risco trombótico.

Palavras-chave: Trombose, trombofilia, hipercoagulabilidade, fatores de risco, variantes trombogénicas, contraceptivos orais, tabagismo, enfarte do miocárdio, acidente vascular cerebral, aborto espontâneo.

Abstract

Introduction: Thrombosis is a pathology in which a blood clot (thrombus) hampers the normal blood flow to a tissue. Under normal conditions, procoagulant, anticoagulant and fibrinolytic pathways regulate hemostasis to avoid pathological clot formation. However, changes in the genes encoding proteins, directly or indirectly involved in hemostatic regulation, can lead to enhanced abnormal blood coagulation, resulting in an increased disposition to the formation of obstructive clots, known as thrombophilia. This state of hypercoagulability is responsible for the development of life-threatening thromboembolic events (like myocardial infarction, ischemic stroke and pulmonary embolism), which usually develop when one or more of both genetic (inherited) and environmental (acquired) risk factors come into play. Pregnant women are also at major risk to manifest thrombosis by recurrent spontaneous abortion episodes. Thus, thrombosis results from the interplay of inherited DNA changes and/or acquired risk factors (cigarette smoking, pregnancy and oral contraceptive or hormone therapy use) that may coexist in the same individual; however, it is preventable. Hence, it is important to recognize the genetic thrombogenic profile of the population to assemble data for identification of high-risk individuals and to develop local guidelines for prevention. With this in mind, this project aims to estimate the prevalence of the *FV G1691A* (*FV Leiden*), *FII G20210A*, *AT Cambridge II*, *PAI-1 4G/5G* and *MTHFR C677T* and *A1298C* thrombogenic variants in the Portuguese population, to establish associations between the presence of these genetic variants and the presence of other increasing risk factors in the development of thrombotic-related events and to identify groups of high-risk individuals in Portugal.

Methodology: 774 e_COR participants from the North, Center, Lisbon, Alentejo and Algarve regions of Portugal were randomly selected to be representative of the Portuguese population. All participants were physically examined and had their personal and recent clinical data collected, as well as blood samples for biochemical analysis and DNA extraction. Participants were genotyped for the *FV Leiden*, *FII G20210A*, *AT Cambridge II*, *PAI-1 4G/5G*, *MTHFR C677T* and *MTHFR A1298C* variants by the TaqMan® OpenArray™ Technology. Allelic and genotypic frequencies were determined to estimate the prevalence of these thrombogenic variants in the Portuguese population. Correlations were made between the genotype data plus personal, biochemical and clinical information available and frequencies of subgroups of individuals with increased risk of thrombosis were determined.

Results: Only 2.2%, 4.1% and 0.6% of the population were carriers of the *FVL*, the *FII* G20210A and the *AT* Cambridge II variants respectively, and no homozygosity was found for these genetic defects. In contrast, the *PAI-1* 4G/5G and the *MTHFR* C677T and A1298C variants were largely distributed among the population: 50.4% and 21.8% were, respectively, heterozygous and homozygous for the *PAI-1* 4G/5G; 43.2% and 12.8% were, in the same order, heterozygous and homozygous for the *MTHFR* C677T; and 39.5% and 9.8% displayed the heterozygous and homozygous genotype of the *MTHFR* A1298C variant. Only 3.6% of the participants did not display inherited genetic risk alleles for neither thrombogenic variants, while the majority (69% of the population) were carriers of at least two different genetic variants. The most prevalent coinheritances were the combination of the *PAI-1* 4G/5G variant with either the *MTHFR* C677T (24%) or the *MTHFR* A1298C (19.8%) variants, in heterozygous and/or homozygous genotypes. High frequencies of individuals with acquired thrombogenic risks were also observed: overweight or obesity (64.7%), hyperhomocysteinemia (19.4%) and cigarette smoking habits (24.3%); which were more prevalent among men. Women can be at especially increased risk by using oral contraceptives (29.5% of this study female population) and hormone therapy treatments (12.3%). Some participants had also developed thrombotic-related events. Among those who suffered a myocardial infarction, it was detected an increased heterozygous frequency for the *FVL* (5.3% *vs.* 2.1%) and *FII* G20210A (15.8% *vs.* 3.7%) and an increased frequency for the *PAI-1* 4G/5G homozygous genotype (4G/4G) (31.6% *vs.* 21.6%). The group with stroke diagnosis revealed a higher heterozygous frequency for the *FVL* (13.0% *vs.* 1.9%) and *MTHFR* C677T (65.2% *vs.* 42.4%) and an increased homozygous frequency for the *PAI-1* 4G/4G variant (30.4% *vs.* 21.6%). Finally, among women who had a spontaneous abort, it was observed a higher frequency of *FVL* (4.5% *vs.* 2%) and *PAI-1* 4G/5G (56.8% *vs.* 47%) heterozygous genotypes and of the *MTHFR* A1298C homozygous genotype (15.9% *vs.* 9.2%, p=0.037). Additionally, women smokers who had spontaneous abortion displayed a higher frequency of the *MTHFR* C677T homozygous genotype (18.2% *vs.* 3%).

Discussion: The *PAI-1* 4G/5G was the most prevalent variant in mainland Portugal, followed by the *MTHFR* C677T and A1298C variants. The estimated prevalence of *FVL* and *FII* G20210A variants was considerably lower, while the *AT* Cambridge II was quite rare. Even though the variants' estimated Portuguese prevalence was within the same range for Caucasian populations, the *FVL* and *FII* G20210A frequencies appear to differ among European countries. In this study, the *FVL*, *FII* G20210A and *MTHFR* C677T variants seemed to impact the onset of arterial thrombosis; so, 0.3% of the population may be at increased risk due to coinheritance of these three

genetic factors. The *PAI-1* 4G/4G genotype may also be involved in the arterial thrombosis pathology, putting 21.8% of the Portuguese population at higher thrombotic-risk, particularly those with the synergistic combination of hypercoagulability and hypofibrinolysis: 1.9% and 2.8% of the population who are simultaneously carriers of the *FVL* or the *FII* G20210A variants, respectively. However, the most prevalent coinheritances in Portugal were found to be the combination of *PAI-1* 4G/5G with the *MTHFR* variants (43.8% of the population), suggesting a thrombophilic genetic profile characterized by impaired fibrinolysis, greater procoagulant activity, possible enhanced inflammation and endothelial dysfunction. Women may also be at particular risk by using oral contraceptives and carrying at least one risk allele of either one of these variants: *PAI-1* 4G/5G (about 20% of all women), *MTHFR* C677T (about 15%) and *MTHFR* A1298C (about 15%). Another female fraction of about 6 to 10%, carriers of *PAI-1* 4G/5G, *MTHFR* C677T or *MTHFR* A1298C variants, is also at risk due to the concurrent use of hormone therapy. Women users of oral contraceptives who were also carriers of the *FVL* (0.2%) and *FII* G20210A (0.7%) variants, or those who do hormone therapy and were found to be FII G20210A carriers (0.7%) are expected to be the ones at high-risk, especially if they are obese and/or smokers. Pregnant carriers of the *MTHFR* A1298C can be at high-risk of spontaneous abortion, whereas the *FVL* and the *PAI-1* 4G/5G variants may also have some influence in such incidence. The *FII* G20210A did not seemed to influence the occurrence of spontaneous abortion, neither the *MTHFR* C677T, at least by itself. Taking into consideration a possible synergistic risk of smoking habits, the *MTHFR* C677T homozygous genotype may potentiate the occurrence of spontaneous abortion and the inheritance of at least one risk allele of either one of the *MTHFR* variants among smokers may indeed enhance the risk of spontaneous abortion.

Conclusion: The Portuguese population was shown to present several genetic and acquired thrombotic risk factors, which can interact and culminate in the onset of the disease. Therefore, awareness about thrombophilia can be critical for prevention, either for the elimination of acquired risk factors (lifestyle changes) or to optimize prophylaxis. Our findings might be important to elaborate health guideline strategies and to identify thrombotic high-risk groups that deserve special attention.

Keywords: Thrombosis, thrombophilia, hypercoagulability, risk factors, thrombogenic variants, oral contraceptives, cigarette smoking, myocardial infarction, stroke, spontaneous abortion.

Abbreviations

A	Adenine
APC	Activated Protein C
APC-R	Activated Protein C Resistance
AT	Antithrombin
BMI	Body mass index
bp	Base Pairs
C	Cytosine
CI	Confidence Intervals
COCs	Combined Oral Contraceptives
CS	Cigarette Smoking
CVD(s)	Cardiovascular Disease(s)
dbSNP	Single Nucleotide Polymorphism Database
DNA	DeoxyriboNucleic Acid
dNTPs	Deoxynucleotides
DVT	Deep Venous Thrombosis
EC(s)	Endothelial Cell(s)
EDTA	EthyleneDiamine Tetraacetic Acid
F	Clotting/Coagulation Factor
FVL	*Factor V Leiden*
G	Guanine
Hcy	Homocysteine
HHcy	Hyperhomocysteinemia
hmz	Homozygous
HT	Hormone Therapy
htz	Heterozygous
HWE	Hardy-Weinberg Equilibrium
INSA	National Institute of Health Dr. Ricardo Jorge
IS	Ischemic Stroke
MGB	Minor Groove Binder
MI	Myocardial Infarction
MS	Methionine Synthase
MTHF	Methylene Tetrahydrofolate
MTHFR	Methylene Tetrahydrofolate Reductase
NA	Not Applicable

NCD(s)	Non-Communicable disease(s)
OC(s)	Oral Contraceptive(s)
PAI-1	Plasminogen Activator Inhibitor-1
PC	Protein C
PCR	Polymerase Chain Reaction
PE	Pulmonary Embolism
PS	Protein S
RFLP	Restriction Fragment Length Polymorphism
RNA	RiboNucleic Acid
SA	Spontaneous Abortion
SAP	*Shrimp Alkaline Phosphatase*
T	Thymine
TBE	Tris/Borate/EDTA
TF	Tissue Factor / Tissue Thromboplastin
THF	Tetrahydrofolate
UDR	Diagnosis and Reference Unit
UV	Ultra-violet
VTE	Venous Thromboembolism (Venous Thrombosis)
wt	Wild-type

Content

Figure List

Table List

1. Introduction

The improvement of public health and expanded sanitation conditions, together with the modern healthcare and medicine developments, had allowed the mortality rates to decrease and to double the average life expectancy, being observed a quadruplication of the world population during the 20th century. However, world globalization, moved by population ageing, has potentiate a transition from infectious (communicable) diseases to chronic (non-communicable) diseases, which have replaced infectious illnesses and malnutrition as the new leading cause of death and disability worldwide[1–5]. Unlike infectious illnesses, **non-communicable diseases (NCDs)** do not have a known causative transmission agent[6]; instead, they result on the combined effect of genetic, physiological, environmental and behavior factors that increase the odds of disorder development[6,7].

According to the *World Health Organization*, NCDs are responsible for 41 million deaths each year (71% of all global deaths) with 15 million people of those who die aged between 30 and 69 years ("premature" deaths)[7]. Hence, NCDs are currently seen as a main emerging public health threat in humanitarian settings, also displaying growing impact over time and high costs to society[3,4]. **Cardiovascular diseases (CVDs)** are the leading contributor to the NCDs' burden[7], being mainly developed due to the combined presence of heritable and environmental factors and the harmful interaction between them. Besides largely preventable, CVDs are responsible for almost one-third of all global deaths and still a remaining cause of morbidity despite the outcome improvements[8–11]. There is therefore an emerging need to improve disease managing, by providing patients' access to palliative care, detection, screening and treatment, and to develop new preventive measures[5,7]. Only by elimination of health risk behaviors, at least 80% of CVDs' events would possibly be prevented and their prevalence would markedly reduce. Predictive medicine should then be a high priority: healthy lifestyles should be promoted among the general population and high-risk individuals must be clearly identified, adopting the best preventive approaches according to the associated risk[8,12].

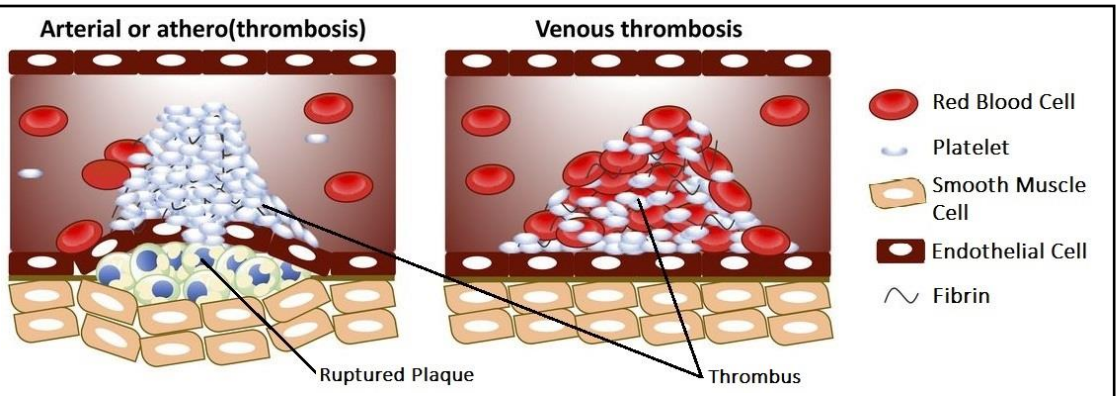
1.1. Thrombosis: An Underlying Pathology Among Cardiovascular Disorders

Thrombosis is a common and complex condition that underlies most CVDs[1,13] and in which a blood clot (known, in this case, as a ‘‘thrombus’’) hampers the normal

blood flow to a tissue, provoking ischemia and tissue death[14,15]. This condition has been historically considered to exist by two pathophysiological forms; **arterial thrombosis** (or atherothrombosis), related to platelet activation and generally occurring following an atherosclerotic plaque rupture that damages endothelial cells and leads to the local formation of a platelet clump; and **venous thrombosis** (or venous thromboembolism), whereas clotting system is activated without endothelial disruption (**Figure 1**). Despite being types of CVDs differing in their etiology, both arterial and venous thrombosis are main causes of high mortality and morbidity worldwide[16–21].

Atherothrombosis is the most common type of thrombosis, with increasing prevalence in developed countries, and occurring frequently in CVDs high-risk individuals[20,22]. It is the central pathological mechanism underlying **myocardial infarction (MI)** and **ischemic stroke (IS)**, together the leading causes of death in developed countries[22,23]. The development of both disorders is very similar and their clinical presentation (as well as their consequences) is dictated by the location (and extent) of the thrombus[24], mainly if it obstructs the coronary or cerebral circulation, respectively affecting the arteries supplying oxygen to the heart (MI) or brain (IS)[24–26]. Although it may not cause symptoms, the formation of a thrombus in the arterial circulation promotes the rapid progression of atherosclerotic lesions[20]. Atherosclerotic plaques are extremely dangerous due to their high content of **tissue factor (TF)**, a property conferring greater thrombogenicity[27]. After plaque erosion or rupture, TF contacts with the circulating blood and triggers the blood coagulation cascade, that stimulates platelet aggregation (essential for thrombus stability). Thereby, thrombus forms over the plaque, leading to vascular occlusion, reduction of blood flow and necrosis of the tissue perfused by the occluded artery[27–29].

The third leading vascular disease (after MI and IS), with a global incidence of about 10 million cases *per* year, is **venous thromboembolism (VTE)**[25,30], being particularly common among European descendants[18]. It comprises two main related clinical events: **deep venous thrombosis (DVT)**, characterized by clots formation in the legs and arms' veins; and **pulmonary embolism (PE)**, when the embolized clots travel to the lungs[18,25,31,32]. Those two incidences often constitute a single clinical event in which DVT is followed by PE, with PE causing most of VTE deaths[25,30]. Besides being a common illness, the underlying pathogenic mechanisms of VTE are only partially known in comparison to those of atherothrombosis[25]. However, it is accepted that the occurrence of venous thrombosis specially relies in blood stasis and hypercoagulability states, much more than endothelial damage (the main determinant of arterial thrombosis) [25,33]; moreover, venous thrombi are mostly formed by fibrin and red blood cells, and less by platelets (the main arterial thrombi component)[25].



Adapted from: Yeung, J., Li, W. and Holinstat, M. (2018). 'Platelet Signaling and Disease: Targeted Therapy for Thrombosis and Other Related Diseases', Pharmacological Reviews. Edited by L. L. Isom, 70(3), pp. 526–548.

Figure 1| Thrombosis pathophysiological forms: Arterial thrombosis (or atherothrombosis) and venous thrombosis (or venous thromboembolism). *While arterial thrombosis develops due to plaque rupture, that causes endothelial cells' damage and leads to the formation of a clot mainly composed by platelets, venous thrombosis relies on the formation of a clot triggered by the coagulation system activation and is mostly composed by fibrin and red blood cells.*

Thrombosis is a typical multifactorial disease and a life-long risk since it involves both environmental and genetic mechanisms[23] and it usually occurs when one or more of the circumstantial factors (genetic as well as acquired risk factors) come into play[13,23]. Both arterial and venous thrombosis are under the significant influence of a genetic component that confers predisposition to thrombotic events[22]. The associated genetic risks are mostly due to mutations that cause an imbalance between procoagulant and anticoagulant forces, affecting the natural anticoagulant mechanisms and leading to a hypercoagulable state, responsible for the development of most life-threatening thromboembolic disorders[13,23]. This hypercoagulable state can also be promoted by cardiovascular risk factors that trigger blood thrombogenicity, being associated to numerous thrombotic complications[27].

1.2. The Role of Hemostasis on the Vascular System Integrity and on Thrombophilia Emergence

The vascular system (also known as circulatory system) is made up of blood vessels, in which blood circulates under pressure through the vasculature. In case of a vascular system breakdown, the damage must be quickly repaired to maintain the blood flow and the vasculature integrity. Any escaping blood needs to be rapidly converted into a gel (the ''clot''), that will plug the hole to prevent blood from leaving the vascular compartment and enable subsequent repair events. Such events aim not only to prevent excessive bleeding, by confining the circulating blood within the veins and arteries, but

also keeping it in its fluid liquid state and under circulation, through inhibition of the intravascular coagulation[14,34–36]. This physiological response to damage is known as **hemostasis,** a complex well-regulated process that allows to *maintain the closed high-pressure circulatory system integrity*[37–39]. Hemostasis is conventionally divided into a primary or a secondary process. While the injury of small vessels triggers the short-lived response of primary hemostasis (associated not only with endothelial damage but also to normal endothelial renewal) secondary hemostasis is required to control bleeding from large wounds and ease hemorrhage arrest[35,37,39,40]. The formation of a blood clot at damaged site is termed coagulation and It involves direct activation, adhesion, and aggregation of platelets (primary hemostasis), together with deposition and maturation of a fibrin network (secondary hemostasis)[41]. As healing is occurring, clots are dissolved through **fibrinolysis**[35,42]. Despite usually described separately, those mechanisms occur all together at the injured vessel site (**Figure 2**)[37] and any step disruption may lead to a state of increased risk of thrombus formation, known as **thrombophilia** or **hypercoagulability**, a common complex CVD, in which the enhanced abnormal blood coagulation (hypercoagulability status) causes an increased disposition to the formation of obstructive (arterial and venous) clots[9,34,43,44]. Thus the circulatory system integrity will depend on the balance between all mechanisms (procoagulant, anticoagulant, and fibrinolytic) implicated in clot formation, which must work together in a harmonized way to enable the coordinated activity of all platelets, vascular and plasma factors[35,39,45].

1.2.1. Primary Hemostasis: The Vascular System and Platelet's Response to Injury

The vascular system has anticoagulant properties that aim to prevent the initiation (and propagation) of the coagulation process. The blood vessels' inside layer consists of **endothelial cells (ECs)** supported by a collagen-rich basement membrane and surrounding layers of connective tissue, forming a smooth unbroken surface to allow the blood fluid passage and prevent turbulence that, otherwise, could trigger the activation of platelets and plasma proteins[35]. The healthy intact ECs are naturally seen as antithrombotic as the endothelium shields circulating blood from thrombogenic subendothelial components (collagen, TF, …) and repel platelets[27,37]. However, the antithrombotic capabilities of ECs decrease if the vascular endothelium is disrupted, with exposure of the underlying subendothelial components, that endorse the activation of vascular wall repair (hemostatic) mechanisms as part of the healing process[27,37,40,41]. The first reply to blood vessels' damage is the local vasoconstriction (also known as vascular spasm), which decreases the blood flow to minimize major further blood losses.

This instant post-injury contraction quickly drops and platelets will interact with the injured vessel wall elements to form a plug to block the gap and close the wound[35,37,39,40]. As soon as the blood-flow is allowed to increase, the plug might be rapidly sheared from the surface[39]. The platelet plug is formed by adhesion of platelets to the exposed subendothelial collagen and to each other (platelet aggregation) and must be then stabilized by fibrinogen. This platelet plug, which is the endpoint of the primary hemostatic mechanisms, is however uncapable to sustain long-term hemostasis and will only provide a temporary seal for the vessel's wound, which must be later reinforced by fibrin (**Figure 2**)[35,37,46].

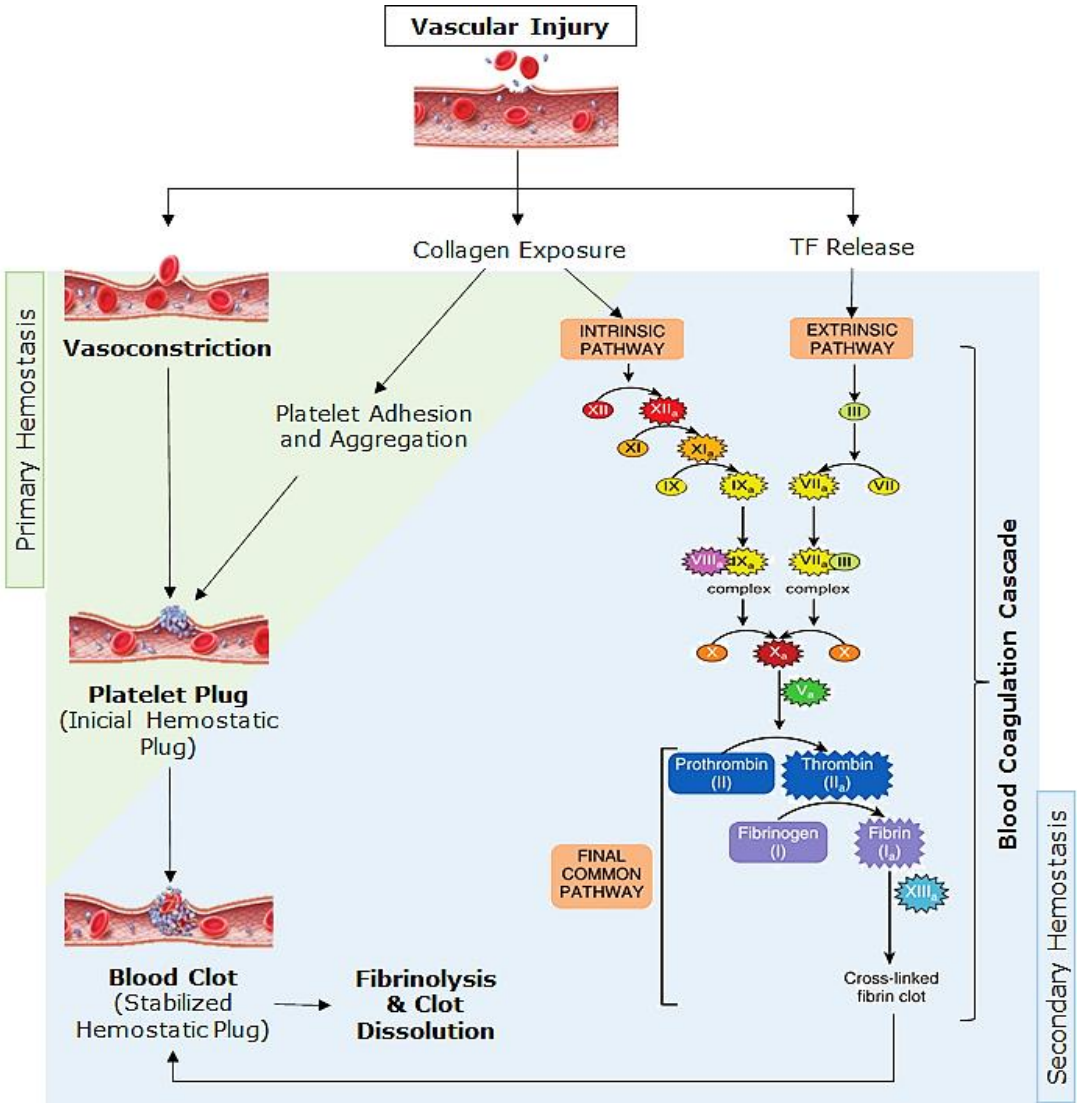


Figure resources: "Human Biology: Concepts and Current Issues", 7th Edition, Michael D. Johnson, West Virginia University©2014|Pearson; "Anatomy & Physiology", by OpenStax College. (https://openstax.org/books/anatomy-and-physiology/pages/18-5-hemostasis, 15 Jan 2020)

Figure 2| An overview of the hemostatic process. *Upon vascular injury, the primary hemostasis triggers vasoconstriction to avoid excessive bleeding. The exposure of collagen and TF mobilizes platelets to form the initial platelet plug (stabilized by fibrinogen) and triggers the coagulation cascade of secondary hemostasis, that generates the required fibrin to strengthen the final blood clot. The clot formation is restrained through antithrombotic control mechanisms (that inhibit coagulation) and, after fulfilling its role, it is dissolved by fibrinolysis.*

1.2.2. The Coagulation System Cascade of Secondary Hemostasis

Primary and secondary hemostasis are not independent from each other; instead there is an interaction between both processes: while the platelet plug is being formed, secondary hemostasis is also occurring[35,40,41,46]. Secondary hemostasis can be seen as *the process that forms a stable fibrin clot over the already-formed platelet plug*[37] and it will be the decisive step in the formation of the hemostatic plug, turning the liquid state of blood into a gel (solid state)[40,42]. Such mechanism involves a cascade of enzyme activation events between soluble plasma proteins, termed **coagulation factors** or **clotting factors (F)**, released into circulation in an inactive form and then activated during the **coagulation cascade** or **clotting cascade** chain reaction, in which each factor activates another[14,37,39,42]. This complex reaction cascade proceeds by two distinct routes, called the intrinsic and the extrinsic pathways, which converge into a final (common) pathway to generates thrombin and fibrin. Fibrin intercalates and binds to platelets, stabilizing and reinforcing the platelet aggregate plug, which results in the final fibrin clot – the secondary hemostasis outcome and the end product of the coagulation cascade (**Figure 2**)[35,40,41,46].

The **extrinsic (or TF) pathway** requires that plasma interacts with something ‘‘extrinsic’’, being triggered due to damage of the vessel’s outside tissue, by exposure of the cell-surface integral-membrane protein TF [14,40–42]. TF release (also known as FIII or tissue thromboplastin) constitutes the first vital event of blood coagulation upon vessel wall impairment. After FVII, activation, TF binds to small amounts of FVIIa. The TF-FVIIa complex is an extremely strong activator of coagulation and, once established, will catalyze the activation FX, that then assemble with FVa. The FXa-FVa complex (also known as “prothrombinase”) is required for the activation of prothrombin (FII) into thrombin (FIIa). Thrombin will allow the generation of fibrin that polymerizes, forming a gel and stabilizing the platelet plug[14,37,39,42]. Still, the TF pathway aims to generate the generate the necessary thrombin amounts to produce enough fibrin[39]. Conversely, the **intrinsic (or contact activation) pathway** is initiated by surface contact with the injured endothelium, through exposure of collagen to the circulating platelets. It causes FXII activation, which leads to the proteolytic activation of FXI, that in turn activates FIX[40–42,47]. FIXa combines with FVIIIa, originating the “tenase” complex (FIXa-FVIIIa), also required for the formation of thrombin, the linking step of the intrinsic and extrinsic pathways[39,41,42]. Although distinct, both the intrinsic and extrinsic pathways are highly connected, culminating in the activation of FX, that initiates the **common pathway**. FXa together with FVa (activated by thrombin), produces thrombin at the damaged site – the central event of secondary hemostasis[37,40–42]. Thrombin breaks soluble fibrinogen into

insoluble fibrin monomers and it also activates FXIII[14,37,40,41]. Fibrin monomers are initially held together by non-covalent interactions, but FXIIIa (the fibrin stabilizing factor) will catalyze the formation of covalent cross-links that turn fibrin into the meshwork that strengthens the final clot[39,40,42].

1.2.3. Inhibition of Coagulation and Clot Dissolution (Fibrinolysis)

As fibrin is being cross-linked, several interactions aim to restrain clot formation to the damaged area, limiting its extent propagation[42]. The coagulation process is carefully monitored by natural anticoagulant proteins, like **protein C** (**PC**; in its active form, **activated protein C; APC**), **protein S (PS)** and **antithrombin (AT)**[34,48]. As a powerful platelet activator, thrombin contributes to the formation of a protective hemostatic plug under normal hemostatic conditions; however, in case of pathologic clotting activation, it may lead to the development of a thrombus[14]. Thrombin can lose its procoagulant properties and activate PC. APC complexes with its cofactor PS to inactivate the clotting cascade factors Va and VIIIa[37,48–50]. AT is the most abundant and important inhibitor of clotting[37]. It is not only the primary physiological inhibitor of the procoagulant factors thrombin (FIIa, the main target) and Xa but also, in less extent, of other activated serine proteases (factors IXa, XIa, XIIa and XIIIa), quickly clearing them from circulation [13,49,51,52]. Meanwhile, as the wound is revascularized and healing is occurring, the fibrin clot is degraded by activation of the fibrinolytic system. Fibrinolysis is the pro-hemorrhagic event of fibrin clot dissolution, that aims to *discard fibrin after its hemostatic function has been fulfilled*. Such mechanism is arbitrated by plasmin, which digests soluble fibrin[35,37,42,48]. Similar to all hemostatic mechanisms, the fibrinolytic response is also under control of several enzymes, which include the plasminogen activators and the **plasminogen activator inhibitor-I (PAI-I)**, respectively required for the activation and inhibition of the fibrinolytic system, that is by catalysis or inhibition of plasmin conversion reaction from plasminogen (**Figure 3**)[53,54].

Under normal conditions, procoagulant, anticoagulant and fibrinolytic pathways regulate hemostasis to avoid pathological clot formation[13]. It is the subtle balance between coagulation and fibrinolysis (clot production and clot dissolution) that will establish the true stability of the fibrin clot[24,35,39]. However, due to the complexity of all the clotting cascade reactions and the multiple interactions of the factors involved, there is a considerable potential for dysfunction that can occur at any stage of the coagulation process[36]. Hemostasis can therefore get out of balance as a result of enhanced levels of clotting factors, anomalies in the anticoagulation factors and reduced fibrinolytic capacity that increases thrombosis susceptibility[24,35,39,55].

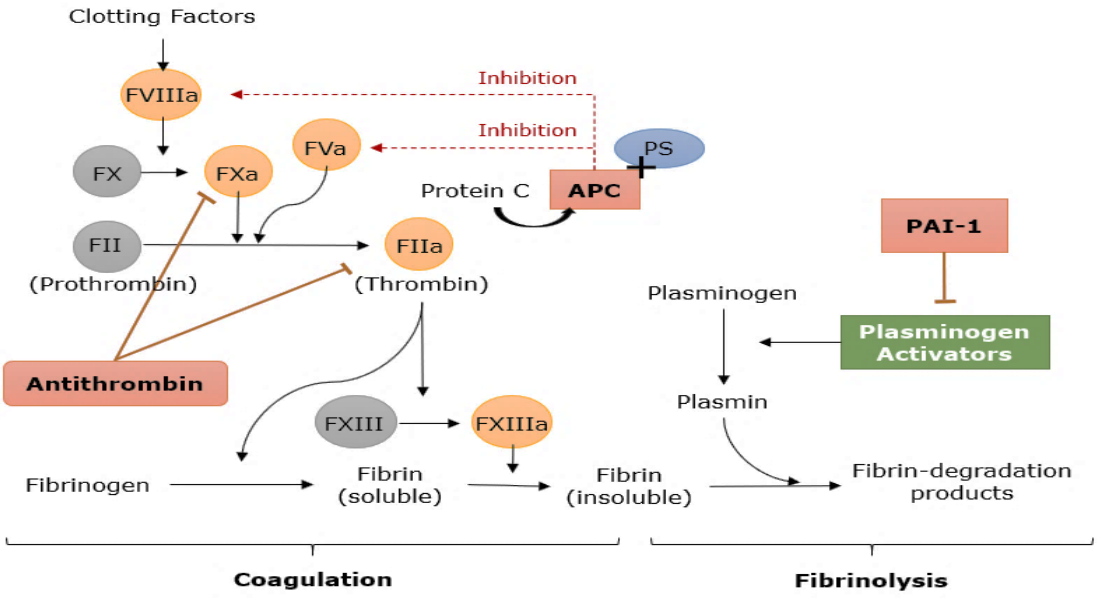


Figure 3| Antithrombotic Mechanisms of Hemostasis*. Thrombin can activate protein C to limit the extent of coagulation. APC will complex with its cofactor PS to inactivate factors Va and VIIIa. Antithrombin inhibits FXa and specially thrombin. Clot dissolution occurs through fibrinolysis, in which plasmin (converted from plasminogen, in a reaction catalyzed by plasminogen activators) digest fibrin. Such process is downregulated by the PAI-1 enzyme, that inhibits the plasminogen activators' action and, consequently, the fibrinolytic response.*

1.3. Risk Factors for Thrombosis Emergence

Numerous modifiable and nonmodifiable risk factors were conventionally recognized as increasing risks for the development of many NCDs. The reduction of modifiable risk factors, instigated by unhealthy and sedentary lifestyles (the main targets for preventive and management strategies) is crucial to control NCDs[3,4,6,7,56]. However, in an attempt to explain thrombophilia in patients without recognized cause and in families with thrombotic tendencies, genetic abnormalities in the coagulation pathway were searched[48]. Through candidate genes sequencing, rare genetic variants associated to increased thrombotic incidence were indeed identified[32,43,48]. These variants had a functional role and influence either in gene expression or in protein activity, thereby respectively interfering in a quantitative or qualitative manner with the hemostatic mechanisms[57]. However, those thrombogenic genetic factors seem to be only mildly increasing risks and they exhibit significant differences regarding geographic and racial distribution. Moreover, several VTE cases and thrombophilic families do not reveal recognized genetic-associated risks. Such findings led to the recognition of thrombophilia as a complex and multifactorial trait, relying on a complex interplay between both genetic and environmental risk factors; additionally, requiring the presence of 2 or more of those factors for its onset[57–60].

1.3.1. Inherited Hypercoagulable States: The Genetics of Thrombosis

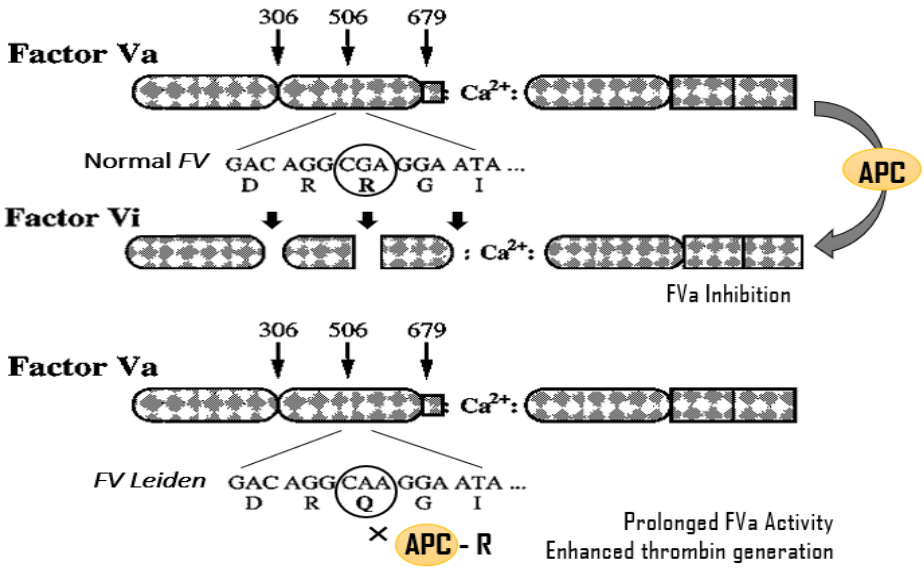
As with all body systems, the coagulation pathway is under homeostatic regulation to maintain the balance between the procoagulant and natural anticoagulant mechanisms. Namely, between clotting at the site of injury (confining the clot near the damaged vessel wall) and the maintenance of the blood fluid state within the vascular compartment. This way, the entire complex blood coagulation process is dictated at several steps by different natural inhibitors, that hamper excessive and uncontrolled clot formation[25,37,38,43]. Any inadvertent clotting either due to gain-of-function mutations of some coagulation factors, as well as inherited abnormalities of the natural inhibitors can disrupt the previous described equilibrium, increasing the procoagulant activity (hyperactive clotting) and decreasing the inhibitors activity (hypoactive anticoagulation), leading to an hypercoagulable state. Additionally, thrombosis can also arise due to hypoactive fibrinolysis[13,25,34,61,62]. Changes in the genes of proteins involved in all these pathways are crucial for thrombosis predisposition[62] and were determined by genetic association studies to be associated to an increased risk of both venous and arterial thrombosis. These variants occur in the genes coding for FV, FII (prothrombin), AT, PAI-1 and the methylene tetrahydrofolate reductase (MTHFR) enzyme, which are directly or indirectly involved in hemostatic regulation[10,13,57]. Many symptomatic patients are indeed carriers of more than one inherited variant regarding those genes (genetic risk factors), which supports the polygenic clinical phenotype of thrombophilia[59].

1.3.1.1. Gain of Function Mutations in Procoagulant Proteins: the FV Leiden and the Prothrombin Variants

Inherited thrombophilia can be caused by several genic factors known to confer a significant prothrombotic phenotype. Among all genetic factors, the discovery of two main inherited defects, a point mutation in the *FV* gene, the *FV* G1691A variant (also known as ***FV Leiden*, *FVL***), and the identification of the *FII (prothrombin)* G20210A variant, markedly change the molecular view of venous thrombosis[23,57,59]. They have emerged as the more frequently found in VTE cases[16,23,60,63] and are among the most often investigated as putative inherited risk factors. Besides, being quite relevant in VTE cases, they do not seem however to be major risk factors for MI or IS, increasing the risk for arterial thrombosis in a less extent[22,25,64].

FV is a blood protein required for normal blood coagulation, thus directly involved in hemostatic regulation and in the development of several complex chronic metabolic

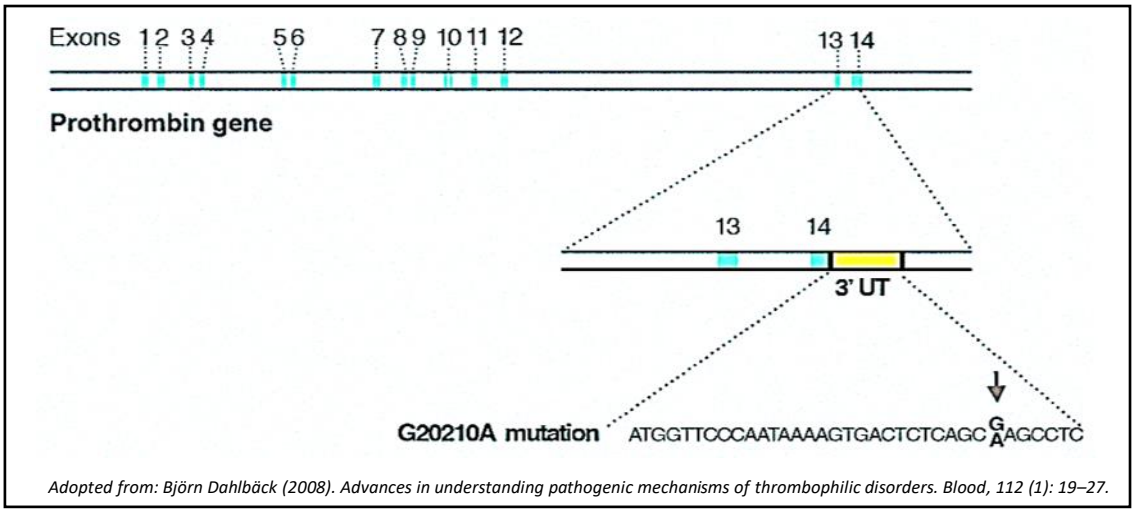
disorders, such as CVDs[10]. FV activity is downregulated by the PC and PS anticoagulant system, a major defense approach against thrombosis. After activated, PC will exert its anticoagulant effect through proteolysis of FV specific arginine bonds. However, changes in the gene encoding for FV (located on 1q23) may generate resistance to the PC-induced proteolysis, impeding its degradation and triggering clot formation. This condition is known as **activated PC resistance (APC-R)** [10,50,65,66] and it is most commonly caused by the *FVL*[50,66]. *FVL* is an autosomal dominant inherited defect[18] in which a guanine (G) is replaced by an adenine (A) at the nucleotide position 1691 of the *FV* gene exon 10, ensuing in a missense mutation that causes poor anticoagulant response to APC. To inactivate FVa, APC needs to recognize an arginine residue at position 506 of the FV molecule, one of the APC cleavage sites[34,55,64]. Though, because of the *FVL* defect, the codon (CGA), coding for the normally present arginine residue, turns into a different codon (CAA) that is translated as glutamine at the 506-cleavage site (R506Q). As a result, APC can no longer inactivate FVa, which becomes resistant to degradation and retains its coagulation abilities (**Figure 4**). Consequently, carriers of this defect are prone to develop hypercoagulability, that may clinically manifest as VTE episodes[23,34]. Heterozygous carriers of the *FVL* mutation are believed to have 5- to 10-fold greater risk of thrombosis, whereas homozygous subjects may be at 50- to 100-fold increased risk[55], with the majority expected to almost certainly develop a thrombotic event over their lifetime[43].



Adapted from: Dahlbäck, B., Hillarp, A., Rosen, S. et al. (1996). "Resistance to activated protein C, the FV:Q506 allele, and venous thrombosis". Annals of Hematology 72, 166–176.

Figure 4| The FVL impairment on the APC inhibitory action. *FVa is inactivated through cleavage of specific arginine bonds by APC, which inhibits its procoagulant activity. However, the presence of the FVL mutation induces resistance to APC (APC-R), as it eliminates one of APC cleavage sites (Arg506), impairing the degradation of FVa and consequently leading to enhanced thrombin generation.*

The *FII* G20210A variant is also inherited in an autosomal dominant manner[67] and it consists in a G to A change at position 20210 of the 3' untranslated region (UTR) of the gene (11p11-q12), that alters mRNA stability (**Figure 5**). As both the 5'-UTR and the 3'-UTR may play regulatory roles in the gene expression, this variant is associated to increased plasma levels (overproduction) of prothrombin[16,23,55,62,64]. Prothrombin is the precursor of thrombin (FIIa), the final effector of the clotting cascade and a key enzyme in hemostasis and thrombosis[50,68]. Moreover, despite being involved in the path of fibrin generation, prothrombin is also involved on platelet activation[68]. Thus, regarding its procoagulant role, prothrombin-related thrombophilia is a significant predisposing factor to VTE[69]: While the heterozygous genotype of the *FII* G20210A variant increases 2 to 3-fold the risk of VTE occurrence, the homozygous genotype may enhance this risk by 5 to 10-fold[18,70]. Moreover, it is associated with an almost 6-fold greater probability of recurrence[25].



Adopted from: Björn Dahlbäck (2008). Advances in understanding pathogenic mechanisms of thrombophilic disorders. Blood, 112 (1): 19–27.

Figure 5| The Prothrombin (FII) G20210A variant. *A nucleotide change (G to A) at position 20210 of the 3' UT region of the prothrombin gene will lead to increased production of prothrombin.*

1.3.1.2. Antithrombin Deficiency: An Anticoagulant Loss of Function

Antithrombin (AT) is a main endogenous circulating anticoagulant, that belongs to the serpin (serine protease inhibitors) superfamily. Under normal conditions, AT is in a repressed state, displaying weak and slow anticoagulant activity[71,72]. To help control coagulation and restrain it to the damaged endothelium boundaries, AT combines with heparin or heparan sulfate cofactors, substantially increasing (≈1000-fold) its inhibitory capacity, by turning into a fast and efficient inhibitor towards its target proteases [13,37,72]. However, rare genetic defects in the gene encoding AT (*SERPINC1*, located on the chromosome band 1q23–25.1) can make this inhibitor prone to structural and functional impairments that compromise hemostasis[13,73].

Inherited in an autosomal dominant manner, hereditary AT deficiency is a rare disease caused by various alterations in the *SERPINC1* gene[45,52,71]. It was the first thrombophilic disorder found[74] and despite its clinical heterogeneity, it is one of the severest risk factors for VTE[52,75]. While significantly reduced AT anticoagulant capacity is known to cause severe thrombophilia, the complete deficiency of this inhibitor might be incompatible with life, being associated to embryonic lethality. For this reason, AT deficient subjects are generally heterozygous and tend to suffer from recurrent thromboembolic episodes since young age. The rare cases of homozygous AT deficiency can only be caused by milder defects responsible for only slight impairments on the AT inhibitory capacity or on heparin binding[13,73,75,76]. Subjects with this pathology can be at 5- to 50-fold increased risk of VTE[52], however risk extension will therefore be highly dependent on the type of mutation that causes the inhibitory deficit[51]. Due to the AT important physiological function in keeping the balance between coagulation and anticoagulation (essential for thrombosis prevention), even modest deficiency of this inhibitor activity, can significantly increase the risk of both first and recurrent thrombotic events[52,58].

Since the first AT mutation was characterized, more than 220 distinct *SERPINC1* variants were found in inherited thrombophilia patients[73,75]. The most prevalent AT mutation is the *AT Cambridge II*, a dysfunctional variant located in the region coding for the AT molecule reactive site, which is described as a single point nucleotide substitution G>T at the position 13268 of the *SERPINC1*[75,77–79]. Such substitution results in the replacement of the normal alanine residue by a serine, causing the synthesis of a dysfunctional AT, with normal cofactor affinity but with borderline or mildly reduced inhibitory activity. Consequently, AT is unable to efficiently inhibit thrombin, increasing the risk of venous and arterial thrombosis[75,77–80]. Subjects with this genetic defect have a 6- to 10-fold increased risk of MI[52] and are at major risk (about 9-fold) of developing VTE[52,74,81]. Moreover, the combination of AT deficiency with other prothrombotic genetic risk factors, namely the *FVL* and *prothrombin* variants, is not an unusual phenomenon and it is a greatly contributor for increasing the risk of thrombosis[75].

1.3.1.3. Plasminogen Activator Inhibitor-I Impaired Fibrinolysis

The fibrinolytic system is a key player in the hemostasis maintenance, as fibrinolysis allows the dissolvement of the blood clots formed during tissue injury[82]. So, variants in factors involved in this system modulation, more precisely those associated with defective fibrinolysis, are believed to propense to clot formation. This is the case of those that enhance the accumulation of endogenous fibrinolytic inhibitors, like the PAI-

1[43,57,83]. PAI-1 is a glycoprotein member of the serpin superfamily of protease inhibitors, synthesized and secreted by ECs and platelets, under both physiological conditions or in response to stress[29,53,54,84]. It is the main biological inhibitor of the fibrinolytic system, quickly restraining fibrinolysis by direct inhibition of the plasminogen activators and, thereby, stopping plasminogen conversion, that is, plasmin formation. Due to its downregulating role in fibrin clots breakdown, PAI-1 has not only a central part in the negative regulation of fibrinolysis but it is also associated to thrombosis[29,54,85–87]. While the blockage of PAI-1 can promote clot degradation, its activation can result in the formation of a thrombus[86], establishing a clear association between PAI-1 and the pathogenesis and progression of various thrombotic vascular events[87]. Enhanced PAI-1 levels, reflected by increased enzyme activity, are a leader cause of fibrinolytic activity impairments (hypofibrinolysis), which appears to compromise the normal fibrin clearance mechanisms[53,82,88] and prone (at least in part) to vascular thrombotic events[53,54].

Several variants within the *PAI-1* gene (*SERPINE1*, 7q21.3–q22) have been proposed to modulate the expression of PAI-1. Among them, one major variant assumed to increase the risk of thrombosis (and the most frequently described and widely studied) is a single G deletion/insertion 675 bases upstream the gene transcriptional start site (i.e. in the promoter region), resulting in alleles with either a 4G or 5G genotype[53,54,82,83,89]. Although both 4G and 5G alleles bind with the transcriptional activator, only the 5G allele can also bind to a repressor protein at the overlapping binding site, reducing transcription and lowering the content of circulating PAI-1[84,90]. Conversely, as the 4G allele is lacking the binding site for the repressor protein, only the transcriptional activator binds, leading to higher gene transcription and elevated PAI-1 plasma levels (**Figure 6**). This way, in opposition to the normal PAI-1 expression

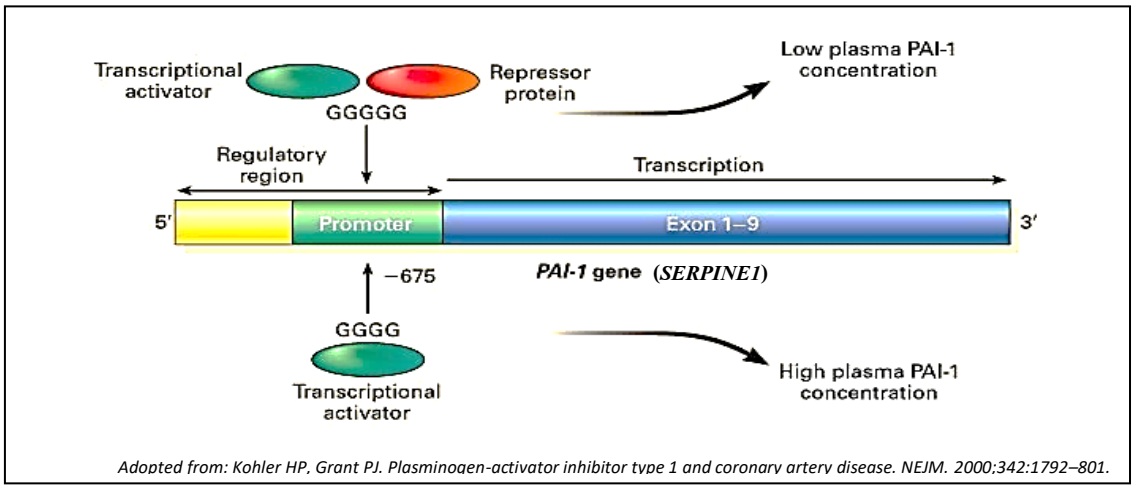


Figure 6| The modulation of PAI-1 expression by the PAI-1 4G/5G genetic variant. *While a 5G allele allows both the transcriptional activator and the repressor protein binding, only the transcriptional activator binds to the 4G allele, leading to higher gene transcription and elevated PAI-1 plasma levels.*

conferred by the 5G allele, a 4G allele expresses almost 6 times more messenger RNA[87,89,90]. Individuals with the 5G/5G genotype (wild-type subjects) are those with the lowest PAI-1 levels, whereas heterozygous (4G/5G) carriers display intermediate levels and homozygous carriers (4G/4G) have the highest PAI-1 concentrations, approximately 25% higher than the wild-type subjects[54,87,88,91]. Moreover, the 4G/4G genotype carriers also present greater procoagulant activity (including fibrin formation), when compared with subjects with 1 or 2 copies of the *5G* allele[53,82,84].

1.3.1.4. Hyperhomocysteinemia due to MTHFR Variants

Homocysteine (Hcy) is a non-essential sulfhydryl and non-protein forming amino acid, produced in all cells by biochemical synthesis, through the metabolism of the essential (dietary) amino acid methionine into cysteine[25,43,92–94]. Besides, being extremely important for the cell's homeostasis[94], Hcy is also toxic as it can cause vessels damage, induce inflammation and disturb liver detoxification, affecting health through several disorders and metabolic pathologies[95,96]. Increased Hcy levels are a mild risk factor for thrombosis, being specially associated with a higher risk to the incidence of arterial thrombosis, but also of VTE[25,93]. Such levels can be raised not only due to Hcy increased production but also due to its reduced metabolism[93], which can occur through a re-methylation process to convert Hcy into methionine again[92]. The metabolic Hcy re-methylation relies on a rate-limiting key enzyme – the **5,10-methylene tetrahydrofolate reductase (MTHFR)**, an important catalyzing agent in the vitamin folate metabolism[97–99]: After folate (vitamin B9) reduction to **tetrahydrofolate (THF)**, THF is converted to **5,10-methyltetrahydrofolate (5,10-MTHF)**; the MTHFR enzyme will then catalytically reduce 5,10-MTHF to **5-methyltetrahydrofolate (5-MTHF)** – a crucial step for Hcy recycling as, via **methionine synthase (MS)** action in the presence of the essential Vitamin B12 cofactor, the 5-MTHF substrate will function as a methyl group donor for Hcy conversion into methionine (**Figure 7**)[92,95,96,98].

Hyperhomocysteinemia (HHcy) is the state of abnormally increased Hcy plasma concentrations and it has been identified as a widely recognized risk factor and strong contributor for several CVDs. Prolonged HHcy promotes atherosclerotic plaques' formation, atherothrombotic events (through endothelial dysfunction and vascular damage) and enhanced inflammation, leading to the so-called thrombophilic profile[92,94,99]. HHcy can be caused by genetic mutations in the *MTHFR* gene (1p36.22), that alter or decrease the MTHFR enzyme activity. The most frequent (and possibly the most deleterious) are the *MTHFR* C677T (rs1801133) and A1298C (rs1801131) variants. Both these enzymatic isoforms are correlated with changes in the MTHFR

activity, being well-known to reduce the 5-MTHF formation (that can reach up to 75% less) and so limiting Hcy conversion back to methionine, causing its accumulation in the blood instead[95–97,100].

The *MTHFR* C677T variant is extremely common and prevalent[16,43,97] and it displays the more pronounced effects, causing worse handicaps[95]. It consists in a point mutation in which a cytosine (C) is replaced by a thymine (T) and it occurs in the coding region of the *MTHFR* exon 4 (c.665C>T), resulting in a highly conserved residue exchange of alanine into valine at the N-terminal catalytic domain region of the enzyme (p.Ala222Val)[22,97,98,100]. The presence of the T-allele results in the synthesis of a more thermolabile and less active form of the MTHFR enzyme, which is associated with <30% specific activity and consequently with a higher level of Hcy[22,43,97,98]. Compared to the normal functional enzyme (CC genotype), heterozygous carriers for the *MTHFR* C677T variant (CT genotype) display less 30-40% enzyme activity, whereas homozygous carriers (TT genotype) have 60-70% decreased MTHFR activity[98,100,101]. The *MTHFR* A1298C variant is also associated with decreased MTHFR activity. In this case, the A is replaced by a C in the *MTHFR* exon 7 (c.1286A>C), causing an amino acid change, from glutamine to alanine, within the C-terminal regulatory domain of the protein (p.Glu429Ala)[23,97,100]. The reduced MTHFR activity is also more pronounced in the homozygous than in the heterozygous state: while heterozygous carriers (AC genotype) have a 20-30% reduction of the enzyme activity, homozygous (CC genotype) exhibit a 40-50% reduction[97,100]. The coexistence of 677TT/1298AA and 677CT/1298AC genotypes lowers MTHFR activity to a great extent, by 60-70% and 40-50%, respectively[100].

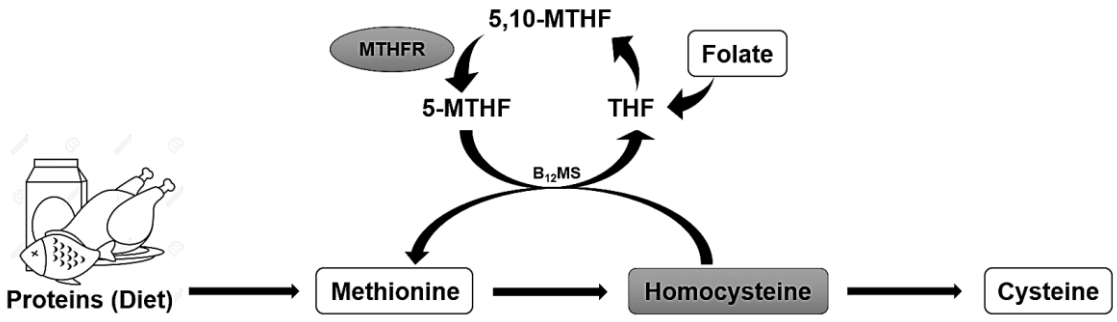


Figure 7| The process of homocysteine formation and the role of the MTHFR enzyme in the recycling of this aminoacid. *The metabolic conversion reaction of the dietary amino acid methionine into cysteine leads to the production of the homocysteine (Hcy) intermediate product. Hcy can be recycled through the folate metabolism, in which THF is converted to 5,10-MTHF. The MTHFR enzyme will then catalytically reduce 5,10-MTHF to 5-MTHF, that acts as a methyl group donor for Hcy re-methylation into methionine again. This last reaction requires the action of the methionine synthase enzyme in the presence of the cofactor vitamin B12 (B_{12}MS).*

1.3.2. Acquired Hypercoagulable States: Other Non-Modifiable and Modifiable Thrombotic Risk Factors

In the past, it was already known that individuals with inherited defects would have increased risks of facing a thrombotic event, if the inherited defect was compounded by a predisposing factor[43] or by exposure to high-risk conditions[31]. The same way that gene–gene interactions result in a synergistic increased effect of thrombosis development, also gene–environment interactions are capable of increasing the overall risk above the sum of individual thrombogenic risk factors[31,102]. Thrombosis is thus a multifactorial disease caused by the interplay of inherited (genetic) and acquired (environmental) risk factors, that may coexist in the same patient[18,22,103–105]. Besides genetic predisposition, multiple factors such as age, cholesterol levels, hypertension, obesity, immobilization, major recent surgeries, cigarette smoking, use of oral contraceptives, cancer, diabetes, inflammations (among others) and acquired hypercoagulable conditions, like pregnancy and puerperium, contribute to raise the incidence of thrombosis and synergistically enhance its risk[22,25,34,55,105,106].

1.3.2.1. Cigarette Smoking: The Hemostatic Systems' Effects

Ranked as the single greatest preventable cause of morbidity and mortality[107], **cigarette smoking (CS)** is an important risk factor for CVDs[108], with enormous health implications at a population level[107,109]. Even though its role as an independent risk factor (or effect modifier) for VTE remains controversial[17,25,110], CS is believed to increase the risk of arterial thrombosis[109], being a risk predictor for acute MI and also linked to increased risk of IS[17,25,110,111]. The mechanisms by which the atherothrombotic risks are increased are however not entirely understood[27]. CS is supposed to affect the main hemostatic systems by interfering with the functions of ECs, platelets, and coagulation factors[108,109]. Active smokers may indeed display significant individual shifts at the coagulation system level, when compared with lifetime nonsmokers[108]. Additionally, it is also suggested that it interferes with the antithrombotic and fibrinolytic functions of the blood and blood vessel walls, which also might be the main cause of most smokers' vascular events[27]. Nevertheless, the underlying pathological mechanisms of CS are still far from being fully understood[108], as the available data regarding its effects on thrombohemostatic and fibrinolytic factors is quite controversial[27]. Among all hypothetical reasons, the increased thrombotic risk conferred by CS are believed to be related to oxidative damage, endothelial dysfunction, increased release of TF and platelet aggregation[27,109]: CS may induce endothelial dysfunction

mechanisms, which activate platelet interaction, leading to the release of TF (and activation of the coagulation cascade), which culminates in the generation of thrombin and ultimately leads to clots' formation[27]. Likewise, there were documented CS effects on hemostasis regarding smokers increased plasma fibrinogen levels when compared to non-smokers[109].

1.3.2.2. Pregnancy and its Thrombotic Outcome Complications

Pregnancy by itself is an acquired and independent risk factor for the development VTE, which is a leading cause of maternal morbidity and mortality. About 1 in each 1000 pregnancies suffer complications due to VTE episodes[25,93,104,112,113]. The risk of such occurrence is in fact 5 times higher in pregnant women when compared to non-pregnant women of similar age[104,112]. During pregnancy, the coagulation system undergoes significant changes to help maintain the placental function and minimize blood loss at delivery. While the circulating levels of procoagulant factors increase, some natural anticoagulant pathways are repressed. Due to these changes, pregnant women develop acquired APC-R, partially related to reduced PS levels, which enhances thrombin generation and leads to a hypercoagulability state that may predispose women to thrombosis and placental vascular complications[25,89,93].

Recurrent **spontaneous abortion (SA)** refers to the loss of a fetus before 20 weeks of gestation in at least 2 consecutive pregnancies, a condition identified in about 2% of women, that remains unknown in over 50% of cases[54,89,114]. Its pathophysiology is not well understood, as its cause is complex and believed to be multifactorial, but it is particularly associated with coagulation and fibrinolytic factors[54,89]. In pregnancy, each factor causing an imbalance between coagulation and fibrinolysis may cause thrombosis and, consequently, recurrent SA episodes[89]. The placental system may thus be compromised by disturbances in hemostasis, mostly caused by maternal inherited thrombophilia[114], which is believed to increase up to 34-fold the risk of pregnancy associated VTE[113]. In fact, about half of pregnancy VTE episodes occur in women with thrombophilic defects[93]; the risk will however depend on the type of abnormality and in the presence of other risk determinants (maternal age, smoking, obesity, prior personal or family history of VTE, …) that significantly raise the risk of obstetric outcomes[31,93,104]. While women with deficiencies of AT, PC or PS are expected to be those at major risk, the FVL or prothrombin variants seem to be lower threats[93]. In contrast, the MTHFR enzyme is particularly important in pregnant women and must function properly, being required for implantation and invasion of the embryo[100,101]. Malfunctions regarding the folate cycle metabolism that result in enhanced Hcy levels (associated with endothelial

dysfunction) may cause placental vascular damage, which in turn is correlated with abortion, recurrent miscarriage and other obstetrical complications and outcomes of gestation, such as pregnancy diabetes[94,98–100]. Finally, the PAI-1 expression might also likely contribute to recurrent SA pathophysiology as PAI-1 can inhibit trophoblast invasion and maternal tissue remodeling[54].

1.3.2.3. The Impact of Exogenous Hormones

Combined oral contraceptives (COCs) are hormonal formulations of an estrogen, that acts to control withdrawal bleeding, associated with a progestogen to inhibit ovulation[33,106,115]. The use of such exogenous hormones, by COCs and also by **hormone therapy (HT)** (with estrogen and with or without a progestogen), for menopause and cancer treatments[31,115,116], is clearly established to be associated not only with an increased risk of VTE but also to the development of MI and IS events (arterial thrombosis)[18,33,93]. Noteworthily, arterial thrombosis is less frequent at reproductive age (1 case for every 5-10 cases of VTE)[33].

Until the end of 1995, it was believed that estrogens were the only hormones responsible for the increased incidence of thrombosis[33,36,115], as the incidence of VTE in young and otherwise healthy women emerged by the time that the oral estrogen compounds became available[31]. Estrogens are linked to changes in lipids and in lipoprotein metabolism and are believed to directly affect the vascular wall, change the factors that promote endothelial dysfunction and modestly modify the plasma levels of several coagulation factors[25,117]. They also lead to decreased levels of the PS and AT natural inhibitors, increased acquired APC-R and increased thrombin generation[33,115]. These had prompted the development of low-dose COCs, in which the estrogen content was progressively reduced in an attempt to diminish the associated risks[33,93,115]. However, even the low-estrogen formulations revealed to increase the incidence of thrombosis[115], being associated with 1.5 to more than 3-fold greater risk of incident VTE (compared with the risk of non-users in reproductive age)[18,93], which had led to acknowledge the also important role of progestogens on thrombosis effects modulation[36]. The COCs' progestogens have different systemic effects mediated by their binding ability (and affinity profile) to the progesterone and other steroid receptors, which will determine its thrombotic risk[33]. As the progestogen dose must remain unchanged for the effective inhibition of ovulation, the chemical composition of progestogens was altered trough time[115]: Compared with 2nd-generation progestogens, those of 3rd-generation COCs lead to higher levels of procoagulant factors (such as FVII and prothrombin), lower levels of natural anticoagulants' (especially PS), and a more

pronounced increase of APC-R, having a 2-fold higher thrombogenic associated risk[33,36,115]. In opposition, isolated progestogens do not seem to interfere in coagulation or fibrinolysis[33] and progestogen-only contraceptives have generally no or little effect on plasma lipoprotein levels and on the coagulation system, conferring a non-significant increase in VTE risk[25,33,106,117].

1.4. Objectives

Thrombotic diseases can lower the quality of life with late onset complications and even result in fatal outcomes. However, they are preventable[118]. The primary goals of public health initiatives are indeed the prediction and prevention of diseases, which is possible by identification of high-risk individuals[9]. If at-risk individuals were identified and managed prior to clinical manifestation, it would be possible to significantly reduce the negative impact of thrombosis, its incidence and high associated mortality and morbidity, as well as the substantial health care associated-expenses[59]. Genetic screenings, whether in asymptomatic individuals or in whose the risk is known to be increased (or who are at a specific phase of life that merits screening) allow to reveal those with predisposing genetic factors and predict the disease risk at the individual level[9]. Screening for genetic variants that predispose to thrombotic incidences is especially relevant for those with acquired and modifiable risk factors, who are already at risk due to non-genetic features and so may benefit most from preventive lifestyle strategies, medication or intervention approaches in the presence of the genetic thrombosis-related defects. As the prevalence of all inherited polymorphic alleles varies among populations[57], it is important to recognize the genetic profiling of each population, not only to assemble data for identification of high-risk individuals, but also to develop local guidelines for prevention. Hence, this project aims to:

- Estimate the prevalence of the *FV G1691A* (*FV Leiden*), *FII G20210A*, *AT Cambridge II*, *PAI-1 4G/5G, MTHFR C677T* and *MTHFR A1298C* thrombogenic variants in the Portuguese population;
- Establish associations between the genetic variants and acquired risk factors in the development of thrombotic-related events (stroke and myocardial infarction) and the occurrence of spontaneous abortion (in women);
- Identify groups of high-risk individuals in Portugal, according to the presence of the thrombogenic variants and the presence of additional increasing risk factors (obesity, hiperhomocysteinemia, cigarette smoking and the use of oral contraceptives or hormone therapy) in the population.

2. Methodology

2.1 Study Sample Design

Population sample is composed by participants from the "Prevalence of cardiovascular risk factors in the Portuguese Population" (e_COR) study, previous developed at the National Institute of Health Dr. Ricardo Jorge (INSA) and designed to characterize cardiovascular risk factors in the Portuguese population[119]. The e_COR studies comprises a stratified proportional sample of 1688 unrelated subjects, 848 men and 840 women aged between 18 and 79 years, from the 5 mainland regions of Portugal – North, Center, Lisbon, Alentejo and Algarve regions. From the total e_COR sample, 774 random Caucasian participants were selected for the present study to be representative of the population. Therefore, the sample dimension was defined according to the following assumptions: (1) The minimum sample size for obtaining national representativeness is 662 subjects (obtained with the program EpiInfo® version 3.5.1 of 2008) for a confidence interval of 99% and according to populational data of INE (Census 2011, Portugal National Institute of Statistics); (2) The most common risk factor MTHFR C677T (rs1801133) (47,5%), according to the previous reported by Branco *et al.* (2009)[120]; (3) The sampling error of 1%. The sample selected was tried to follow the demographic scheme and gender distribution of the adult population resident in Portugal in 2011 across those 5 main regions (Nomenclature of Territorial Units for Statistics II, 2011 Census, Statistics Portugal). Thus, from the total of the 774 participants designated, it was randomly selected 287 participants from North (138 males and 149 females), 180 from Center (85 males and 95 females), 216 from Lisbon (100 males and 116 females), 57 from Alentejo (28 males and 29 females) and finally 34 from Algarve (16 males and 18 females). Pregnant or postpartum women up to 3 months; inability to understand/speak Portuguese; declared mental disease; residence outside the study region and telephone contact failure after 3 different attempts at different days and hours were all exclusion criteria applied to the study sample. The e_COR study was approved by the National Data Protection Commission and National Institute of Health Ethic Committee and participants gave informed consent to each aspect of the study.

2.2 Data Collection

All participants were physically examined by measure of weight, height and waist circumference (digital scale SEC-899 and SEC-217 stadiometer, CEAS; flexible tape

SEC-201, SECA) and body mass index (BMI) was calculated as weight in kilograms (0.5 kg was removed from each participant) divided by squared height in meters (kg/m^2). Subjects were assigned as "underweight" if BMI < 18.5 kg/m^2, "normal (healthy) weight" if 18.5 ≤ BMI < 25 kg/m^2, "overweight" if 25 ≤ BMI < 30 kg/m^2 and, finally, BMI ≥ 30 kg/m^2 was considered as obese. Personal data and recent clinical information, which included the occurrence of the thrombotic-related events MI and stroke and smoking habits were also obtained from a questionnaire application. Women were also asked about the use of OCs and/or HT and the number of pregnancies and miscarriages.

2.3 Blood Collection, Sampling and Biochemical Analysis

After fasting for about 12 hours, approximately 16 mL of blood was withdrawn from each participant (1 serum gel tube of 7.5 mL, 3 *ethylenediamine tetraacetic acid* (EDTA) tubes of 2.7 ml) to determine the fasting biochemical profile and for DNA extraction. Blood was maintained at rest for between 30 minutes and 3 hours, and then centrifuged at 3,000 rpm for 15 minutes. After centrifugation, serum and plasma were stored at 2-8 °C controlled temperature, until transportation (in properly refrigerated conditions) to INSA, where samples were processed within 36 hours. All participants' biochemical parameters were determined at the Diagnosis and Reference Laboratory Unit (UDR) of INSA. Hcy levels were determined in the autoanalyzer Cobas Integra 400 plus (Roche, Risch-Rotkreuz, Switzerland) by an enzymatic method. Normal physiological values of blood serum Hcy are expected to range between 5 and 15 µmol/L, with fasting levels greater than 15 µmol/L considered as elevated and an indicative of clinically relevant HHcy. HHcy is considered as mild when Hcy values ranges from 15 to 30 µmol/L, moderate for values between 30 and 100 µmol/L and severe for values above 100 µmol/L[25,92,94]. The remaining EDTA-anticoagulated blood was stored at –20°C until DNA extraction for genetic analysis.

2.4 Genetic Analysis and Genotyping Assays

2.4.1 DNA Extraction

Genomic DNA from each participant was extracted from the EDTA peripheral blood leukocytes according to the salting out method described by Lahiri *et al.* (1991)[121], with some adaptations of the protocol. The whole blood previous collected in the EDTA tubes was homogenized and transferred into a 15 mL tube, where it was also added an

equal volume of salt solution, containing 10 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 2 mM EDTA and 25 mL of Triton X-100 per liter. For every milliliter of blood, it was also pipetted to the tubes 25 µL of IGEPAL® detergent (Sigma-Aldrich). Tubes were well mixed and centrifuged at 2200 rpm for 10 min at 18ºC (Centrifuge 5810-R, *Eppendorf*). Therefore, the supernatant was discarded and the nuclear pellet was washed with a low salt buffer containing 10 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$ and 2 mM EDTA (1mL for milliliter of blood), prior to a second centrifugation, this time at 1600 rpm for 10 min at 18ºC. This washing step was repeated, and, after discarding the supernatant, pellet was resuspended in a high salt buffer containing 10 mM Tris-HCl pH 7.6, 10 mM KCl, 10 mM $MgCl_2$, 0.4 M NaCl and 2 mM EDTA (160 µL for milliliter of blood). 10% SDS was also added to the tubes (10 µL for blood milliliter), which were then thoroughly mixed and incubated for 10 min at 55°C. All tube contents were transferred to a 2 mL microtube, following addition of a concentrated 5 M NaCl solution (60 µL for milliliter of blood), and centrifuged at 13200 rpm for 20 min in a microcentrifuge (Centrifuge 5415-R, *Eppendorf*) to precipitate the protein content. The supernatant (containing DNA) was pipetted to a new tube with 100% ethanol and tubes were slowly inverted till DNA precipitation becomes visible. The precipitated DNA strands were then captured from absolute ethanol and washed in 70% ethanol solution. DNA was dried at room air temperature and finally suspended in 150 µL of TE (10 mM Tris and 1 mM EDTA, pH 8.0) buffer (Ambion®) in a threaded tube. DNA tubes were preserved in 4ºC for a minimum of 24-48h until quantitative and qualitative analysis.

DNA concentration and quality were measured using NanoDrop™ 1000 Spectrophotometer (Thermo Fisher Scientific™) and, at the same time, samples were normalized (by TE addition) to final concentrations about 250-350 ng/µL. The DNAs' purity was evaluated by the optical density ratios A_{260}/A_{230} and A_{280}/A_{260} shown on nanodrop. The integrity of the DNAs was confirmed by UV fluorescence gel electrophoresis: 1% (w/V) agarose gel (Citomed), prepared in 100 mL of **Tris/Borate/EDTA (TBE)** buffer 1X (TBE 10X ultrapure, pH 8.4, 1.0M Tris, 0.9M Boric Acid, 0.01 M EDTA, Invitrogen) and stained with 2% SYBR™ Safe DNA Gel Stain (Invitrogen™). After polymerization of the gel (for at least 1h), samples were added to the gel in the following proportion: 1 µL of DNA, 5 µL of deposition solution and 5 µL of deionized water. The electrophoresis was run at 90V for 50 min (Bio-Rad Power Pac 3000 voltage source) in TBE 1X and the gel was therefore visualized on the Safe Imager™ 2.0 Blue-Light Transilluminator (Invitrogen™). Samples were stored at -4ºC until all study assays were performed.

2.4.2 TaqMan® OpenArray™ Genotyping

Participants were genotyped for the thrombogenic variants by the TaqMan® OpenArray™ Technology, which allows to identify slight variations in genes within a population, providing the analysis of many of those alterations across numerous samples at once. The TaqMan® OpenArray™ Genotyping System requires a specific TaqMan® OpenArray™ Genotyping Plate (Applied Biosystems™), a microscope slide–sized plate, with 48 subarrays, each with 64 through-holes (nano-wells), that comprises a total of 3072 through-holes preloaded with the assays of interest (pre-designed or customized). Each through-hole of the TaqMan® OpenArray™ Genotyping Plate is filled with approximately 33 nL of PCR reaction mixture of reagents (probes and primers), dried-down in the plate and retained via surface tension, due to the hydrophilic coating of the wells' interior *versus* the hydrophobic plate surface[122,123]. The genotyping plates used in this study had the standard layout of 32 variants for 96 samples and were ordered and customized for the assays of the intended thrombogenic variants, listed on **Table 1**.

All experimental parameters, such as the optimal primers for each assay, were designed by the manufacturer, according to the information provided upon request of the service. To run the plates, samples are mixed with the master mix (provided by manufacturer) and loaded with the automated QuantStudio 12K Flex OpenArray AccuFill™ System. Samples are amplified by PCR, with a common pair of primers to both wild-type and mutant sequences, and specific Minor Groove Binder (MGB) probes, that are labeled at the 5′ end with different fluorescent reporter dyes, one for each allele of the target SNP. The allele 1 probe (for the wild-type sequence) is labeled with the VIC® fluorophore and the allele 2 probe (for the "mutant" sequence) is labeled with the FAM®

Table 1| Genetic variants analyzed by the TaqMan® OpenArray™ Technology for this study.

Variant Common Name	Gene	dbSNP	Gene Position and Alleles (GRCh38.p12)	DNA Alteration	Protein Change
FV G1691A (*FV Leiden*)	*F5*	rs6025	chr1: 169549811 C>T	NM_000130.4: c.1601G>A	p.Arg534Gln
FII (Prothrombin) G20210A	*F2*	rs1799963	chr11: 46739505 G>A	NM_000506.3: c.*97G>A	NA
Antithrombin Cambridge II	*SERPINC1*	rs121909548	chr1: 173904038 C>A	NM_000488.3: c.1246G>T	p.Ala416Ser
PAI-1 4G/5G	*SERPINE1*	rs1799762	chr7: 101126429 dupG	NM_000602.4: c.-817dupG	NA
MTHFR C677T	*MTHFR*	rs1801133	chr1: 11796321 G>A	NM_005957.4: c.665C>T	p.Ala222Val
MTHFR A1298C	*MTHFR*	rs1801131	chr1: 11794419 T>G	NM_005957.4: c.1286A>C	p.Glu429Ala

PAI-1 - Plasminogen Activator Inhibitor type-1; MTHFR - Methylene Tetrahydrofolate Reductase dbSNP - Single Nucleotide Polymorphism Database; NA – not applicable.

fluorophore. During PCR, each probe anneals specifically to its complementary sequence (between the forward and reverse primer sites) and only those probes that hybridize to their specific allele (match) are cleaved by DNA polymerase. Thus, the detection of fluorescence signals generated by cleavage during PCR amplification allows to differentiate each allele and reveal which ones are present in the sample[122,123].

Samples had to simultaneously fulfill the DNA purity and concentration requirements for this technology assays. Only samples with DNA concentrations of ≥50ng/mL and with preferential purity of ≥1.7 (with the optimal $A_{260/280}$ and $A_{260/230}$ ratios comprised between 1.8-2 and 2-2.2, respectively) were selected, as previous accessed by NanoDrop® 1000 spectrophotometer (Thermo Fisher Scientific™). The integrity of the DNA was also previously evaluated by UV fluorescence gel electrophoresis (1% agarose gel, stained with 2% SYBR™ Safe). Samples were then diluted to final concentrations of 50 ng/µL (with a minimum DNA amount of 500 ng, as required for the TaqMan® OpenArray™ genotyping reactions) in 96-well plates. The final concentration of the normalized DNA samples was checked using nanodrop and then sent to *Fagron Genomics* company (Terrassa, Barcelona) for the OpenArray® screens. Genotyping was performed on the QuantStudio™ 12K Flex Real-Time PCR System instrument (Applied Biosystems™) following the manufacturer's instructions and results were analyzed with both the QuantStudio™ 12K Flex software and Taqman® Genotyper application (Thermo Fisher Scientific™) and sent back to INSA in CSV format. 16 participants of the study have already been previous genotyped and were used to validate the assays, by confirming the concordance of calls with the expected genotype. PCR-Restriction Fragment Length Polymorphism (RFLP) and Sanger sequencing were also performed and used to verify and revalidate the results.

2.4.3 Polymerase Chain Reaction – Restriction Fragment Length Polymorphism Assays

395 participants were also genotyped for the *FII* G20210A (rs1799963) and *MTHFR* A1298C (rs1801131) variants by PCR-RFLP. The PCR-RFLP details, as the specific primers and enzymes for both SNPs, the relevant amplified gene fragments and the subsequent digestion of the PCR products, are described in **Table 2**, according to the previous report by Szpecht *et al.* (2017)[65]. All PCR and restriction reactions were performed using the Biometra® T3000 Thermocycler (Analytik Jena AG), GeneAmp® PCR System 9700 thermocycler (Applied Biosystems™, CA, USA) and the Biometra®

Tprofessional Basic Gradient 96 (Analytik Jena AG). DNA samples from subjects whose genotype was already known were used as internal controls for enzymatic digestion.

The PCR amplification reactions were carried out (for both genes) in a final volume of 25 µL, with approximately 100-200 ng of genomic DNA, 0.2 mM of each desoxynucleotide triphosphate (dNTP) (dATP, dCTP, dGTP, dTTP; 100 mM, Bioline), 20 pmol of each of the forward (F) and reverse (R) primer (Invitrogen™), 0.75 mM $MgCl_2$ (Bioline), 10X PCR buffer (Bioline) and 1.25 U of DNA polymerase (Bioline). All amplifications were performed according to the following thermocycling program: initial denaturation of 3 minutes at 95ºC; followed by 35 cycles with denaturation at 94ºC for 45 seconds, annealing at 60ºC for 30 seconds and extension at 72ºC for 1 minute; a final extension of 5 minutes at 72ºC was performed. The PCR products were analyzed on a 2% (w/V) agarose gel (Citomed), prepared in 100 mL of TBE 1X (TBE 10X ultrapure, pH 8.4, 1.0M Tris, 0.9M Boric Acid, 0.01 M EDTA, Invitrogen™) and stained with 2% SYBR™ Safe DNA Gel Stain (Invitrogen™). After polymerization of the gel (for at least 1h), each sample was applied in the gel, by adding 5 µL of PCR product and 5 µL of deposition solution. A 100 bp DNA Ladder (Fermentas) was also applied on the gel. The electrophoresis was run at 90V for 40 min (Bio-Rad Power Pac 3000 voltage source) in TBE 1x. The agarose gel was visualized on the transilluminator (Safe Imager™ 2.0 Blue-Light Transilluminator, Invitrogen™).

The amplified PCR fragments for the *FII* G20210A and *MTHFR* A1298C variants were respectively digested with *HindIII* and *MboII* restriction enzymes, as listed in **Table 2**, for 4h at 37ºC. The digestion reactions were performed in final volumes of 50 µL, with 10 µL of the PCR products, 1 µL of enzyme (10 U), 5 µL of each provided specific enzyme buffer 1X and the remaining volume with deionized water. To inactivate the enzymes, samples were then maintained inside the thermocycler for an additional 20 min period at 80 °C (for *HindIII*) or at 65ºC (for *MboII*). A 4% (w/V) gel of a mixture of agaroses

Table 2| Primers and enzymes used in the PCR-RFLP assays and description of its products.

Variants	Sequence of Primers	Restriction enzyme	PCR product size (bp)	Restriction Products
FII G20210A	5′ TCT AGA AAC AGT TGC CTG GC 3′ 5′ ATA GCA CTG GGA GCATTG AAG C 3′	*HindIII* (New England Biolabs)	345	GG: 345 bp; GA: 345, 322, 23 bp; AA: 322, 23 bp
MTHFR A1298C	5′ CTT CTA CCT GAA GAG CAA GTC 3′ 5′ CAT GTC CAC AGC ATG GAG 3′	*MboII* (New England Biolabs)	256	AA: 176, 30, 28, 22 bp; AC: 204, 30, 28, 22 bp; CC: 204, 30, 22 bp

PCR – Polymerase Chain Reaction; RFLP – Restriction Fragment Length Polymorphism; bp – base pairs; MTHFR – Methylene Tetrahydrofolate Reductase

(Citomed and MetaPhor™ Lonza), at the same proportion, was prepared in 100 mL of TBE 1X (TBE 10X ultrapure, pH 8.4, 1.0M Tris, 0.9M Boric Acid, 0.01 M EDTA, Invitrogen) and stained with 2% SYBR™ Safe (Invitrogen™). After gel polymerization, 15 µL of the RFLP products along with 5 µL of deposition solution were added to the gel, as well as the 100 bp ladder (Fermentas). Electrophoresis migration occurred at 70V for about 2h (Bio-rad Power Pac 3000 voltage source) in TBE 1x. Afterwards, the agarose gels were visualized by UV fluorescence (Safe Imager™ 2.0 Blue-Light Transilluminator, Invitrogen™).

2.4.4 Sanger Sequencing

The DNA samples of 10 participants were also genotyped by Sanger sequencing for the *FII* G20210A (rs1799963) and *MTHFR* A1298C (rs1801131) variants, to verify the previous PCR-RFLP genotyping results. The PCR products were first purified, by elimination of the primers and dNTPs excess trough enzymatic digestion with a combination of two hydrolytic enzymes, the Exonuclease 1 and *Shrimp Alkaline Phosphatase* (SAP), combined in a commercial product termed ExoSAP-IT (Amersham Pharmacia Biotech). 1µL of ExoSAP-IT was added to 2.5 µL of the PCR product and kept for 15 minutes incubation at 37ºC, followed by 15 minutes at 80ºC, in the thermocycler (Biometra® T3000, Analytik Jena AG). Samples for sequencing reaction of each SNP were prepared according to the *Big Dye® Terminator Cycle Sequencing Ready Reaction Kit* (Applied Biosystems) in 10 µL volumes, with 2 pmol of the same forward primers used in the PCR-RFLP procedure, 2 µL of *Terminator Ready Reaction Mix* and 1 µL of the purified PCR product, following the described thermocycling protocol: 1 minute for 96ºC and then 25 cycles of 96ºC for 10 seconds, 58ºC for 5 seconds and 55ºC for 4 minutes. Samples were then kept at -4ºC until sequencing (3100 Genetic Analyser, Applied Biosystems) at the Genetics Department of INSA. The results were analyzed with the Staden Package software.

2.5 Statistical Analysis

Allelic and genotypic frequencies for each genetic variant (*FV G1691A*, *FII G20210A*, *AT Cambridge II*, *PAI-1 4G/5G, MTHFR C677T* and *MTHFR A1298C)* were determined by gene counting and expressed in percentual values (%) for the determination of the estimated prevalence of such variants in the Portuguese population. The 95% confidence intervals (CI) were also established for the variants' genotypic

frequencies and the genotypes' distribution was tested for **Hardy-Weinberg equilibrium (HWE)** by the chi-square test (χ^2). The critical χ^2 value at the 5% level of significance with one degree of freedom is 3.841, and so if $\chi^2 < 3.841$, the variant is considered as not deviated from the HWE[124,125]. Genotypic frequencies were also determined (by gene counting) among subgroups of the population for comparing possible variations. For these parallel analyses, the population was divided according to region (North, Center, Lisbon, Alentejo and Algarve), sex and age of the participants. Similarly, frequencies of thrombogenic allele carriers with other risk factors (obesity, HHcy, CS, OC and HT use) were also determined to assess groups of high-risk in the population.

The presence of several risk factors was correlated among the population and associations were made between the genotype data (thrombogenic variants) plus the clinical, biochemical and personal data available. The genotypes of participants who suffered or not thrombotic-related events (myocardial infarction and stroke) or SA (in the female population) were analyzed with or without the presence of other thrombotic-risk factors. As the evaluated group of women with SA was composed by more than 30 participants, it was additionally tested by the Pearson Chi-square test, carried out using SPSS software (version 25, SPSS Inc.). The hypothesis of equality of two proportions was tested against the alternative that they are not equal, with the p-value for the inclusion and exclusion of variables set at 0.05 (i.e. p-value < 0.05 was considered as statistically significant).

3. Results

3.1 Study Sample Characterization

The sample of 774 participants (367 males and 407 females) were grouped according to several parameters of thrombotic risk relevance (age, BMI, Hcy levels, CS habits, OC and HT use), as well as according to the incidence of thrombotic-related events (MI and stroke) and SA (in the female population) (**Table 3**). A major part of the

Table 3| Characteristics of the study sample participants.

	Male (N=367)		Female (N=407)		Total (N=774)	
	n	%	n	%	n	%
Age						
18-34	110	30.0	125	30.7	235	30.4
35-64	136	37.0	158	38.8	294	38.0
65-79	121	33.0	124	30.5	245	31.6
Body Mass Index						
Underweight	6	1.6	14	3.4	20	2.6
Normal (Healthy) Weight	110	30.0	142	34.9	252	32.6
Overweight	165	44.9	161	39.6	326	42.1
Obese	85	23.2	90	22.1	175	22.6
Without Info[a]	1	0.3	0	0.0	1	0.1
Homocysteine Levels						
Physiological Values	266	72.5	358	88.0	624	80.6
Mild Hyperhomocysteinemia	92	25.1	48	11.8	140	18.1
Moderate Hyperhomocysteinemia	9	2.4	1	0.2	10	1.3
Cigarette Smoking						
Yes	113	30.8	75	18.4	188	24.3
No	254	69.2	332	81.6	586	75.7
Oral Contraceptive Use						
Yes	-	-	120	29.5	120	15.5
Not anymore	-	-	183	45.0	183	23.6
Never	-	-	104	25.5	104	13.4
Hormone Therapy						
Yes	-	-	50	12.3	50	6.5
No	-	-	138	33.9	138	17.8
Unknown[b]	-	-	219	53.8	219	28.3
Myocardial Infarction						
Yes	12	3.3	7	1.7	19	2.5
No	355	96.7	399	98.0	754	97.4
Possible[c]	0	0.0	1	0.3	1	0.1
Stroke						
Yes	12	3.3	11	2.7	23	3.0
No	355	96.7	395	97.1	750	96.9
Possible[c]	0	0.0	1	0.2	1	0.1
Spontaneous Abortion						
Yes	-	-	44	10.8	44	5.7
No	-	-	363	89.2	363	46.9

(a) Weight undetermined due to participants' physical limitations.
(b) Participants were unaware of the answer or did not provide the information.
(c) *Participants may have suffered the event; however. it was not confirmed by diagnosis.*

study sample revealed to be overweight (42.1%) and obese (22.6%), while most of the participants (80.6%) displayed normal Hcy serum levels (the remaining displayed mild or moderate HHcy). There were no participants with severe HHcy and even the moderate condition was present in only a few individuals. It is worth noting that the higher incidence of HHcy was observed among male individuals: 92 males against 48 females displayed the mild condition, whereas 9 males *versus* only 1 female presented moderate HHcy. Men with CS habits were also more prevalent. Among women, 44 had experienced one or more SA.

3.2 Estimated Prevalence and Frequency Distributions of the Thrombogenic Variants in the Portuguese Population

The genotypes of all six thrombogenic variants evaluated in the study were determined for all the 774 participants. The genotypic and allelic frequencies of each variant were calculated (**Table 4**) to estimate the prevalence of these variants in the Portuguese population (**Figure 8**). In this population, all variants' genotypic distributions were consistent with the HWE (χ^2 <3.841).

Table 4| Genotypic and allelic frequencies of the thrombogenic variants among the sample.

	Genotype Frequency				Allelic Frequency			HWE
		n	%	CI 95%		n	%	χ^2
***FV G1691A* (*FV Leiden*)**	G/G	757	97.8	96.8 - 98.8	G	1531	98.9	0.095
	G/A	17	2.2	1.2 - 3.2	A	17	1.1	
	A/A	0	0	-				
FII G20210A	G/G	742	95.9	94.5 - 97.3	G	1516	97.9	0.345
	G/A	32	4.1	2.7 - 5.5	A	32	2.1	
	A/A	0	0.0	-				
AT Cambridge II	G/G	769	99.4	98.8 - 99.9	G	1543	99.7	0.008
	G/T	5	0.6	0.1 - 1.2	T	5	0.3	
	T/T	0	0.0	-				
PAI-1 4G/5G	5G/5G	215	27.8	24.6 - 30.9	5G	820	53.0	0.099
	4G/5G	390	50.4	46.9 - 53.9	4G	728	47.0	
	4G/4G	169	21.8	18.9 - 24.7				
MTHFR C677T	C/C	341	44.1	40.6 - 47.6	C	1016	65.6	1.460
	C/T	334	43.2	39.7 - 46.6	T	532	34.4	
	T/T	99	12.8	10.4 - 15.1				
MTHFR A1298C	A/A	392	50.6	47.1 - 54.2	A	1090	70.4	2.025
	A/C	306	39.5	36.1 - 43.0	C	458	29.6	
	C/C	76	9.8	7.7 - 11.9				

The genotypic and allelic frequencies of the genetic variants Factor V Leiden (FVL), Factor II (FII) G20210A, Antithrombin (AT) Cambridge II, Plasminogen Activator Inhibitor type-1 (PAI-1) 4G/5G and Methylene Tetrahydrofolate Reductase (MTHFR) C677T and A1298C are expressed in percentage (%), according to the number (n) of participants who display each variant respective genotype among the total sample (N=774). The confidence intervals (CI) of 95% determined for each genotype are also expressed in percentage (%). The chi-square (χ2) value for the Hardy-Weinberg equilibrium (HWE) is received for value of χ2 <3.84; p < 0.05 for one degree of freedom.

As seen in **Figure 8**, *FV*, *FII* and *AT* variants were rare among the sample individuals, which displayed a major frequency of the wild-type genotype for these genetic variations. In fact, 99.4%, 97.8% and 95.9% of subjects did not carry any pathogenic allele for the *AT*, *FV* and *FII* genetic variations respectively, and the ones that carried one pathogenic allele (heterozygosity) were a small minority (0.6%, 2.2% and 4.1%, for *AT*, *FV* and *FII*, respectively). None of the subjects exhibited the homozygous condition for these variants and the *AT* variant was indeed almost absent in the population. In contrast, the *PAI-1* and the *MTHFR* variants were widely distributed among the population. The heterozygous condition for those variants was extremely common and their frequency was not very different from the one observed for their wild-type genotypes. Additionally, the *PAI-1* homozygous genotype (4G/4G) was almost as frequent (21.8%) as the wild-type one (5G/5G). Regarding the *MTHFR* variants, the *MTHFR* C677T was the most prevalent, with most individuals carrying at least one pathogenic allele (43.2% CT and 12.8% TT genotypes against 44.1% CC wild-type genotype). Finally, half of the participants revealed to be wild-type subjects for the *MTHFR* A1298C (50.6%), followed by a smaller prevalence of the heterozygous condition (39.5%) and a much smaller fraction of homozygous (9.8%) (**Figure 8**).

The collected data was also weighted by geographic region and gender of the participants (**Figure 9)**. Regarding the distribution of the genotypic frequency across the continental regions of Portugal (**Figure 9A**), the Center appears to be the one with the greatest number of *FVL* (3.3%) and *FII* G20210A (5%) heterozygous carriers. None of the participants from Alentejo and Algarve regions presented the *FVL* and the *AT*

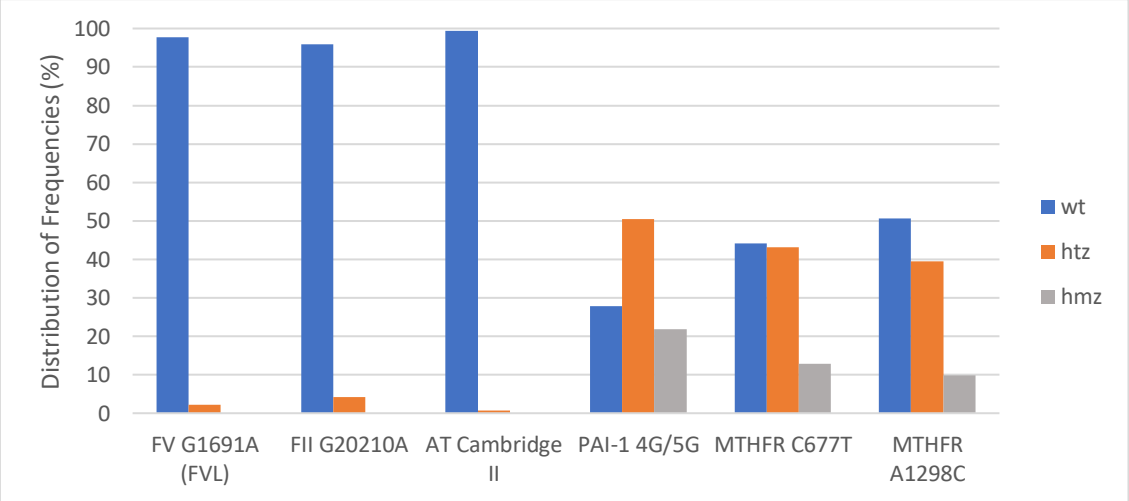


Figure 8| Estimated prevalence of the thrombogenic variants in the Portuguese population. *The estimated prevalence of the thrombogenic variants Factor V Leiden (FVL), Factor II (FII) G20210A, Antithrombin (AT) Cambridge II, Plasminogen Activator Inhibitor type-1 (PAI-1) 4G/5G and Methylene Tetrahydrofolate Reductase (MTHFR) C677T and A1298C are expressed in percentage (%) according to each variant respective wild-type (wt), heterozygous (htz) and homozygous (hmz) genotype.*

heterozygous genotypes; nevertheless, these regions had the lowest number of genotyped individuals. The *PAI-1* heterozygous genotype frequency was higher in Alentejo (57.9%), followed by the North (53.0%) and Center regions (53.3%). The homozygous frequency of this variant was quite similar to the one observed for the wild-type genotype in North, Center and Alentejo, but the Center displayed a higher frequency of homozygous individuals (23.9%). The *MTHFR* C677T revealed similar frequencies of heterozygous and wild-type genotypes across regions, with the Center and Lisbon displaying a slightly higher frequency of heterozygous than wild-type subjects. The homozygous frequency is higher in Center (17.2%) and smaller in Alentejo (8.8%) regions. Regarding the *MTHFR* A1298C variant: the frequency of wild-type individuals was higher, decreasing for the heterozygous and homozygous genotypes. Comparing all regions, Lisbon displayed a higher heterozygous frequency (43.1%) whereas North

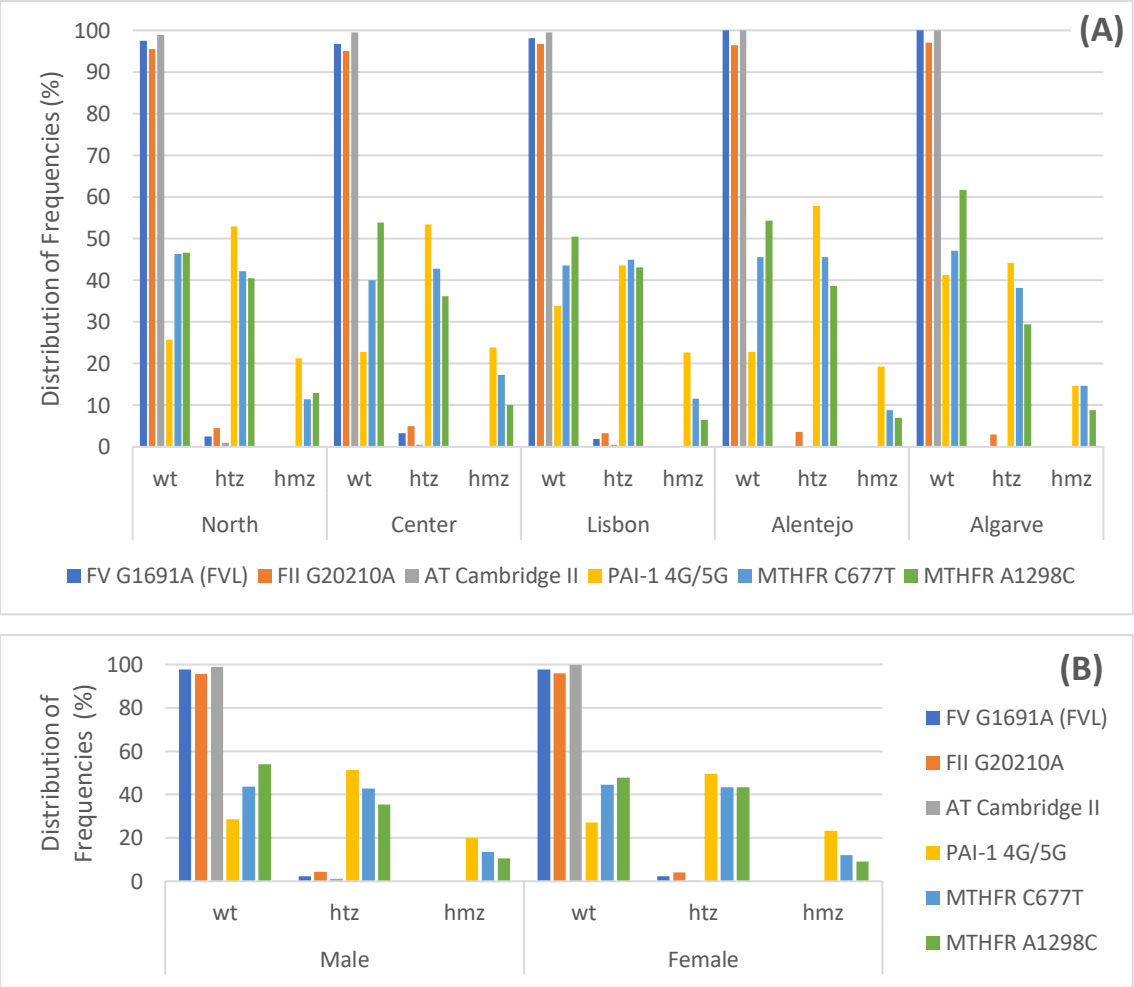


Figure 9| Genotypic frequencies of the thrombogenic variants according to the population characteristics: (A) Portugal continental regions and (B) gender of the participants. *The genotypic frequencies for the Factor V Leiden (FVL), Factor II (FII) G20210A, Antithrombin (AT) Cambridge II, Plasminogen Activator Inhibitor type-1 (PAI-1) 4G/5G and Methylene Tetrahydrofolate Reductase (MTHFR) C677T and A1298C variants are expressed in percentage (%) for each populational group according to the wild-type (wt), heterozygous (htz) and homozygous (hmz) genotypes.*

presented a higher frequency of homozygous subjects (12.9%). According to this study sample, the frequency of the variants' genotype was similar in both genders (**Figure 9B**). Despite no significant differences were found among them, comparing to men, women displayed higher frequencies of *MTHFR* A1298C heterozygous (43.2% females *vs.* 35.4% males) and *PAI-1* 4G/4G homozygous (23.3% females *vs.* 20.2% males) carriers.

3.3 Thrombogenic Variants Frequency (and other Increasing Risk Factors) in Participants with Thrombotic-Related Events

3.3.1 Myocardial Infarction

To evaluate the possible influence of the thrombogenic variants on the occurrence of MI episodes, the frequency of the variants' genotypes was determined for the participants who suffered (n=19) or not (n=754) a MI event (**Table 5**), which are

Table 5| Genotypic frequencies among participants who suffered or not a myocardial infarction event.

Variant	Genotype	Without MI		Diagnosed MI	
		n	%	n	%
***FV G1691A* (*FV Leiden*)**	G/G	738	97.9	18	94.7
	G/A	16	2.1	1	5.3
	A/A	0	0.0	0	0.0
FII G20210A	G/G	726	96.3	16	84.2
	G/A	28	3.7	3	15.8
	A/A	0	0.0	0	0.0
AT Cambridge II	G/G	749	99.3	19	100.0
	G/T	5	0.7	0	0.0
	T/T	0	0.0	0	0.0
PAI-1 4G/5G	5G/5G	210	27.9	4	21.1
	4G/5G	381	50.5	9	47.4
	4G/4G	163	21.6	6	31.6
MTHFR C677T	C/C	330	43.8	10	52.6
	C/T	327	43.4	7	36.8
	T/T	97	12.9	2	10.5
MTHFR A1298C	A/A	378	50.1	13	68.4
	A/C	302	40.1	4	21.1
	C/C	74	9.8	2	10.5

The genotypic frequencies of the genetic variants Factor V Leiden (FVL), Factor II (FII) G20210A, Antithrombin (AT) Cambridge II, Plasminogen Activator Inhibitor type-1 (PAI-1) 4G/5G and Methylene Tetrahydrofolate Reductase (MTHFR) C677T and A1298C among the number (n) of participants who did not suffer from a myocardial infarction event ("Without MI", n=754) and those who had a diagnosis ("Diagnosed MI", n=19) are expressed in percentage (%).

represented in **Figure 10**. Compared to the participants without a MI diagnosis (n=754), it was possible to observe a decreased frequency of the wild-type genotypes for the variants *FVL* (94.7% *vs.* 97.9%), *FII* G20210A (84.2% *vs.* 96.3%) and *PAI-1* 4G/5G (21.1% *vs.* 27.9%) in the group of participants diagnosed with MI (n=19). On the other hand, this group displayed a higher frequency of the heterozygous genotype for the *FVL* (5.3% *vs.* 2.1%) and *FII* (15.8% *vs.* 3.7%) variants and an increased frequency for the homozygous genotype for the *PAI-1* variant (31.6% *vs.* 21.6%). The heterozygous and homozygous frequencies of both *MTHFR* variants were not increased among the participants who developed MI. Yet, it was analyzed if the participants presenting raised levels of Hcy (mild and moderate HHcy) also displayed increased occurrence of MI. While among the participants with mild HHcy (n=140) only 5.7% had endured a MI episode, among those with moderate HHcy (n=10), the incidence of these event was of 10% (1/10). It was also intended to evaluate, if the presence of the CS acquired risk factor in association with the thrombogenic variants might had influence on the development of MI events. Yet such association was not possible to establish since, among the participants who developed MI, only one individual was also a smoker. This smoker participant who experienced MI was also heterozygous for the *FII* G20210A variant and for the *PAI-1* variant, displaying a wild-type genotype for all the other thrombogenic variants.

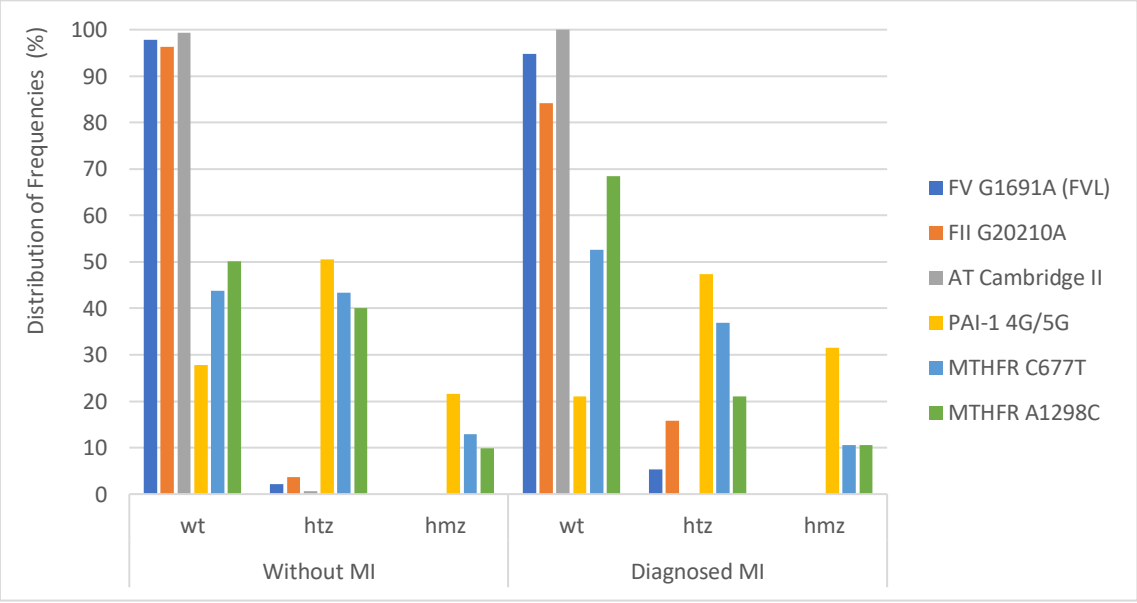


Figure 10| Genotypic frequencies among the group of participants who have never developed a myocardial infarction event and those with a diagnosis. *The genotypic frequencies for the Factor V Leiden (FVL), Factor II (FII) G20210A, Antithrombin (AT) Cambridge II, Plasminogen Activator Inhibitor type-1 (PAI-1) 4G/5G and Methylene Tetrahydrofolate Reductase (MTHFR) C677T and A1298C variants are expressed in percentage (%), according to the wild-type (wt), heterozygous (htz) and homozygous (hmz) genotypes, for the group of participants who have never developed a myocardial infarction event ("Without MI", n=754) and for the group of participants who had ("Diagnosed MI", n=19).*

The frequency of the variants' genotypes was also determined among women users or previous users of OCs who had developed or not MI **(Figure 11)**. Among those OC users and previous users (n=302), 4 women (1.3%) were diagnosed with MI, against 298 (98.7%) who never faced such event. Since there were no cases of current OC users, these women were all previous users, being unknown if the occurrence of the thrombotic event overlapped with the period of OC use. Compared to the frequency of *FII* G20210A heterozygous women without MI diagnosis (n=8, 2.7%) (**Figure 11A**), the frequency of *FII* G20210A heterozygous who experienced MI (n=1, 25%) was higher (**Figure 11B**). However, due to the reduced number of OC users and ex-users diagnosed with MI, no significance can be extrapolated. Regarding HT, no associations were made because there were no participants who had developed MI during such treatments.

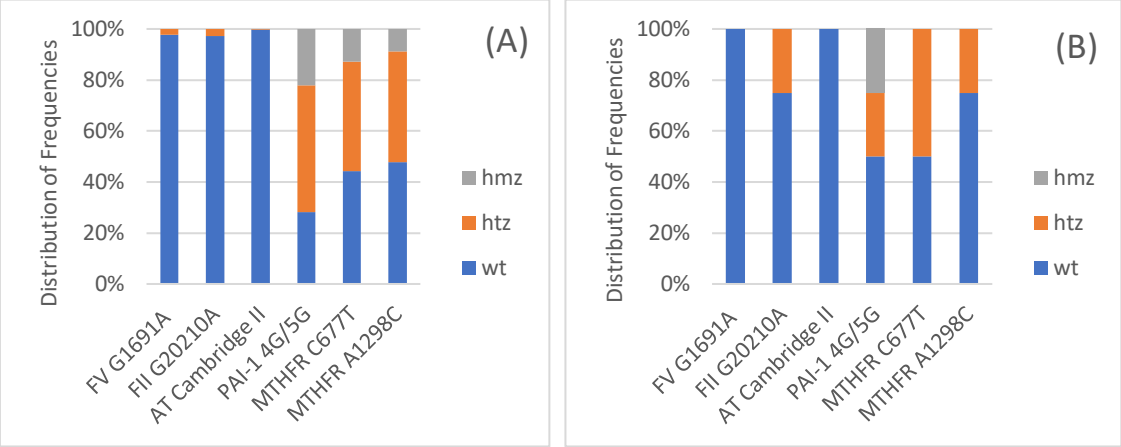


Figure 11| Genotypic frequencies for each thrombogenic variant in women users and previous users of oral contraceptives: (A) Without myocardial infarction and (B) Diagnosed myocardial infarction. *The wild-type (wt), heterozygous (htz) and homozygous (hmz) genotypic frequencies are expressed in percentage (%) for each one of the following variants: Factor V Leiden (FVL), Factor II (FII) G20210A, Antithrombin (AT) Cambridge II, Plasminogen Activator Inhibitor type-1 (PAI-1) 4G/5G and Methylene Tetrahydrofolate Reductase (MTHFR) C677T and A1298C; among the group of women participants users and previous users of oral contraceptives (A) who have never developed a myocardial infarction event (n=298) and among (B) those who were diagnosed with a myocardial infarction (n=4).*

3.3.2 Stroke

The genotypic frequencies of the variants were determined in participants that developed stroke (n=23) and in those who did not (n=750) (**Table 6** and **Figure 12**). The stroke group revealed a higher heterozygous frequency for the *FVL* (13.0% *vs.* 1.9%) and *MTHFR* C677T (65.2% *vs.* 42.4%) variants and a higher homozygous frequency for the *PAI-1* 4G/5G (30.4% *vs.* 21.6%) variant. The frequency of the *MTHFR* C677T homozygous carriers (8.7%) was however decreased compared to that of individuals without stroke diagnosis (12.9%).

Table 6| Genotypic frequencies among participants who have developed stroke and those who not.

Variant	Genotype	Without Stroke		Diagnosed Stroke	
		n	%	n	%
FV G1691A (FV Leiden)	G/G	736	98.1	20	87.0
	G/A	14	1.9	3	13.0
	A/A	0	0.0	0	0.0
FII G20210A	G/G	719	95.9	23	100.0
	G/A	31	4.1	0	0.0
	A/A	0	0.0	0	0.0
AT Cambridge II	G/G	745	99.3	23	100.0
	G/T	5	0.7	0	0.0
	T/T	0	0.0	0	0.0
PAI-1 4G/5G	5G/5G	210	28.0	5	21.7
	4G/5G	378	50.4	11	47.8
	4G/4G	162	21.6	7	30.4
MTHFR C677T	C/C	335	44.7	6	26.1
	C/T	318	42.4	15	65.2
	T/T	97	12.9	2	8.7
MTHFR A1298C	A/A	379	50.5	12	52.2
	A/C	296	39.5	10	43.5
	C/C	75	10.0	1	4.3

The genotypic frequencies of the genetic variants Factor V Leiden (FVL), Factor II (FII) G20210A, Antithrombin (AT) Cambridge II, Plasminogen Activator Inhibitor type-1 (PAI-1) 4G/5G and Methylene Tetrahydrofolate Reductase (MTHFR) C677T and A1298C among the number (n) of participants who did not suffered a stroke event ("Without Stroke", n=750) and those who had a diagnosis of stroke ("Diagnosed Stroke", n=23) are expressed in percentage (%).

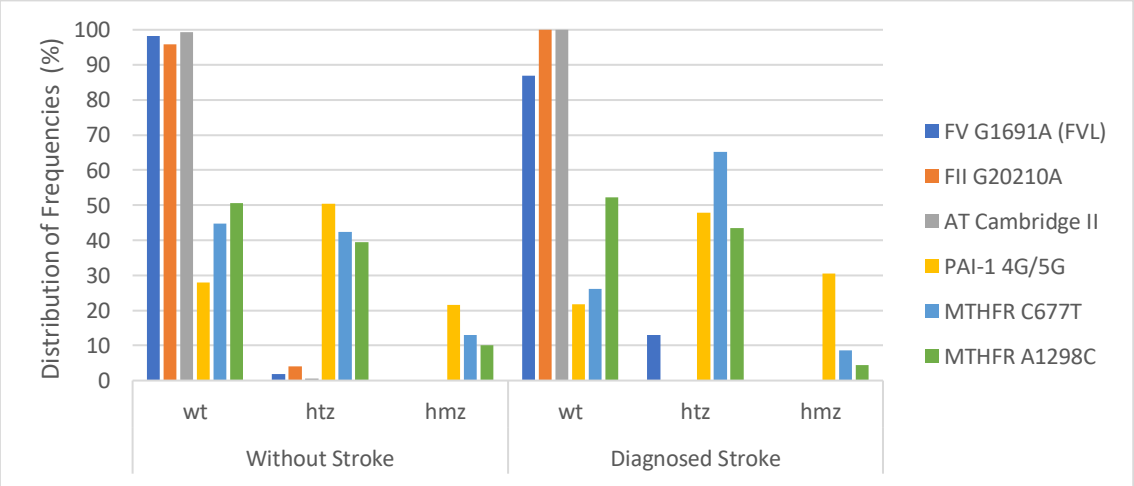


Figure 12| Genotypic frequencies among the group of participants who have never developed a stroke event and those with a diagnosis. *The genotypic frequencies for the Factor V Leiden (FVL), Factor II (FII) G20210A, Antithrombin (AT) Cambridge II, Plasminogen Activator Inhibitor type-1 (PAI-1) 4G/5G and Methylene Tetrahydrofolate Reductase (MTHFR) C677T and A1298C variants are expressed in percentage (%), according to the wild-type (wt), heterozygous (htz) and homozygous (hmz) genotypes, for the group of participants who have never developed a stroke event ("Without Stroke", n=750) and for the group of participants who had ("Diagnosed Stroke", n=23).*

It was evaluated if the group with previous stroke episodes might have developed such event due to the presence of the CS increasing risk factor together with the presence of the thrombogenic variants (**Figure 13**). Among the group diagnosed with stroke (n=23), 17.4% (4/23) were current smokers against the remaining 82.6% (19/23) who did not smoke (this subgroup can however include ex-smokers). Compared to the subgroup of non-smokers (**Figure 13A**), the subgroup of smokers who had developed stroke (**Figure 13B**) showed a higher frequency for the heterozygous genotype of the *FVL*, *PAI-1* 4G/5G and *MTHFR* C677T variants. No wild-type individuals for the *PAI-1* 4G/5G and *MTHFR* C677T variants were observed in the smokers' subgroup. It was also intended to evaluate if higher Hcy levels could contribute to stroke development. However, it was verified that among participants with mild HHcy (n=140), only 6.4% had a stroke diagnosis, whereas among participants who displayed moderate HHcy, none had experienced a stroke event. No correlation could thus be made among these increasing risk factors.

The occurrence of stroke was also assessed in women users or previous users of OCs in association with their thrombogenic genotypes (**Figure 14**). Among the group of users and previous users of OCs (n=302), 8 women (2.6%) had suffered a stroke episode. In this scenario, there were also no cases of current OCs users and so these women were all previous users of OCs. Compared to the other subgroup of OC users/ex-users who have never developed stroke (n=294, 97.4%) (**Figure 14A**), the subgroup of women who suffered stroke (**Figure 14B**) displayed a higher frequency of heterozygosity

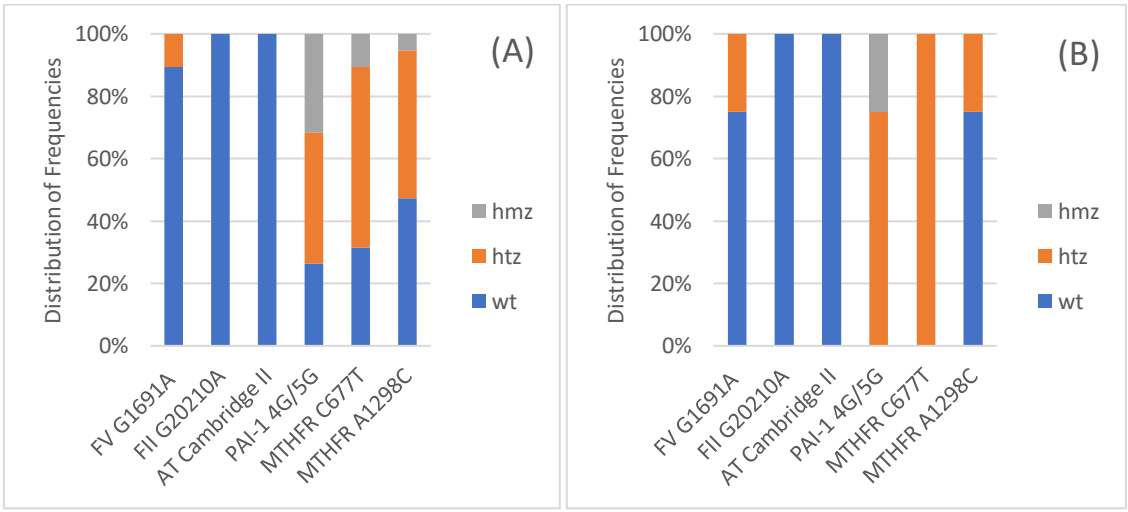


Figure 13| Genotypic frequencies for each thrombogenic variant in participants with a stroke diagnosis: (A) Non-smokers and (B) Smokers. *The wild-type (wt), heterozygous (htz) and homozygous (hmz) genotypic frequencies are expressed in percentage (%) for each one of the following variants: Factor V Leiden (FVL), Factor II (FII) G20210A, Antithrombin (AT) Cambridge II, Plasminogen Activator Inhibitor type-1 (PAI-1) 4G/5G and Methylene Tetrahydrofolate Reductase (MTHFR) C677T and A1298C; among the group of (A) non-smokers who developed stroke (n=19) and (B) smokers who developed stroke (n=4).*

for the *FVL* (25% *vs.* 1.7%), *PAI-1* 4G/5G (62.5% *vs.* 48.6%), *MTHFR* C677T (62.5% *vs.* 42.2%) and *MTHFR* A1298C (50% *vs.* 43.2%) variants. However, no *MTHFR* A1298C homozygous carriers were observed among women who used to take OCs and suffered stroke. There were also no cases in this study sample of women users of HT who had developed stroke.

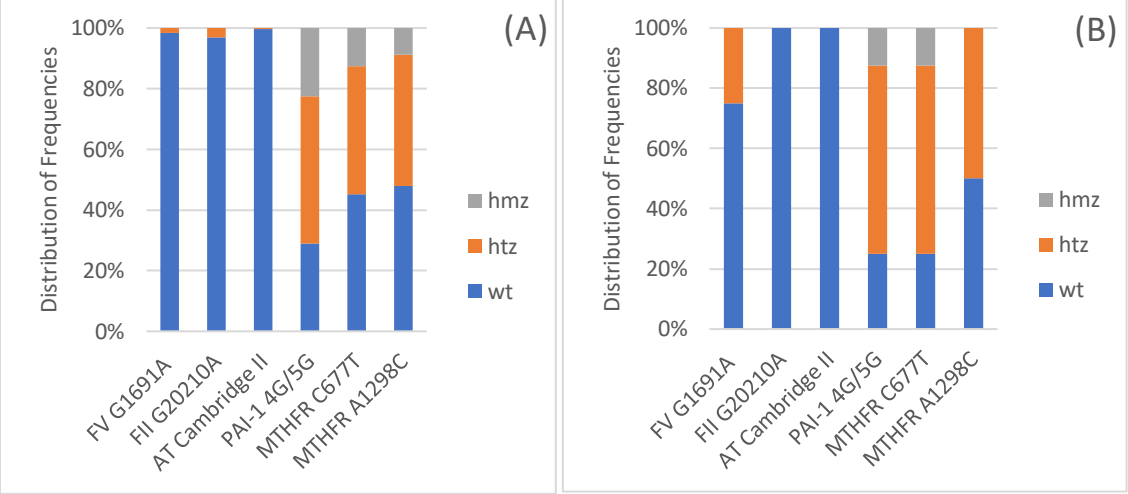


Figure 14| Genotypic frequencies for each thrombogenic variant in women users and previous users of oral contraceptives: (A) Without stroke and (B) Diagnosed stroke. *The wild-type (wt), heterozygous (htz) and homozygous (hmz) genotypic frequencies are expressed in percentage (%) for each one of the following variants: Factor V Leiden (FVL), Factor II (FII) G20210A, Antithrombin (AT) Cambridge II, Plasminogen Activator Inhibitor type-1 (PAI-1) 4G/5G and Methylene Tetrahydrofolate Reductase (MTHFR) C677T and A1298C; among the group of women participants users and previous users of oral contraceptives (A) who have never developed a stroke event (n=294) and among (B) those who were diagnosed with stroke (n=8).*

3.4 Thrombogenic Variants Frequency (and other Increasing Risk Factors) in Women who suffered a Spontaneous Abort

Participant women who had previous been pregnant (n=295) were divided into two groups, with (n=44) or without (n=251) a SA episode, and their genotypic frequencies were determined (**Table 7** and **Figure 15**). Compared to the women without a SA event, the SA group showed a lower frequency of the wild-type genotypes for the *FVL* (95.5% *vs.* 98.0%) and the *PAI-1* (20.5% *vs.* 28.7%) variants. Instead, among this group, there was a higher frequency of the heterozygous genotypes: 4.5% *vs.* 2%, for *FVL*; and 56.8% *vs.* 47%, for the *PAI-1* 4G/5G. Interestingly, despite the higher frequency of wild-type *MTHFR* A1298C among females with SA (56.8%), when compared to the wild-type females without SA (43.4%), and the decreased heterozygous frequency (from 47.4% to 27.3%, from one group to another), it was observed a higher frequency of the

Table 7| Genotypic frequencies in women participants who suffered or not a spontaneous abort.

Variant	Genotype	Without SA		SA		p-value
		n	%	n	%	
FV G1691A (FV Leiden)	G/G	246	98.0	42	95.5	0.281
	G/A	5	2.0	2	4.5	
	A/A	0	0.0	0	0.0	
FII G20210A	G/G	238	94.8	42	95.5	0.607
	G/A	13	5.2	2	4.5	
	A/A	0	0.0	0	0.0	
AT Cambridge II	G/G	251	100	44	100	-
	G/T	0	0.0	0	0.0	
	T/T	0	0.0	0	0.0	
PAI-1 4G/5G	5G/5G	72	28.7	9	20.5	0.426
	4G/5G	118	47.0	25	56.8	
	4G/4G	61	24.3	10	22.7	
MTHFR C677T	C/C	108	43.0	24	54.5	0.288
	C/T	111	44.2	17	38.6	
	T/T	32	12.7	3	6.8	
MTHFR A1298C	A/A	109	43.4	25	56.8	0.037*
	A/C	119	47.4	12	27.3	
	C/C	23	9.2	7	15.9	

*The genotypic frequencies of the genetic variants Factor V Leiden (FVL), Factor II (FII) G20210A, Antithrombin (AT) Cambridge II, Plasminogen Activator Inhibitor type-1 (PAI-1) 4G/5G and Methylene Tetrahydrofolate Reductase (MTHFR) C677T and A1298C among the number (n) of female participants who never suffered a spontaneous abort (SA) ("Without SA", n=363) and those who had suffered a SA ("SA", n=44) are expressed in percentage (%). *MTHFR A1298C frequency variations among the two groups ("Without SA vs. SA") is shown to be statistically significant (p=0.037).*

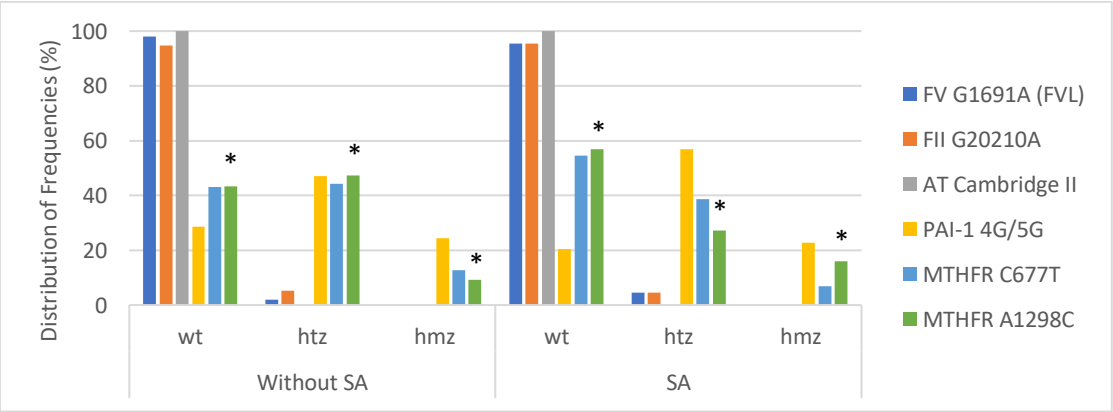


Figure 15| Genotypic frequencies among the group of women who have never suffered a spontaneous abort and the group of women who developed such event. *The genotypic frequencies for the Factor V Leiden (FVL), Factor II (FII) G20210A, Antithrombin (AT) Cambridge II, Plasminogen Activator Inhibitor type-1 (PAI-1) 4G/5G and Methylene Tetrahydrofolate Reductase (MTHFR) C677T and A1298C variants are expressed in percentage (%), according to the wild-type (wt), heterozygous (htz) and homozygous (hmz) genotypes, for the group of women who have never developed a spontaneous abort ("Without SA", n=251) and for the group who had ("SA", n=44). *MTHFR A1298C frequency variations among the two groups was shown to be statistically significant (p=0.037).*

homozygous condition (15.9% against 9.2%, in women with and without SA, respectively). *MTHFR* A1298C frequencies were shown to be statistically significant (p=0.037). The group of women who had at least one SA (n=44) was also split in two subgroups according to women smokers (n=11, 25%) and non-smokers (n=33, 75%) to evaluate the presence of the CS increasing risk factor together with the presence of the thrombogenic variants (**Figure 16**). Compared to the subgroup of non-smoker women who had a SA (**Figure 16A**), the subgroup of smokers (**Figure 16B**) showed a higher frequency for the *FVL* (9.1% *vs.* 3%) and the *PAI-1* 4G/5G (72.7% *vs.* 51.5%) heterozygous genotypes and for the *MTHFR* C677T (18.2% *vs.* 3%) and *MTHFR* A1298C (27.3% *vs.* 12.1%) homozygous genotypes. Despite the increased higher frequency of the *MTHFR* C677T homozygous genotype, this group had a substantial raised frequency of wild-type subjects (72.7% *vs.* 48.5%) and decreased frequency of heterozygosity (9% *vs.* 48.5%).

Due to the previous described important role of the MTHFR enzyme during pregnancy, it was also correlated the presence of *MTHFR* risk alleles of either the C677T or A1298C variants in women who suffered (n=44) or not a SA (n=251). Interestingly, the group of women who had previous suffered a SA displayed a lower frequency of carriers of at least one *MTHFR* risk allele (for either one of the variants) than women who had never suffered a SA (75% *vs.* 88.8%, respectively). In other words, the group of women who had experienced a SA event displayed a higher frequency of the wild-type

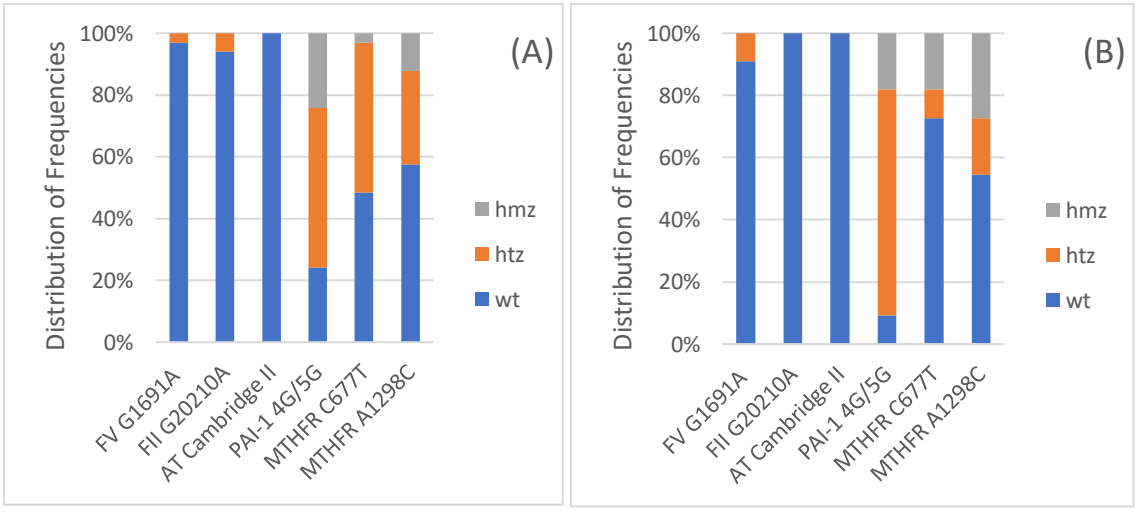


Figure 16| Genotypic frequencies for each thrombogenic variant among women who suffered a spontaneous abort: (A) Non-smoker women and (B) Women smokers. *The wild-type (wt), heterozygous (htz) and homozygous (hmz) genotypic frequencies are expressed in percentage (%) for each one of the following variants: Factor V Leiden (FVL), Factor II (FII) G20210A, Antithrombin (AT) Cambridge II, Plasminogen Activator Inhibitor type-1 (PAI-1) 4G/5G and Methylene Tetrahydrofolate Reductase (MTHFR) C677T and A1298C; among the group of (A) non-smoker women who suffered a spontaneous abort (n=33) and (B) smokers who suffered a spontaneous abort (n=11).*

genotype for both the *MTHFR* variants, compared to women who never had a SA (25% *vs.* 11.2%, respectively). Therefore, the distribution of the *MTHFR* risk alleles among the group of women with SA was accessed and it is represented in **Figure 17**. The maximum of *MTHFR* risk alleles carried by women was 2 (either from the same variant or one from each variant); no women carried either 3 or 4 risk alleles. It is possible to see that 25% (n=11) of the women who suffered a SA (n=44) did not carry any *MTHFR* risk alleles (i.e. displayed the wild-type genotype for both the C677T and A1298C variants); while in the same proportion (n=11, 25%) they were carriers of only one risk allele for the *MTHFR* C677T. It was also found a frequency of 15.9% (n=7) of women carriers of two risk alleles for the *MTHFR* A1298C only, followed by a frequency of 13.6% (n=6) of women carrying a risk allele for the *MTHFR* C677T and another risk allele for the *MTHFR* A1298C.

It was also analyzed if women who displayed at least one of the *MTHFR* variants risk alleles would have a higher frequency of SA if they were also smokers (n=46) or not (n=249) (**Figure 18**), being observed an increase of the SA frequency among women who had the two thrombogenic factors CS and inherited *MTHFR* risk alleles (23.9%, n=11;), compared to those who did not display the CS acquired risk factor (13.3%; n=33). Among women with mild HHcy (n=48), 6.3% had suffered a SA; whereas no cases were detected among women who had undergone such event and had moderate HHcy. Thus, no correlation could be established between these increasing risk factor and the SA pregnancy outcome.

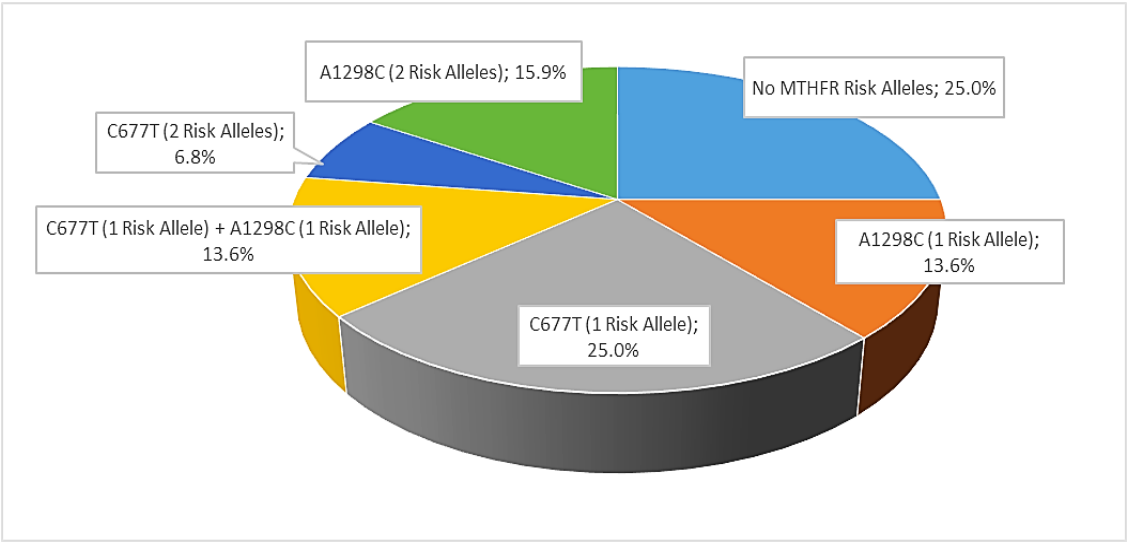


Figure 17| Distribution of MTHFR risk alleles among women who had suffered a SA event. *Distribution of the Methylene Tetrahydrofolate Reductase (MTHFR) C677T and A1298C risk alleles among women who had suffered a spontaneous abort (SA) event (n=44). The frequency of women carriers is expressed in percentage (%).*

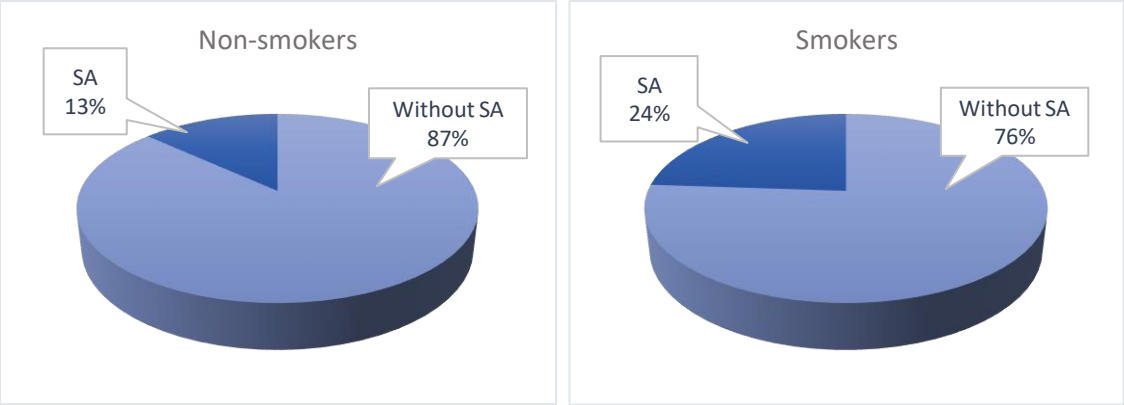


Figure 18| Frequencies of non-smoker women and smoker women carriers of a Methylene Tetrahydrofolate Reductase (MTHFR) gene risk allele who had or hadn't suffered a spontaneous abort. *Frequencies of women carriers of either the C677T or A1298C MTHFR variants risk allele, who are non-smokers (n=249) and smokers (n=46) and who had or hadn't suffered a spontaneous abort event ("SA" or "Without SA" subgroups, respectively) are expressed in percentage (%).*

3.5 Incidence of Multiple Thrombotic Risk Factors in the Population

As previously described, along with inherited genetic defects and increased age (non-modifiable risk-factors), the presence of other acquired risks, such as obesity, raised levels of Hcy, CS habits and the use of OCs and HT are all synergistic factors for thrombotic incidence. So, the study participants who presented such acquired risks were grouped according to either they displayed one or two of the thrombogenic alleles (heterozygous and homozygous genotypes, respectively) for each variant to evaluate the frequency of this high-risk individuals in the sample, that is, to deduce possible groups of high-risk individuals in Portugal.

3.5.1 Frequencies of Individuals with Genetic Risk Factors, Excess Weight and Hyperhomocysteinemia

According to our sample, more than half of the population (64.7%) are overweight (42.1%) and obese (22.6%) (**Figure 19**). In **Figure 20** is represented the frequency of participants (among all sample) who displayed a heterozygous or homozygous genotype for each variant and simultaneously displayed excess of weight (overweigh and obese subjects). Additionally, the frequencies of the obese participants only (group of major increased weight risk) who displayed the genetic defects was also represented: 0.8% (n=6), 1% (n=8) and 0.3% (n=2) of all the study participants were heterozygous for the *FVL*, *FII* G20210A and *AT* Cambridge II variants, respectively; while 12.4% (n=96), 9.6% (n=74) and 8.5% (n=66) were heterozygous, in that order, for the *PAI-1* 4G/5G, *MTHFR*

C677T and *MTHFR* A1298C variants. As previously seen, there were no homozygous subjects in this study population for the *FVL*, *FII* G20210A and *AT* Cambridge II; homozygosity was only present for the remaining variants. Thus, it was determined a frequency of 5% (n=39) obese *PAI-1* homozygous individuals and of 3.2% (n=25) and 1.7% (n=13) obese homozygous carriers of the *MTHFR* C677T and *MTHFR* A1298C variants, respectively.

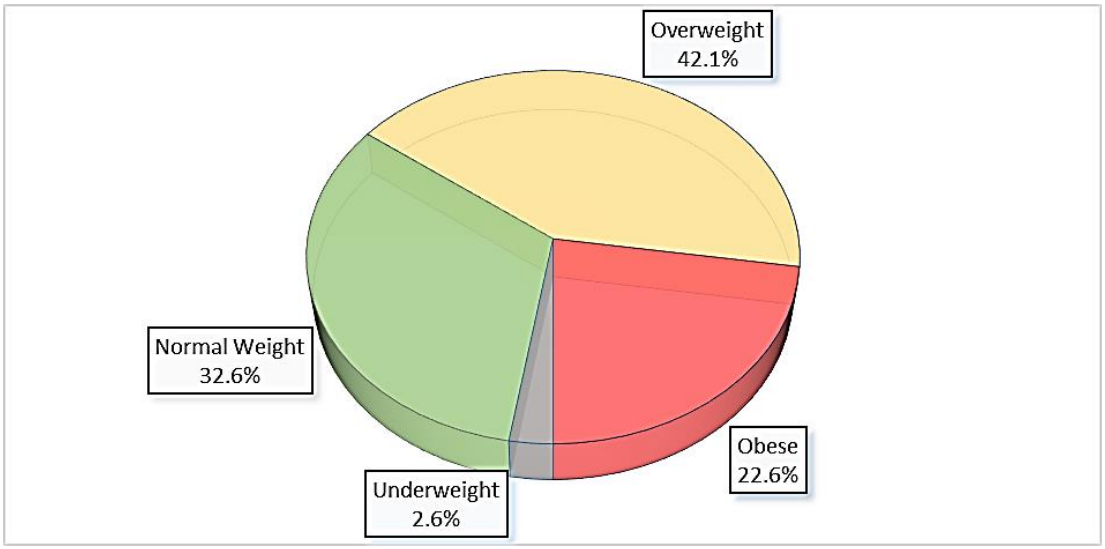


Figure 19| Frequency of the population participants in each weight class. *Frequency of study individuals who are underweight, display the normal (healthy) weight and those who have excess weight, that is overweight or obese subjects, expressed in percentage (%) for the total of 774 participants of the study.*

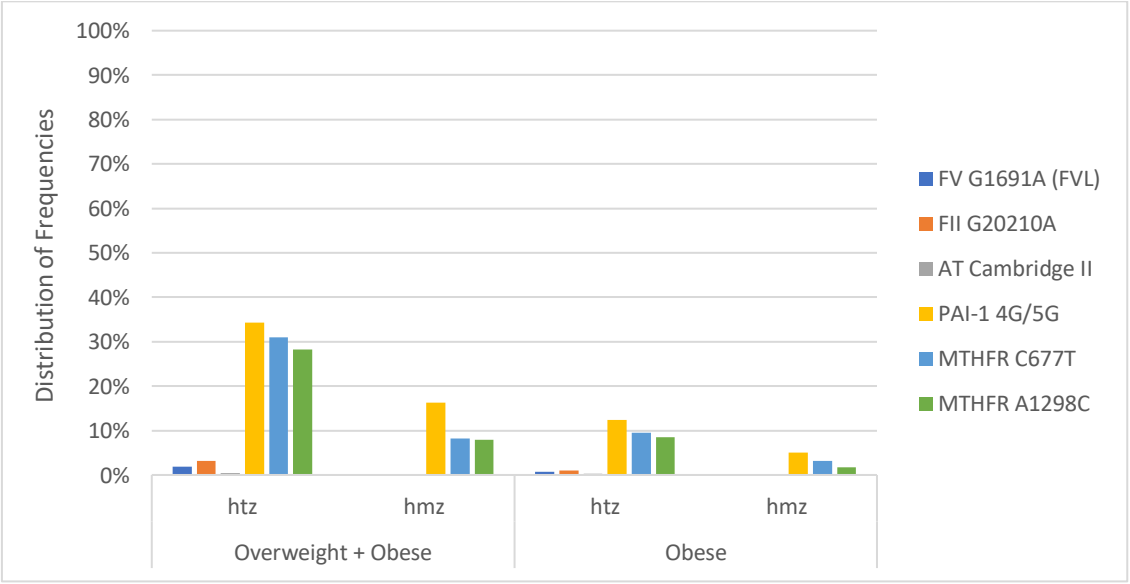


Figure 20| Frequency of individuals who are carriers of thrombogenic risk alleles and display excess weight. *The frequency of heterozygous (htz) and homozygous (hmz) individuals for the Factor V Leiden (FVL), Factor II (FII) G20210A, Antithrombin (AT) Cambridge II, Plasminogen Activator Inhibitor type-1 (PAI-1) 4G/5G and Methylene Tetrahydrofolate Reductase (MTHFR) C677T and A1298C variants who display excess of weight ("Overweight + Obese"), as well as for those "Obese", are expressed in percentage (%) for the total population sample (774 participants).*

It was also tackled the content of subjects with raised levels of Hcy (mild and/or moderate HHcy) who carry thrombogenic alleles (**Figure 21**). As previously determined, there were no cases of severe HHcy, only mild HHcy (n=140, 18.1%) and the moderate condition (n=10, 1.3%) were present. Analyzing the **Figure 21**, it was possible to observe that among subjects with moderate HHcy (the group of participants with higher Hcy levels), no heterozygous individuals were found for the *FVL*, *FII* G20210A and *AT* Cambridge II variants. Therefore the frequency of participants with the heterozygous genotypes for those variants, who also displayed raised levels of Hcy (*"Mild and Moderate HHcy"* group), was, respectively 0.15% (n=1), 0.4% (n=3) and 0.3% (n=2), but all of them displayed mild levels of Hcy. Regarding just the participants with moderate HHcy: 0.4% (n=3) of the population sample were heterozygous and 0.3% (n=2) were homozygous for the *PAI-1* 4G/5G variant; 0.3% (n=2) and 1% (n=8) were heterozygous and homozygous for the *MTHFR* C677T, respectively; and, at last, concerning the *MTHFR* A1298C variant, only 1 heterozygous subject (0.15% of all population) displayed such raised Hcy levels. It was thus possible to observe that among the "Moderate HHcy" group there was a higher frequency of the *MTHFR* C677T homozygous genotype compared to any other genotypic frequency for the remaining variants.

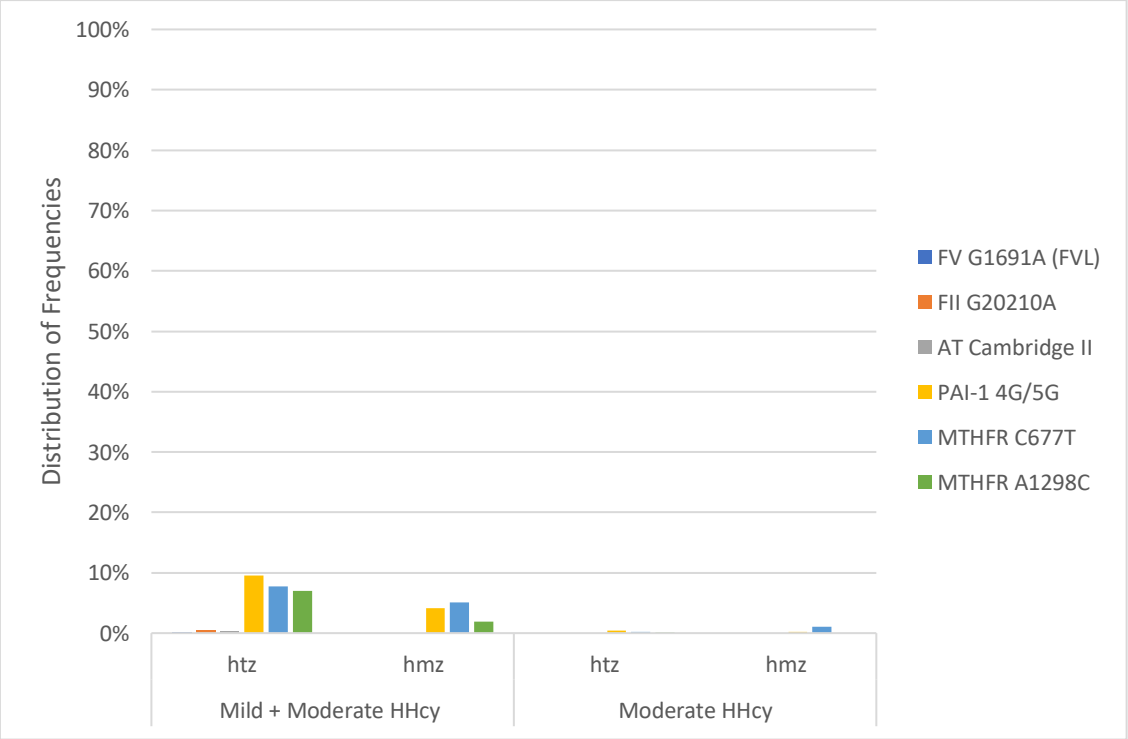


Figure 21| Frequency of the population subjects with increased homocysteine levels. *Frequency of heterozygous (htz) and homozygous (hmz) individuals for the Factor V Leiden (FVL), Factor II (FII) G20210A, Antithrombin (AT) Cambridge II, Plasminogen Activator Inhibitor type-1 (PAI-1) 4G/5G and Methylene Tetrahydrofolate Reductase (MTHFR) C677T and A1298C variants with raised homocysteine levels, that is mild or moderate hyperhomocysteinemia (HHcy) ("Mild + Moderate HHcy"), or with "Moderate HHcy". Frequencies are expressed in percentage (%) for the study population (774 individuals).*

When the risk factors “obesity” and “moderate HHcy” were crossed to determine the frequency of the population participants who were obese and simultaneously displayed such higher Hcy levels (high-risk subjects), 4 individuals (0.5% of the population) were found to display these two acquired thrombogenic risks. Moreover, by analyzing their genetic data, it was also observed that none of those individuals displayed the wild-type genotype for the *MTHFR* C677T variant. In fact, one displayed the heterozygous genotype (this participant was also homozygous for the *PAI-1* 4G/5G variant), whereas the other three revealed to be homozygous carriers. In addition, two of those individuals also carried the *PAI-1* variant (one heterozygous and other homozygous). All these participants were wild type for the remaining variants.

3.5.2 Frequencies of Individuals with Cigarette Smoking Habits in the Presence of Other Risks

The frequency of individuals with CS habits, inherited thrombogenic alleles and weight excess was also assessed (**Figure 22**). It was found that 12.5% (n=97) of all individuals are smokers with weight excess (overweight or obese subjects), with 4% (n=31) being obese smokers. Taking into consideration the genotypic data shown in Figure 22, while 0.5% (n=4) of the study participants were smokers and heterozygous for the *FVL*, half of those (n=2) were also overweight or obese; the frequency of smokers’ *FII* G20210A carriers was of 0.6% (n=5), and one individual was simultaneously overweight; and 0.3% (n=2) were smokers’ heterozygous for the *AT* Cambridge II who were simultaneously overweight or obese. On the other hand, among smoker carriers of one *PAI-1* 4G/5G risk allele (13.2% of all participants, n=102), several were obese (n=17, 2.2% of all participants); whereas among smoking carriers of two *PAI-1* risk alleles (4.1%, n=32), a quarter displayed the obesity risk factor (1% of the population, n=8). Regarding the smokers’ heterozygous for the *MTHFR* C677T (9.3%) and A1298C (10%) variants, a frequency of 1.7% and 1.8% (n=13 and n=14), for each of the previous *MTHFR* variants respectively, were simultaneously obese. For the *MTHFR* C677T homozygous smokers (3.1%, n=24), 3 were found to be obese (0.4% of all participants), while for the *MTHFR* A1298C homozygous cases (1.9%, n=15) just one subject was obese.

It was also evaluated the frequency of smokers who also presented HHcy (**Figure 23**). Compared to the frequency of subjects with CS habits in the population, the frequency of those who also revealed raised levels of Hcy was substantially decreased and, in turn, among those with these thrombogenic factors, the frequency of those with moderate HHcy was even more lower. It was found that 4.5% (n=35) of all individuals

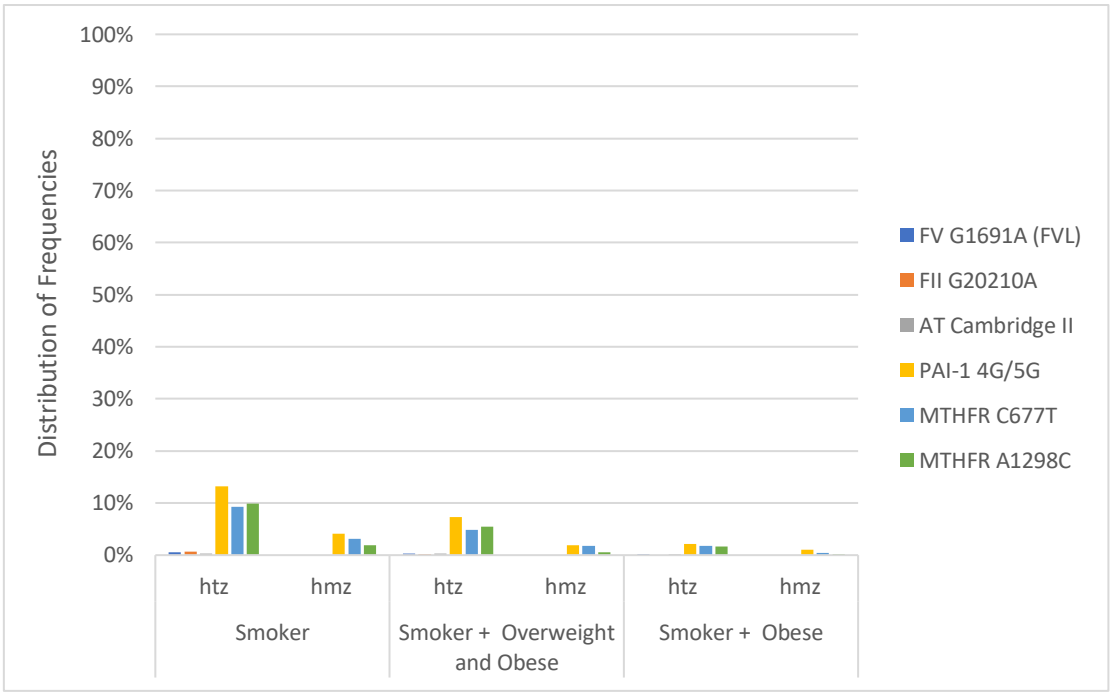


Figure 22| Frequency of individuals with cigarette smoking habits who are carriers of thrombogenic risk alleles and who also display weight excess. *Frequency of heterozygous (htz) and homozygous (hmz) individuals for the Factor V Leiden (FVL), Factor II (FII) G20210A, Antithrombin (AT) Cambridge II, Plasminogen Activator Inhibitor type-1 (PAI-1) 4G/5G and Methylene Tetrahydrofolate Reductase (MTHFR) C677T and A1298C who smoke ("Smoker" group) and who also display excess weight ("Smoker + Overweight and Obese" subgroup) or who in addition are obese ("Smoker + Obese" subgroup). Frequencies are expressed in percentage (%) for the study population (n=774) individuals.*

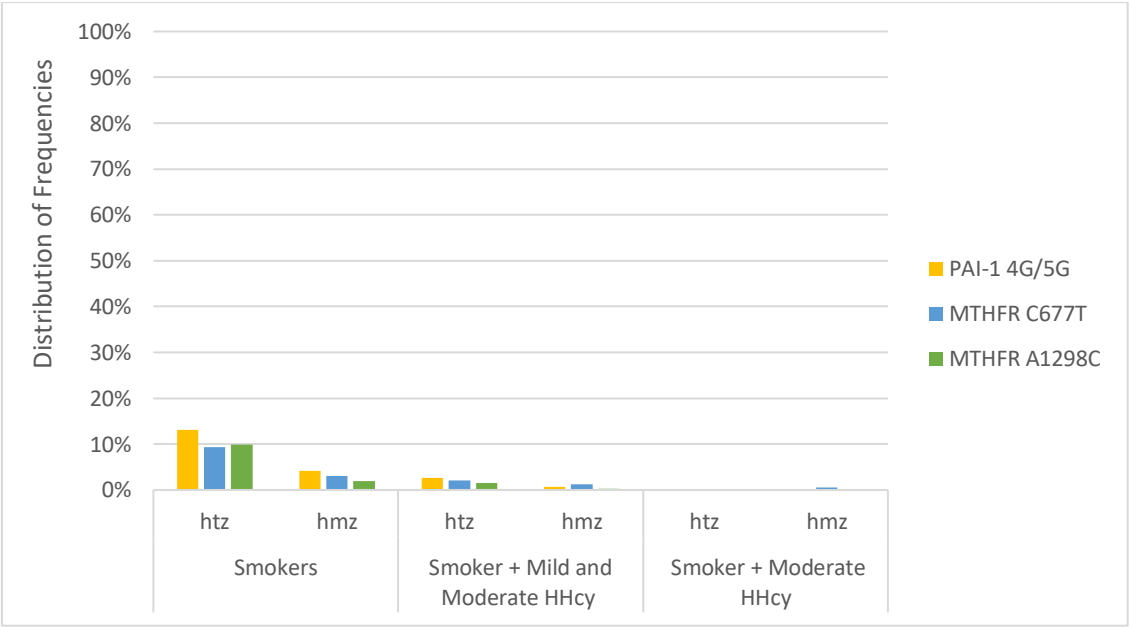


Figure 23| Frequency of individuals with cigarette smoking habits who are carriers of thrombogenic risk alleles and display raised homocysteine levels. *Frequency of Plasminogen Activator Inhibitor type-1 (PAI-1) 4G/5G, Methylene Tetrahydrofolate Reductase (MTHFR) C677T and MTHFR A1298C heterozygous (htz) and homozygous (hmz) individuals with cigarette smoking habits ("Smokers") who also display raised levels of homocysteine, that is, mild or moderate hyperhomocysteinemia (HHcy) ("Smoker + Mild and Moderate HHcy") and of those who display moderate HHcy ("Smoker + Moderate HHcy") among the study population (774 individuals). Frequencies are expressed in percentage (%).*

were smokers with HHcy (mild or moderate) and that 0.65% (n=5) were smokers with the moderate HHcy condition. There were no cases of carriers of *FVL*, *FII* G20210A and *AT* Cambridge II pathogenic alleles who simultaneously displayed CS habits and raised levels of Hcy. However, 2.6% (n=20) and 0.65% (n=5) of the study participants were smokers with HHcy who were heterozygous and homozygous for the *PAI-1* variant, respectively. In the same way (presenting the CS habits and HHcy), 2.1% (n=16) and 1.3% (n=10) were heterozygous and homozygous for the *MTHFR* C677T, whereas 1.6% (n=12) and 0.3% (n=2) were heterozygous and homozygous for the *MTHFR* A1298C. From all participants, there was still 0.13% (n=1) smokers with moderate HHcy who also carried one pathogenic allele for each one of these variants. Additionally, 0.5% (n=4) of the population was homozygous for the *MTHFR* C677T who displayed CS habits and moderate HHcy.

3.5.3 Frequencies of Oral Contraceptives and Hormone Therapy Users with Genetic and Acquired Risks

In the sample group of women (n=407), 29.5% (n=120) were current users of OCs. Along with the use of oral contraception, 2.2% of the female population (n=9) were obese, 8.1% (n=33) were smokers and 0.25% (n=1) displayed all those risks. Additionally, 2% of the women sample (n=8) were users of OCs who were found to have increased levels of Hcy: 1.7% (n=7) displayed mild HHcy and 0.3% (n=1) had moderate HHcy. Taking into account the presence of the thrombogenic variants in OC users (**Figure 24**), only 1 woman was found to be heterozygous for *FVL* (0.2% of the female sample) whereas 3 (0.7%) were heterozygous for the *FII* G20210A variant. Additionally, despite none of them being obese, the heterozygous *FVL* woman and one of the *FII* G20210A heterozygous carriers were simultaneously smokers. There were no cases of OC users' carriers of the *AT* Cambridge II risk allele. The frequency of OC users heterozygous for either *PAI-1* 4G/5G, *MTHFR* C677T or *MTHFR* A1298C variants was relatively higher, when compared to the other thrombogenic variants, being respectively 13.8% (n=56), 11.3% (n=46) and 13.5% (n=55). Along with the use of OCs, 1%-1.2% of the women sample (4 to 5 females), were carriers of one risk allele for each of the previous variants and were also obese, whereas 2.7%-4.4% were heterozygous smokers (11 to 18 females) (**Figure 24A**). The frequency OC users homozygous for those variants was 6.4% (n=26), 3.9% (n=16) and 1.7% (n=7) for the *PAI-1* 4G/5G, *MTHFR* C677T and A1298C variants, respectively. None of the obese OC users was homozygous for the *MTHFR* A1298C and the frequency of *PAI-1* homozygous women

was very low (0.25% of the female sample, n=1). The *MTHFR* C677T homozygous condition is the most frequent across OC users who were obese (0.5%, n=2), while the *PAI-1* 4G/5G homozygous genotype is the more frequent in OC users who were also smokers (1.5%, n=6). Concerning the homozygosity for the *MTHFR* variants, the frequency of OC users at increased risk due to the CS risk factor was about 0.5%-0.7% (2 and 3 women, respectively homozygous for the C677T and A1298C variants) (**Figure 24B**). It was also found one obese woman who took OCs and displayed CS habits. This woman was heterozygous for both the *MTHFR* C677T and A1298C variants.

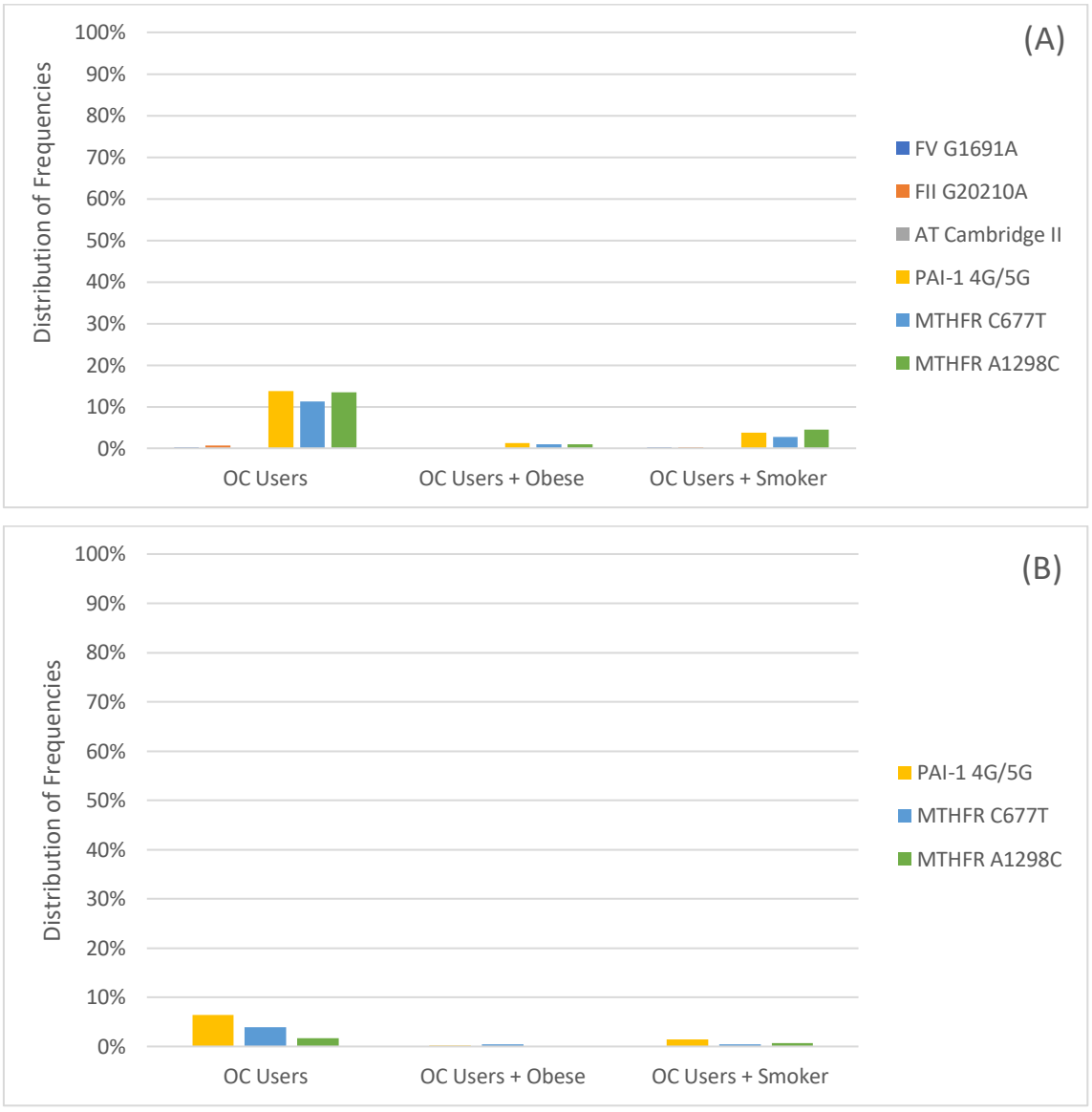


Figure 24| Frequency of women users of oral contraceptives carriers of thrombogenic variants and who display the additional increasing risk factors obesity and cigarette smoking habits: (A) heterozygous carriers and (B) homozygous carriers. *The frequencies of oral contraceptive users ("OC users") carriers of thrombogenic risk alleles for the Factor V Leiden (FVL), Factor II (FII) G20210A, Antithrombin (AT) Cambridge II, Plasminogen Activator Inhibitor type-1 (PAI-1) 4G/5G and Methylene Tetrahydrofolate Reductase (MTHFR) C677T and A1298C and who are also obese ("OC users + Obese") or who smoke ("OC users + Smoker"), among the women study participants (n=407) are expressed in percentage (%).*

Regarding HT users among our female population, 12.3% (n=50) was under the thrombogenic effects of exogenous hormones from HT treatments, 3.7% of women (n=15) were also obese, 1.5% (n=6) were smokers and 0.75% (n=3) displayed all those risks. None of the women with these acquired thrombogenic risk factors were carriers of pathogenic alleles for the *FVL* and the *AT* variants **(Figure 25)**. However, 0.7% (n=3) of women doing HT treatments were heterozygous for the *FII* G20210A variant. Among those, one woman (0.25% of the female population) was also found to be obese, while another (0.25%) was also a smoker. The frequency of HT users who were heterozygous for the *PAI-1* 4G/5G, *MTHFR* C677T and *MTHFR* A1298C variants was, respectively, 6.6% (n=27), 5.2% (n=21) and 5.4% (n=22). In the same order, 2.5% (n=10), 0.5% (n=2)

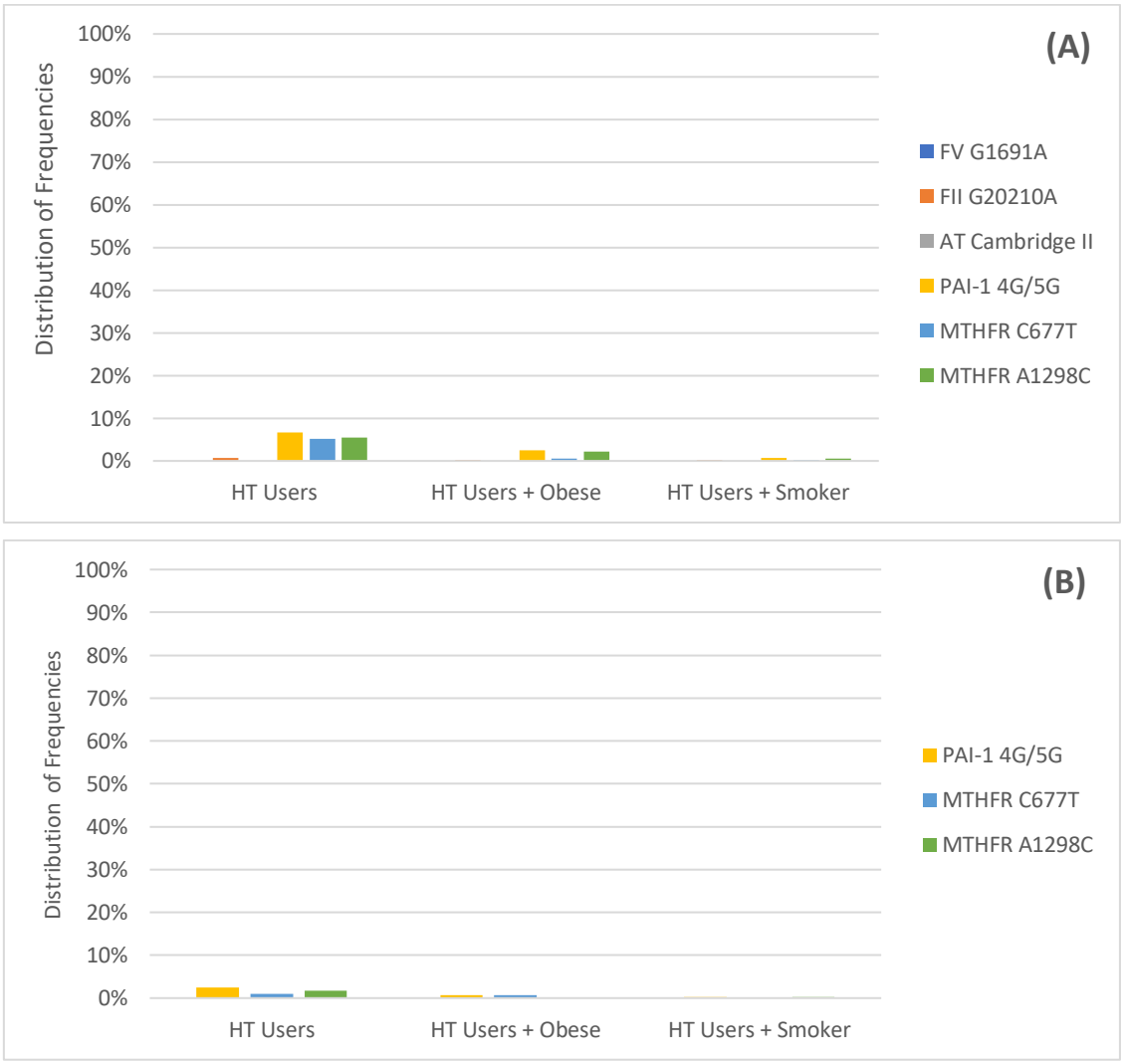


Figure 25| Frequency of women users of hormone therapy carriers of thrombogenic variants and who display the increasing risk factors obesity and cigarette smoking habits: (A) heterozygous carriers and (B) homozygous carriers. *The frequencies of hormone therapy users ("HT users") carriers of thrombogenic risk alleles for the Factor V Leiden (FVL), Factor II (FII) G20210A, Antithrombin (AT) Cambridge II, Plasminogen Activator Inhibitor type-1 (PAI-1) 4G/5G and Methylene Tetrahydrofolate Reductase (MTHFR) C677T and A1298C and who are also obese ("HT users + Obese") or who smoke ("HT users + Smoker"), among the women study participants (n=407) are expressed in percentage (%).*

and 2.2% (n=9) were simultaneously obese. Finally, 0.7% (n=3), 0.25% (n=1) and 0.5% (n=2) were HT users who smoke and carry a pathogenic allele for the *PAI-1* 4G/5G, *MTHFR* C677T and *MTHFR* A1298C variants **(Figure 25A)**. The frequency of females who did HT and were homozygous for the PAI-1, MTHFR C677T and the MTHFR A1298C variants was 2.5% (n=10), 1% (n=4) and 1.7% (n=7), respectively. For obese HT users, a frequency of 0.7% (n=3) was found for the *PAI-1* 4G/5G and the *MTHFR* C677T homozygous condition; whereas for smoker HT users, 0.2% (n=1) were found to be homozygous for either the *PAI-1* 4G/5G or *MTHFR* A1298C variants **(Figure 25B)**. Additionally, there were 3 obese HT users who also displayed CS habits and none of those was wild-type for the *PAI-1* 4G/5G variant: while one displayed the *PAI-1* 4G/5G heterozygous genotype only, another was heterozygous for the *MTHFR* A1298C and the remaining one was homozygous for the *PAI-1* variant (4G/4G genotype) who simultaneously displayed the heterozygous genotype for the *MTHFR* A1298C variant.

3.5.4 Multiple Inherited Genetic Risk Factors

As previously acknowledge, since more than one inherited genetic factor can exist in the same individual, it was accessed the number of thrombogenic variants present in each participant (**Figure 26**). As represented in Figure 26, from all participants genotyped in this study (n=774), only 3.6% (n=28) did not display inherited genetic risk alleles for none of the thrombogenic variants. On the other hand, 27.4% (n=212) were carriers of one (among the six) thrombogenic variants (either in the heterozygous or homozygous condition), 51.4% (about half of the population, n=398) carried 2 different thrombogenic variants, 16.9% (n=131) carried simultaneously 3 different thrombogenic variants and a small minority of 0.4% (n=3) and 0.3% (n=2) were carriers of 4 and 5 variants, respectively. According to our data, most individuals in this population were carriers of two variants only, with the most frequently found combinations being the coinheritance of the *PAI-1* 4G/5G with the *MTHFR* C677T (n=186, 24%) or with *MTHFR* A1298C (n=153, 19.8%). There were no individuals in this population sample who carried all six thrombogenic variants at once.

The frequency of *FVL* and/or *FII* G20210A carriers who also displayed additional thrombogenic genotypes (high-risk fractions of the population due to inherited combined genetic risks) was determined and represented in **Figure 27**. Among *FVL* and *FII* G20210A carriers, no subjects were carriers of the *AT* Cambridge II genetic variant, and so it was also known that the frequency of participants who were carriers of 5 genetic risks (Figure 26), only lack the *AT* variant in their genetic code. None of the *FVL* carriers

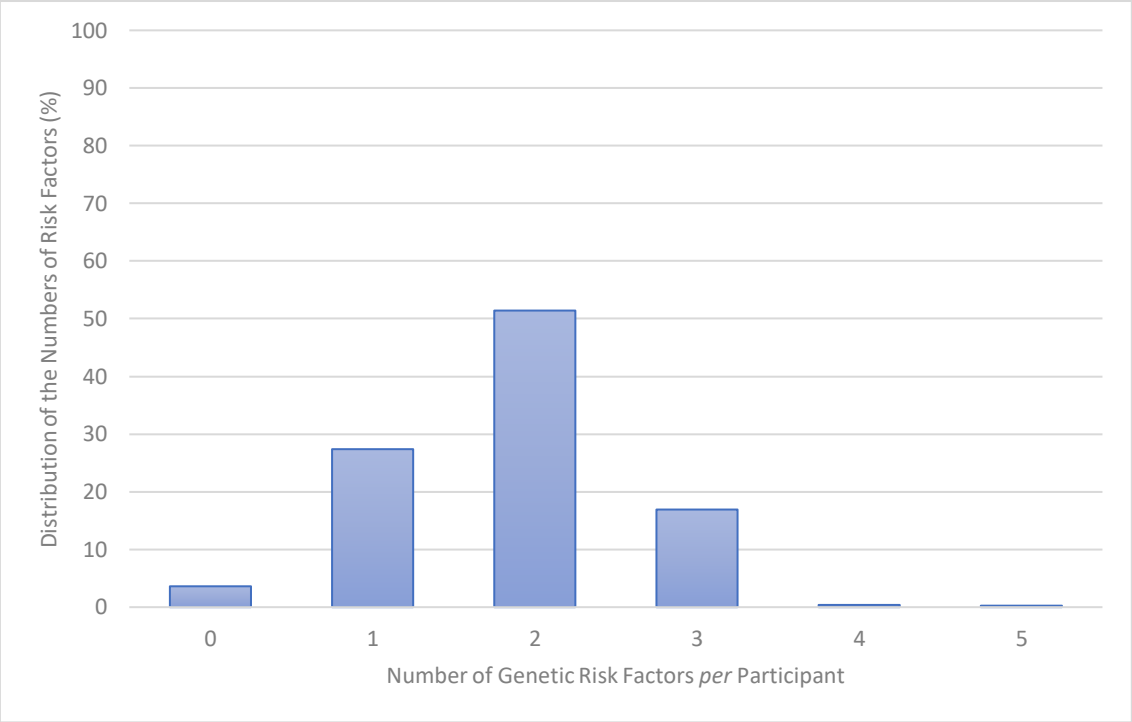


Figure 26| Distribution of the numbers of genetic risk factors present in each participant among the population. *Frequency of participants who carry one or more genetic variants (as well as those with none thrombogenic variant) among the total sample (774 subjects), expressed in percentage (%).*

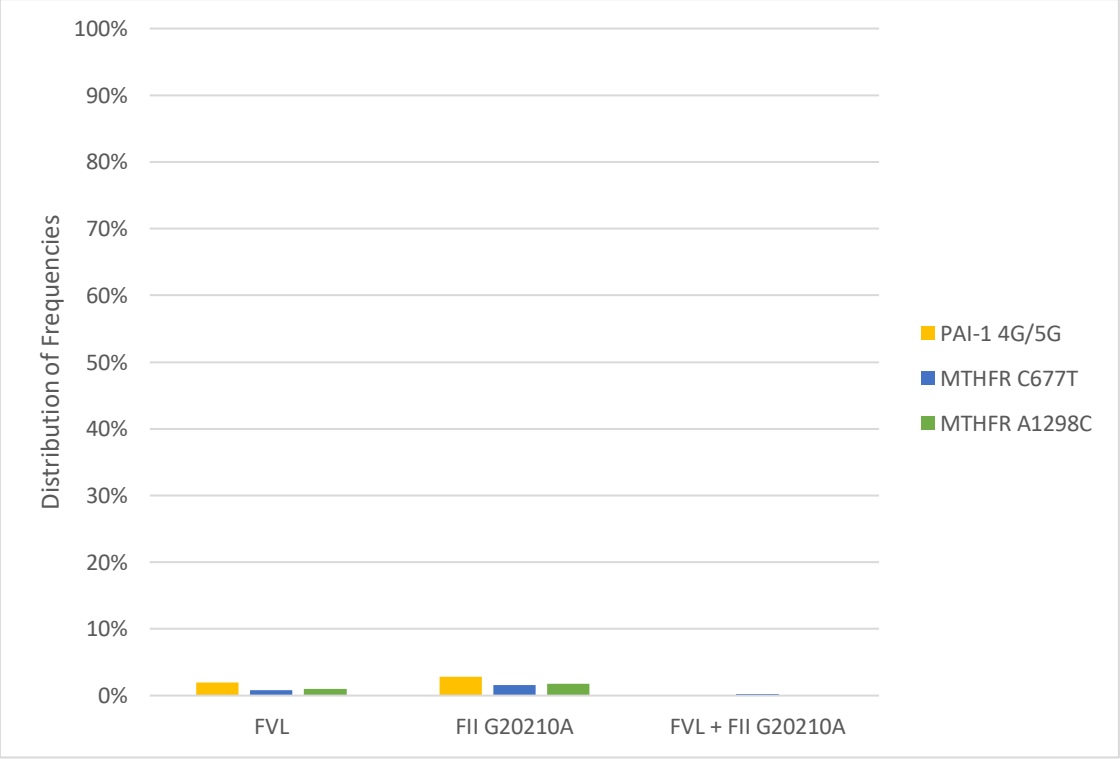


Figure 27| Frequency of Factor V (FV) G1691A or/and Factor II (FII) G20210A carriers who also present additional genetic variants. *The frequency of carriers of the FV G1691A, of the FII G20210A and of both FV G1691A and FII G20210A among the sample (a total of 774 subjects) is expressed in percentage (%), according to either they also carry the Plasminogen Activator Inhibitor type-1 (PAI-1) 4G/5G, Methylene Tetrahydrofolate Reductase (MTHFR) C677T and MTHFR A1298C variants.*

(2.2%, n=17) displayed that genetic risk factor only. Among those participants who were carriers of the *FVL* but who did not carry the *FII* G20210A variant: 1.9% (n=15) were also carriers of one or two risk alleles for the *PAI-1* variant, whereas 0.8% (n=6) and 1% (n=8) were carriers (either heterozygous or homozygous) of the *MTHFR* C677T or the *MTHFR* A1298C, respectively. Regarding the *FII* G20210A variant carriers (4.1%, n=32), only 4 subjects were shown to be wild type for all other variants. Thus, along the carriers of the *FII* G20210A heterozygous genotype (but who were wild-type for the *FVL*), 2.8% (n=22), 1.6% (n=12) and 1.8% (n=14) also carried one or two risk alleles for the *PAI-1* 4G/5G, *MTHFR* C677T and *MTHFR* A1298C variants, respectively. Within these fractions of the population who were either *FVL* or *FII* G20210A heterozygous and simultaneously carriers of the additional thrombogenic genotypes, some participants may present more than the two genetic risks evaluated, as they were not filtrated according to either they were wild type for the other variants. On the other hand, 0.3% of all participants (n=2) were found to be simultaneously heterozygous for the *FVL* and *FII* G20210A and carriers of a risk allele for the *MTHFR* C677T, displaying these 3 genetic risk factors.

Following the previous analysis, the frequency of individuals who carry one or two DNA changes in the *MTHFR* gene (C677T and A1298C) was also accessed and crossed along with the presence of the other variants within the same individual (**Figure 28**). It was found that there were no cases of participants who displayed at the same time a heterozygous and homozygous genotype for those variants as well as there were no subjects' homozygous for both gene changes. It was found, however, that 19.6% (n=152) were heterozygous for both gene variants. In addition, small frequencies of participants displayed the heterozygous condition for both *MTHFR* variants plus other variants: 0.1% (n=1) were simultaneously carriers of *FVL* and 0.3% (n=2) were simultaneously carriers of the *FII* G20210A or the *AT* Cambridge II variants. A higher frequency of 10.2% individuals (n=79) were heterozygous for the MTHFR C677T and A1298C and carried a risk allele for the *PAI-1* 4G/5G variant, while 3.9% (n=30) carried the two pathogenic alleles of the *PAI-1* 4G/5G variant.

Although not separately represented, the *PAI-1* 4G/5G coinheritance with other variants was previously described and inferred from both Figure 27 and Figure 28, with the exception for the simultaneous inheritance with the *AT* Cambridge II. However, as only 5 *AT* carriers (0.6% of the population) were found, the frequency of participants with this variant and other genotypic combinations was extremely low. Still, just like for the *FVL* carriers, none of the *AT* Cambridge II carriers displayed that genetic risk factor only: 2 of them (0.3% of the population) were double heterozygous of the *MTHFR* C677T and A1298C variants and were also heterozygous for the *PAI-1* 4G/5G (carrying 4 genetic

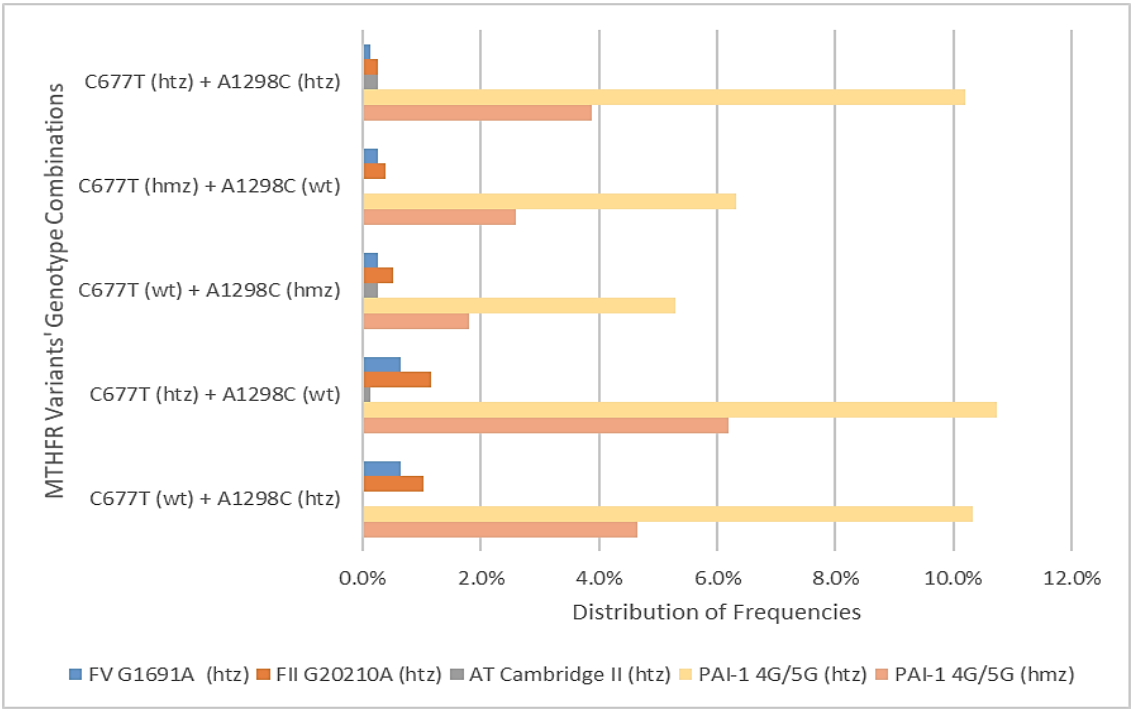


Figure 28| Frequency of carriers of at least one risk allele for either the MTHFR C677T and or the MTHFR A1298C variants who also present additional genetic risk alleles. *The frequency of individuals who carry one or more risk alleles in the MTHFR gene (C677T and/or A1298C variants), that is who are wild-type (wt) and heterozygous (htz) or homozygous (hmz) for either the C677T or the A1298C variants, as well as are heterozygous (htz) for both variants, and simultaneously are heterozygous (htz) or homozygous (hmz) for the Factor V (FV) G1691A, Factor II (FII) G20210A, Antithrombin (AT) Cambridge II or for the Plasminogen Activator Inhibitor type-1 (PAI-1) variants are expressed in percentage (%), for a total of 774 participants of the study sample.*

risks); other 2 (0.3%) were homozygous for the *MTHFR* A1298C (with one of those being also heterozygous for the *PAI-1* 4G/5G) and the remaining one (n=1, 0.15%) was simultaneously heterozygous for the *MTHFR* C677T and homozygous for the *PAI-1* 4G/5G variant.

4. Discussion

4.1 The Thrombogenic Variants incidence in the Portuguese Population and in other European Countries

According to the results obtained in the present study group of participants, from the thrombogenic variants evaluated, the *PAI-1* 4G/5G and the MT*HFR* gene variants were the most frequently found among the mainland Portuguese population. These variants were widely distributed across the continental regions (from Northern to Southern), being quite frequent in both male and female individuals. On the other hand, the *FVL*, *FII* G20210A and *AT* Cambridge II variants were shown to be markedly rarer among Portuguese individuals and were not found in the homozygous condition. All the variants were in HWE. This is important as HWE claims that allele frequencies remain constant from one generation to the next, and so, in the absence of any disturbing factors, the genotypic proportions must persist unchanged over generations. Contrarily, if the genotypic distribution of the variants was deviated from the HWE, it would mean that the gene was evolving (changing)[126,127]. Therefore, such results reflect the absence of the natural selection (positive or negative) influence among the variants' genotypic distribution, suggesting that although they are not a limiting factor for the perpetuation of the human species, none of the variants confer an adaptative advantage to the population.

4.1.1 Plasminogen Activator Inhibitor-1 4G/5G Variant

From all thrombogenic variants analyzed in this study, the *PAI-1* 4G/5G appeared to be the most prevalent in Portugal, as only 27.8% displayed the 5G/5G wild-type genotype compared to 50.4% (half of the population) who were heterozygous individuals and 21.8% who presented the homozygous genotype (4G/4G). The 4G/5G (heterozygous) genotype was thus by far the most common among Portuguese individuals. These Portuguese data did not differ considerably for the *PAI-1* 4G/5G deletion/insertion frequencies stated for several other European countries by Ruiz-Quezada *et al.* (2004). In this latter study, the 5G/5G, 4G/5G and 4G/4G genotypic frequencies were found to be (respectively): 27%, 52% and 21% in Spain; 25.4%, 50.2% and 24.4% in Italy; 24%, 49.7% and 26.3% in France; 24%, 49% and 27% in England; 21.4%, 49.7% and 28.9% in Germany; and 21-23%, 51-53% and 26-27% in Netherlands[88]. In all those countries the heterozygous genotype was always the most

prevalent (and even similar among them), while the wild-type and homozygous frequencies inversely differed among the populations of those countries. Without any doubt, the genotypic frequencies among Spanish individuals were the closest to those obtained for the Portuguese sample. Such results were expected as, due to the countries proximity, the possible effect of geographical differences is not as pronounced as for the other countries. In fact, Portugal and Spain were the ones that display higher frequencies of wild-type individuals; in the other European countries such frequency was slightly decreased, being observed an increased frequency for the homozygous condition instead. The frequency of the *PAI-1* variant was also previously determined in Portuguese healthy unrelated individuals (162 subjects) from all central regions of Portugal by Maia *et al.* (2004). These authors determined a frequency of wild-type individuals substantially higher (39% for the 5G/5G genotype), followed by a smaller frequency of heterozygous individuals (38%) and a slight higher (although very similar) frequency of homozygous (23% of 4G/4G genotype)[128]. However, this data was reported more than a decade ago and it was not representative of the Portuguese population, which can explain the observed differences.

4.1.2 Methylene Tetrahydrofolate Reductase Gene Variants

The *MTHFR* C677T was estimated to be the second most prevalent genetic variant in Portugal: against 44.1% of wild-type subjects, 43.2% of the sample population were heterozygous and 12.8% were homozygous for this genetic variant. The higher prevalence of the *MTHFR* C677T variant is then followed by the *MTHFR* A1298C for which half of the population was wild type (50.6%) whereas, among the other half, 39.5% and 9.8% individuals display the heterozygous and homozygous condition, respectively. The frequency of both the *MTHFR* C677T and A1298C variants has been demonstrated to differ amongst world geographical regions, depending on race and ethnic groups and being also influenced by folic acid demand[97,100]. Both variants displayed a higher heterozygous frequency (in case of the *MTHFR* C677T, very near to the one determined for the wild-type genotype), and smaller (although quite prevalent) homozygous frequencies. Still, the estimated prevalence of the *MTHFR* homozygous conditions, determined by us, was in concordance with the expected for Caucasian populations: up to 15% of Caucasians are homozygous for the *MTHFR* C677T and up to 10% for the *MTHFR* A1298C[100]. Not only the genotypic frequencies but also the coexistence of *MTHFR* genotypes was stated to be influenced by the interaction of the genetic-environmental natural selection[100]. According to previous reports, the coexistence of the

MTHFR C677T and A1298C homozygous and heterozygous genotypes (carriers of three risk alleles) and double homozygosity (carriers of the 4 thrombogenic alleles) was inexistent in many European populations[100]. Equally, these *MTHFR* coinheritances were also not found among our Portuguese population. On the other hand, it was found a frequency of 19.6% individuals heterozygous for both *MTHFR* gene variants, which was also in concordance for the estimated frequency of double (C677T and A1298C) heterozygous for the Caucasian population (15-20%)[100].

Nefic *et al.* (2018) reported the genotypic frequencies of the *MTHFR* C677T in several European countries (wild-type, heterozygous and homozygous genotypes, respectively): 40.4%, 45.2% and 14.4% in Spain; 41.2%, 48.7% and 10.1% in France; 35.4%, 48.7% and 15.9% in Italy; 45.3%, 45.3% and 9.4% in Germany; 47.2%, 41.9% and 11% in United Kingdom; and 50.3%, 41.4% and 8.3% in Denmark[97]. The estimated *MTHFR* C677T prevalence for Portugal is therefore in agreement with the range of the European genotypic frequencies; however, it was quite more similar to the previously determined by Wilcken *et al.* (2003) for the Spanish population (n=601) of 44% for both wild-type and heterozygous subjects and 11.8% for the homozygous genotype carriers[129]. Similar to what was previously observed regarding the genotypic frequencies of *PAI-1* 4G/5G, also the Portuguese *MTHFR* C677T genotypic profile was similar to the one determined for Spain. The estimated prevalence of the *MTHFR* C677T determined in this project was indeed closer to that of the Spanish population than to the previous genotypic frequencies of Portuguese Caucasians revealed by Castro *et al.* (2003). In this previous Portuguese study, a frequency of 43.6% *MTHFR* C677T wild-type subjects, 46.2% heterozygous and 10.3% homozygous individuals were found among 117 healthy volunteers[130]. However, this study was also not representative of the Portuguese population as it did not analyze enough individuals to achieve representativeness.

The Nefic *et al.* (2018) report also displayed the wild-type, heterozygous and homozygous genotypic frequencies of the *MTHFR* A1298C across several European countries: 48.4%, 41.5% and 10.1% in Spain; 51.5%, 40.9% and 7.6% in France; 49%, 42.8% and 8.2% in Italy; 50%, 42% and 8% in Germany; 46%, 44% and 10% in United Kingdom; and 47.7%, 43.6% and 8.7% in Sweden[97]. The European genotypic profile for the *MTHFR* A1298C was very similar among all those countries. Likewise, the estimated prevalence for this thrombogenic variant in Portugal was also in agreement with those stated frequencies, being especially similar to the genotypic frequencies of the Spanish and Italian populations. Regarding the *MTHFR* A1298C genotypic frequencies determined by Castro *et al.* (2003), the frequency of wild-type subjects (49.6%) was quite alike to the one from this study (50.6%), whereas the heterozygous frequency was higher (44.4% *vs.* 39.5%) and the homozygous frequency was lower (6% *vs.* 9.8%)[130].

Another study by Branco *et al.* (2009), this time in the São Miguel Azores island, have also determined the frequency of the *MTHFR* C677T and A1298C genetic changes among a large sample of 469 blood donors, concluding that 34.5%, 47.5% and 17.9% of the sample population were respectively wild-type, heterozygous and homozygous for the *MTHFR* C677T variant; while 56.3%, 37.7% and 6% displayed those genotypes for the *MTHFR* A1298C[120]. Interestingly, comparing to the results of mainland Portugal, the prevalence of the *MTHFR* C677T variant was apparently higher in the Azorean Island, whereas the *MTHFR* A1298C appeared to be more present in mainland Portugal. These findings likely support the influence of the population geographical differences in the genotypic frequency distribution of the *MTHFR* variants. Moreover, as this study is from an island, constituting an isolated small population separated from the larger mainland Portuguese population, the results can be prone to diverge due to the so-called *founder effect*. This effect arises when a small group separates from a larger population, reproducing only among themselves, and so leading to different allele frequencies than the original population[127]. The allelic frequencies of both the *MTHFR* variants (41.7% and 24.8% respectively for the *MTHFR* C677T and A1298C risk alleles) differ among the smaller population of the Azorean Island compared to those in the general population of continental Portugal (34.4% and 29.6% respectively for the *MTHFR* C677T and A1298C risk alleles); however, once again this study was not representative of the population and, so, comparisons cannot be made. Additionally, the higher *MTHFR* C677T homozygous frequency among the Azorean population can also suggest consanguinity as, in addition to the island social isolation, widespread cousin-cousin marriages can empower the founder effect[127].

4.1.3 Factor V Leiden and Factor II (Prothrombin) G20210A

As described in the literature, a significant geographic variation was observed for the prevalence of the *FVL* and *FII* G20210A variants which, despite differing from population to population, are almost exclusive of Caucasian populations[23,57]. These variants were shown to be highly present among white individuals of European origin, being quite uncommon (almost null) in African and Asian populations[31,50,66,131]. Based on our sample, the estimated prevalence of the *FVL* (2.2%) and *FII* G20210A (4.1%) in mainland Portugal was significantly lower when compared to the estimated for the other thrombogenic variants. Moreover, they were only found in the heterozygous condition. Such results were however anticipated as, according to the established for the general Caucasian populations, *FVL* frequency is expected to range between 1-2.5%[57], while *FII* G20210A frequency is believed to range 1-5%[69]. Furthermore, the frequency of *FVL*

homozygous subjects is stated to be about 1 in 2500[93] and, as our sample of participants comprises only 774 individuals, it was anticipated why no homozygous carriers were found. The *FII* G20210A homozygous condition was also extremely rare (until 2006, just 70 cases were reported worldwide) and mostly analyzed in case report studies[132,133]. For these reasons, no homozygous frequency of reference was found.

Data regarding the frequency of *FVL* in several European populations was collected in a previous report by Lucotte and Mercier (2001), in which the *FVL* was estimated to be present in a frequency of 1% for Portugal; 3.3% for Spain; 1-3.5% for France; 1.8-3% for Italy; 3% for Netherlands; 2.8% for England; 7.8-8.8% for Germany and, at last, 8.4% for Greece[66]. In line with these data, among those European countries, Greece and Germany displayed a higher incidence of *FVL* carriers while Portugal is shown to be the country with less *FVL* carriers. However, according to the estimated prevalence of the *FVL* (in the present study), its frequency was now higher and more similar to that obtained for England in the Lucotte and Mercier (2001) study. Regarding the *FII* G20210A, more than two decades ago, Rosendaal *et al.* (1998) have reported a prevalence of 3% for this variant among southern European countries, which was nearly twice as high as its prevalence in northern Europe (1.7%)[134]. More recently, Jadaon (2011) have reported higher *FII* G20210A frequencies for healthy European populations, though the same pattern was observed: it appears to be two times more present among south Europe (France, Spain, Italy, Greece...) than in the Northern countries (United Kingdom, Netherlands, Germany, Denmark, Sweden,…), respectively 0.7-8% and 0-2.9%[135]. The estimated prevalence determined in this project for Portugal was however higher than the previously estimated for the southern European region.

Nevertheless, similar Portuguese frequencies to those obtained in this project were reported by Mansilha *et al.* (2006), who determined a frequency of 2% and 5% heterozygous carriers for the *FVL* and *FII* G20210A variants correspondingly, among healthy unrelated individuals from the north of Portugal. No homozygous carriers were also found for neither variants on the Mansilha *et al.* (2006) study[136]. Therefore, the frequency of these variants among Portuguese individuals appeared to remain constant in both studies, being only observed in the current study a very small increase on the *FVL* estimated prevalence and a minor decrease in the *FII* G20210A. Once again, this latter study was not representative of the population as these frequencies were only obtained from a sample of 100 (healthy) individuals from Northern region. On the other hand, the Branco *et al.* (2009) Azorean study, reported a significant higher frequency of *FVL* heterozygous subjects (9.8%) and a smaller frequency (although not very different) of 3.2% for the *FII* G20210A heterozygous condition. Additionally, they also reported one case of homozygosity for the *FII* G20210A variant (homozygous genotypic frequency of

0.2%)[120], while among our sample no homozygous were found. Therefore, while the allelic frequency of the *FVL* and *FII* G20210A risk alleles in our study was found to be of 1.1% and 2.1%, respectively; in the Azorean island it was determined to be 4.9% and 1.8%. Again, this higher allelic frequency for the *FVL* variant in the Azorean Island can be caused by the founder effect and consanguinity. It is possible that, after the founder population event, such allele (present in one copy) was amplified while the population was still small, causing a strong genetic drift[127,137].

4.1.4 Antithrombin Cambridge II

Among all variants studied in this project, the *AT* Cambridge II was found to be the rarest, being determined an estimated prevalence of only 0.6% for the Portuguese population. This is in agreement with the fact that, despite being the most frequent cause of antithrombin deficiency in Caucasian populations[78], the *AT* Cambridge II is a low prevalent polymorphism, estimated to be present in only about 0.2%-0.3% of the general population[74,75] and mostly in the heterozygous condition[51]. The genotype data for this study participants revealed however a double frequency of that estimated for the general population, which can indicate that this variant prevalence varies among region and ethnicity, being possible to hypothesize that while in Portugal (and possibly in other European Caucasian populations) there is a slightly higher incidence of this genetic alteration, its incidence in other populations worldwide must be almost null (just like described for the *FVL* and *FII* G20210A variants). Still, the estimated Portuguese prevalence of *AT* Cambridge II remained very low. There were also no homozygous individuals among the sample of participants, which was expected according to the previous stated concerning its rareness and being predominantly present in the heterozygous state.

Only a few studies have determined the frequency of this thrombogenic defect (largely in patients, to analyze the influence of such genetic flaw in pathologies' onset) and so, very little information is available regarding this variant frequency among general populations. However, a Spanish study by Corral *et al.* (2007) studied this variant among patients with VTE and had determined a frequency of 0.2% among control individuals. Until then, the *AT* variant frequency had only been determined in a western Scotland population in which it was found to be present in 1 for each 630 individuals (0.16%)[59]. Posteriorly, another study from Sanchez *et al.* (2008) determined an equal frequency, to that estimated here for Portugal, of 0.6% among healthy French Caucasians. No homozygous individuals for the *AT* Cambridge II mutation were reported in the Sanchez *et al.* (2008) study too[79]. With this data we can only conclude that the *AT* variant is

indeed rare, as it is not possible to make trustable comparisons with these population frequencies since, due to its rareness, a large number of participants (than that used in case-control studies) is required to determine its real prevalence.

4.2 Individuals at Thrombotic Increased-Risk in the Population

The pathogenesis of thrombosis has been previous known to have a multifactorial origin. As already described, the combination of genetic (intrinsic) and acquired (extrinsic) risk factors will establish each individual thrombotic risk[93]. Nevertheless, when the synergy of these intrinsic and extrinsic factors (the cumulative thrombotic risk) exceeds the critical threshold (the sum of the individual risk factors), thrombosis develops[83,93]. Proper thromboprophylaxis can prevent exceeding this critical threshold, however, even under treatment, thrombosis can arise if the thrombotic risk overcome the protective mechanisms[93]. The risk of developing any thrombotic event (venous or arterial) is well-known to exponentially increase with age, possibly due to increased immobilization and increased levels of some plasma clotting factors. However, when it occurs at younger age, its incidence may have been influenced by the presence of double or multiple gene defects[16,62], since genetic predisposition has a key role in the occurrence and evolvement of thrombosis[53]. Moreover, the high prevalence of thrombosis and its known environmental influences, suggests that multiple genes of varying effects will be involved in determining its susceptibility[105].

4.2.1 The Genetic Component of Thrombosis Development

Despite well-recognized that both forms of thrombosis are under the significant influence of genetics, since early has been discussed the overlap (or lack of it) regarding the genetic predisposing component for VTE and arterial thrombosis. The *FVL*, *FII* G20210A and *MTHFR* C677T risk alleles are indeed believed to be more correlated to VTE than to arterial thrombosis, being controversial whereas they may also confer a modest risk for the pathogenesis of arterial thrombotic events[23]. However, in the present study, the heterozygous genotype for the *FVL* was found to be increased in participants who had developed both MI or stroke episodes; while the frequency of the *FII* G20210A heterozygous genotype was raised in participants who suffered an MI event, and the *MTHFR* C677T heterozygous genotype in participants' diagnosed with stroke. In line with these associations, this study results seem to support the possible influence of such variants in the onset of arterial thrombotic-related events. Moreover, recent data

concerning these variants as risk factors for arterial thrombotic events have emerged: Chiasakul *et al.* (2018) suggested the association of the *FVL* and *FII* G20210A with an increased risk of arterial IS (stronger in the homozygous than in the heterozygous state)[138], whereas Li *et al.* (2017) have demonstrated that the *FII* G20210A variant may indeed represent a risk factor for MI, as they found a statistically significant association of MI with either the heterozygous or homozygous genotype of the *FII* G20210A variant[68]. It is thus possible that 2.2% and 4.1% of the population may be at thrombotic risk due to respectively carrying the *FVL* and the *FII* G20210A risk alleles. Still, the heterozygous genotype for either the *FVL* or the *FII* G20210A by themselves are considered relatively weak risk factors, in opposition to their homozygous state[63,69]. Favorably, no cases of homozygosity were found for either the *FVL* or the *FII* G20210A variants among our population. It was however detected that 0.3% of the population carried the *FVL* and *FII* G20210A genetic defects, simultaneously, and may therefore display an exacerbated risk. In fact, according to the literature, coinheritance of *FVL* and *FII* G20210A variants may confer a substantial increase in the thrombotic risk[43,103,139], which can be six times higher when compared to wild-type subjects for both genes[140]. Moreover, such fraction of the population was found to simultaneously carry the *MTHFR* C677T heterozygous genotype (which had also revealed an increased frequency in participants diagnosed with stroke). No other variant (besides the *MTHFR* C677T, in the heterozygous condition) was present in participants with the *FVL* and *FII* G20210A thrombophilia risk factors.

The *MTHFR* C677T, as well as the *MTHFR* A1298C, were also previously stated to be important risk factors for stroke, by themselves or in combination[141]. The *MTHFR* A1298C homozygous genotype was indeed found to increase the risk for stroke, being suggested that it could be a predictor for future stroke incidence[142]. Another study reported a significant positive association between *MTHFR* C677T homozygous genotype and the risk for MI, but no significant association was found with the *MTHFR* A1298C variant[143]. However, as latter mentioned, only the *MTHFR* C677T heterozygous frequency was raised among stroke diagnosed participants. It was therefore expected that, if these variants were somehow related to an increased arterial thrombotic incidence, the frequency of wild-type subjects who developed such events would be decreased whereas the heterozygous or homozygous genotypic frequencies would be raised, compared to those who did not develop thrombosis (either in both stroke and MI diagnosed groups). Thus, and since the *MTHFR* C677T is quite prevalent in the population, it is possible that the higher heterozygous frequency observed in such participants with a stroke diagnosis can be due to coinheritance of the *MTHFR* C677T heterozygous genotype with the *FVL* variant.

Our results revealed that participants who have developed either MI or stroke also displayed a higher frequency for the *PAI-1* 4G/5G homozygous genotype. In fact, patients that suffered IS, MI or other thrombotic disorders were found to display increased plasma PAI-1 levels[144], which can suggest that these individuals may carry at least one *PAI-1* 4G/5G risk allele. A previous study by El-Aziz *et al.* (2015) concluded that the *PAI-1* homozygous genotype, but also the presence of only a 4G risk allele, are associated with MI risk[145]. More recently, a systematic review and meta-analysis by Liu *et al.* (2018) also established a significant association between the *PAI-1* 4G/5G variant (heterozygous and homozygous genotype) and MI occurrence[86]. Additionally, an association between the homozygous *PAI-1* 4G/5G and the stroke incidence was reported by Akhter *et al.* (2017)[87]. All these data may indeed suggest the influence of the *PAI-1* 4G/5G homozygous genotype in arterial thrombosis, and so, about 22% of the Portuguese population (*PAI-1* 4G/5G homozygous) can be at higher risk for developing a thrombotic event. Moreover, in addition to homozygosity, and according to the described in the literature, subjects with multiple thrombogenic variants (which occurs rarely) are also expected to be at high-risk[63]. Thus, and since it was observed an incredible high prevalence of this thrombogenic variant among the population, most of the *PAI-1* 4G/5G carriers may be at higher-risk if they are also carriers of additional thrombogenic defects. The risk is especially high if there is a synergistic combination of hypercoagulability and hypofibrinolysis[84]. Hence, 1.9% and 2.8% of the population may be at even greater risk by carrying the *PAI-1* 4G/5G variant and simultaneously the *FVL* and *FII* G20210A variants, respectively. Fortunately, no *FVL* and *FII* G20210A double heterozygous participants were also carriers of at least one *PAI-1* 4G/5G risk allele (4G).

Antithrombin deficient homozygous subjects are also believed to be at increased risk for both arterial and venous thrombosis[58], despite controversial results on the antithrombin deficiency and the *AT* Cambridge II role in thrombosis incidence[58,79,146]. However, like the obtained for the *FVL* and *FII* G20210A variants, no *AT* Cambridge II homozygous individuals were found among this population. Moreover, the small minority of the population (0.6%) who carried this thrombogenic variant has never developed a MI or stroke event and so no associations could be established. This 0.6% of the population who carried this variant may be however at increased risk due to the combined presence of the *AT* Cambridge II and other increasing genetic risks. Although no *AT* Cambridge II carriers simultaneously carry the *FVL* and *FII* G20210A genetic variants, none of those presented that genetic risk only. In fact, two of the *AT* carriers may be at higher risk since they were found to present 4 genetic risks at once (double heterozygosity for the *MTHFR* C677T and A1298C variants and heterozygosity for the *PAI-1* 4G/5G variant).

Finally, it is important to acknowledge that, according to the present study, only 3.6% of the Portuguese population did not carry any risk alleles for the *FVL*, *FII* G20210A, *AT* Cambridge II, *PAI-1* 4G/5G, *MTHFR* C677T and A1298C thrombogenic variants, with the others 96.4% carrying at least one thrombogenic risk allele for any of those. Yet, despite the episodic nature of thrombosis suggesting that (1) the presence of an inherited risk factor *per se* is not enough to trigger a thrombotic event and that (2) it is required the interaction with other inherited or acquired risk factors to onset the disorder[49,50], only 27.4% of all participants carried a single thrombogenic variant (either in the heterozygous or homozygous condition). Therefore, the remaining majority of 69% can be at increased predisposition to thrombosis by carrying at least two different genetic variants. The co-existence of two genetic thrombotic defects in the same individual may greatly enhance the overall risk[43,103,139] by the double or triple of the risk associated to a single genetic variant[147].

Most study participants (51.4%) were determined to carry exactly two genetic risk factors, with the most prevalent coinheritance being the combination of the *PAI-1* 4G/5G with either the *MTHFR* C677T (n=186, 24%) or with *MTHFR* A1298C (n=153, 19.8%) variants. Therefore, 43.8% of the population (subjects with these combined inherited variants) are prone to display: (1) impaired fibrinolysis (due high basal PAI-1 levels, that can be increased up to 25% in the *PAI* 4G/5G homozygous condition[88,91]), (2) greater procoagulant activity (including fibrin formation)[84] and (3) increased Hcy levels (due to a less active MTHFR enzyme). This last point can result in prolonged HHcy, which promotes the formation of atherosclerotic plaques, enhanced inflammation and endothelial dysfunction, leading to the occurrence of atherothrombotic events[92,94,99,100]. Both the C677T and A1298C enzymatic isoforms decrease the MTHFR activity, impairing the methylation of Hcy and hence leading to its blood accumulation (HHcy)[95–97,100]. However, comparing to the *MTHFR* A1298C variant, the *MTHFR* C677T presents the most deleterious effects: it decreases the MTHFR activity to a great extent (less 30-40% and 60-70%, for the heterozygous and homozygous genotype *vs.* the 20-30% or 40-50% reduction conferred by the A1298C heterozygous and homozygous genotypes) and also increases the MTHFR thermolability (which is not affected by the A1298C variant)[95,100,148]. This way, the coinheritance of the *PAI-1* 4G/5G with the *MTHFR* C677T was not only the most prevalent but it is also expected to be most harmful than the coinheritance of the *PAI-1* 4G/5G with the *MTHFR* A1298C. Nevertheless, the 43.8% of the population who displayed the previous described thrombophilic profile is at enhanced risk of thrombosis incidence only due to the harmful effects of the genetic component. For this reason, it is quite important to acknowledge the existence of acquired

(increasing) risk factors among the population, which are usually expected to be more common than inherited thrombophilia[147].

4.2.2 Synergy of Acquired and Genetic Thrombotic Risks

In addition to the genetic component, obesity is well-known to confer a higher risk for thrombosis[93]. This acquired risk factor is quite worrisome on the Portuguese population, as only 32.6% of all participants had a normal healthy weight, compared to the big majority of overweight (42.1%) or obese (22.6%) individuals. Furthermore, almost a quarter of the population had acquired CS habits (24.3%) and 19.4% of the population may be at increased risk due to a state of mild to moderate HHcy. No cases of severe HHcy were found among the participants. Still, regarding all these acquired risk factors, it was found that 4% of the population were obese smokers and 4.5% were smokers with HHcy (0.65% display moderate HHcy). Both the CS and HHcy risk factors were found to be more prevalent among men.

CS is reported to be an important increasing risk factor for stroke, increasing up to 50% the risk of such thrombotic incident in smokers. The risk increases with the increasing cigarette consumption, due to CS procoagulant and atherogenic effects[144]. When the genetic component of the participants diagnosed with stroke was crossed with the presence of CS habits, it was found a higher frequency for the *PAI-1* 4G/5G heterozygous genotype. The *FVL* and the *MTHFR* C677T heterozygous frequencies were also higher, however such frequency was already increased in the stroke diagnosed group, even without taking into consideration CS habits. Interestingly, none of the smoker participants who had developed stroke was found to be wild-type for either the *PAI-1* 4G/5G or the *MTHFR* C677T variants, which can reinforce the idea that CS habits can synergistically increase the risk of stroke development in *PAI-1* 4G/5G and *MTHFR* C677T carriers. No associations were however performed regarding MI occurrence as, in this study sample, only one smoker participant had developed MI. Curiously, this subject was also found to present a heterozygous genotype for the *FII* G20210A variant, which was already associated to MI (even without taking into consideration other risk factors) and for the *PAI-1* 4G/5G variant. It is however relevant to mention that the group of non-smoker participants can comprise ex-smokers who had previously developed MI (or stroke), being unknown if they were under the influence of this acquired risk factor when they developed the thrombotic event. Thus, not only *FVL* and/or *FII* G20210A heterozygous and/or *PAI-1* 4G/5G homozygous subjects are expected to have an exacerbated thrombotic risk (especially if they also smoke, but even

if they don't) but also those who are heterozygous for the *PAI-1* 4G/5G or/and the *MTHFR* C677T variants can be at higher risk if they display the CS acquired risk.

According to the literature, smoking is an exogenous stressor involved in the upregulation of PAI-1, associated with increased concentrations of this inhibitor[149], due to the cigarettes' nicotine component that induces PAI-1 production[144]. This way, CS displays a fibrinolysis modulation effect, since fibrinolysis requires the conversion of the plasminogen into plasmin by the plasminogen activators, which are rapidly inactivated by PAI-1[144]. Moreover, smoker carriers of the *MTHFR* C677T and /or the A1298C variants (who are prone to have reduced MTHFR activity) are expected to have a higher loss of the Hcy re-methylation process due to the additional induced demethylation associated to CS[107]. Surprisingly, no increased frequency of the *MTHFR* A1298C heterozygous or homozygous genotypes was found among smoker participants who developed stroke. Therefore, smoker individuals' carriers of a at least 1 risk allele for either the *PAI-1* 4G/5G (about 17% of the population) or the *MTHFR* C677T (about 12% of the population) variants can be at higher risk of thrombosis due to a possible synergistic effect conferred by the CS habits. Moreover, from all six thrombogenic variants evaluated, these two were also the more commonly coinherited among this study population. So, as there is a large part of the population that presented both thrombogenic risk factors, also a big part of the population must have increased risk due to CS habits. It is also plausible that about 3% of the population are at even higher risk due to also carrying at least one *PAI-1* 4G/5G risk allele and the obesity risk factor. In the same way, whereas 1.7-1.8% were smoker heterozygous for each one of the *MTHFR* variants and simultaneously obese, 0.4% of all participants (n=3/774) were smoker obese subjects homozygous for the *MTHFR* C677T and one obese smoker was found to be homozygous for the *MTHFR* A1298C. It is also described in the literature that the several pathological effects of CS are quickly reduced after quitting[27], and so quitting smoking must display rapid relieving benefits regarding the synergistic thrombotic risk conferred by such acquired risk.

Moderate HHcy participants showed a higher incidence of MI (almost the double) compared to those with mild HHcy (10% *vs.* 5.7%), suggesting that higher levels of Hcy may contribute to the incidence of thrombotic events. In fact, HHcy is associated with increased risk of vascular events and with a significantly higher risk of developing an early VTE event, being thus a factor of increasing risk for thrombosis in general[140,150]. As reported in the literature, severe HHcy is rare, the mild and moderate conditions are more common[93] and can be caused by the *MTHFR* genetic mutations (usually by the *MTHFR* C677T variant), accompanied by nutritional deficiency of folate and vitamin cofactors[93,120]. A recent study by Munshi *et al.* (2019) have indeed emphasized that

thrombosis patients with the *MTHFR* C677T variant have significantly elevated Hcy levels (compared to the A1298C variant) and are at higher risk of rapid progression of their disease[151]. There were no cases of carriers of the *FVL*, *FII* G20210A and *AT* Cambridge II pathogenic alleles who simultaneously display CS habits and raised levels of Hcy. However, small fractions of the population may be at higher risk by presenting these risk factors in association to the remaining genetic variants, which were indeed previously described to display increased frequencies among smoker participants who suffered a stroke. Therefore, about 3% carried at least one *PAI-1* 4G/5G risk allele, whereas also about 3% were carriers of *MTHFR* C677T risk alleles. Lastly, although our results did not shown an increased frequency for the *MTHFR* A1298C pathological genotypes in participants with thrombotic related-events, about 2% were carriers of the *MTHFR* A1298C risk allele, who may also be at higher risk due to a possible synergistic effect (with both HHcy and CS habits, and probable due to the existence of more genetic risks among these individuals). Among all population, 0.13% (n=1/774) were smokers' heterozygous carriers for each one of these variants that presented the moderate HHcy condition, while in the same conditions there were 0.5% (n=4/774) homozygous carriers for the *MTHFR* C677T.

4.2.2.1 Women at Increased Risk of Thrombosis

In general, women are expected to be at major thrombotic risk than men since, in addition to the common acquired risk factors, they can also be under influence of the exogenous hormones' effects, granted by OCs and HT use. Although exogenous hormonal use appears to be safe for healthy women (that is, with low baseline risk)[152], current OC or HT use are associated with a 2 to 6-fold increased risk of VTE[147]. It was however suggested that women with inherited thrombophilia who take OCs can be at 20- to 30-fold higher VTE risk (compared to nonusers with thrombophilia or users without the genetic variants)[93]. It is also suggested that, obesity alone can involve approximately a 24-fold increase in the VTE risk among OC users[153]. On the other hand, women under HT are not expected to reach such enhanced risk as HT has a lower estrogen dose and, consequently, a lower biological potency[31,154]. Although the risk of HT may be significantly lower than that of OCs, it is possible that it might be increased by the older age of the HT users, who are in menopause.

Based on this study, it is believed that about 29.5% of Portuguese women are current users of OCs and that, despite not using it anymore, 45% were previous users (who possibly interrupted their use only by entering in menopause). Additionally, women who use OCs and are obese (2.2% of the women population) or have CS habits (8.1%)

may be at increased risk, particularly if they present all those risks (0.25% of the female population). On the other hand, 12.3% of women were users of HT, but the prevalence of this acquired risk factor among Portuguese women can possibly be higher as 53.8% of the female participants were unaware if they were under HT treatments. Nevertheless, 3.7% of the female population are at increased risk by also presenting the obesity risk factor, whereas 1.5% had CS habits. Comparing to the OC users who were simultaneously obese and smokers (0.25%, n=1/407), a higher frequency of smoker obese HT users was found (0.75%, n=3/407).

Among the participating women currently using OCs, there were no cases of thrombotic occurrence, only previous users had confirmed such incidences, being unknown if the timing of OC use overlapped with the thrombotic events. According to the present study, the group of women previous users of OCs who had developed MI also presented a higher *FII* G20210A variant frequency. This particular variant may be especially important for women's thrombogenic risk since a previous study had reported a 4-fold increased risk of MI in women, while for men the risk was about less than half (1.5-fold)[134]. It was also verified a higher frequency of *FVL* heterozygous women among previous users of OCs *vs.* non-users (25% *vs.* 1.7%) in association to stroke occurrence, and of *PAI-1* 4G/5G, *MTHFR* C677T and *MTHFR* A1298C heterozygous genotypes. In fact, OCs were previously associated with a 3.3- to 4.2-fold increased risk of VTE in carriers of *FVL*[25], while more recent data defends that *FVL* carriers may have a 35-fold increased risk of VTE, associated with the usage of exogenous hormones (either OCs or HT)[18]. This huge risk difference may depend obviously of the carriers' genotype, but it can also rely on the additional presence of the acquired risks (for example, CS, HHcy, obesity and additional genetic risks). Due to the influence of the several factors that can synergistically increase the overall thrombotic risk in the presence of exogenous hormones, some studies suggested that women who are seeking hormonal contraception should be screened for the presence of other risk factors[117,152].

The use of OCs is mostly reported to further increase the risk of VTE associated with *FVL* or *FII* G20210A[147]. While the synergistic interaction of OC use with the *FVL* is probably related to the fact that both risk factors lead to the APC-R phenomenon[154] (up to 60% of women who experienced VTE during OC use were also found to be APC resistant)[31], the combination of OC and the *FII* G20210A variant is believed to have a multiplicative effect on overall thrombotic risk as a result of increased prothrombin levels[154]. Thus, women under the effects of OCs who were also carriers of the *FVL* (0.2%) and *FII* G20210A (0.7%) variants may be at high-risk of thrombosis, as OCs may potentiate the thrombogenic effects of these variants, especially in the presence of other acquired risk factors. Those who use OCs and display the more prevalent variants in the

Portuguese population, that is, carriers of at least one risk allele for the *PAI-1* 4G/5G (about 20% of all women participants), *MTHFR* C677T (about 15%) and *MTHFR* A1298C (about 15%) may also be at higher risk, even if in less extent when compared to the *FVL* and *FII* G20210A variants. However, these women can be at an even higher risk if, despite being heterozygous for these variants, they are also obese (1%-1.2%) or smokers (2.7-4.4%). If they are homozygous instead, the risk is expected to be enhanced, but the high-risk population fractions were smaller among the female population: only 0.25% (n=1/407) and 1.5% (n=6/407) were *PAI-1* 4G/5G homozygous who take OCs and are respectively obese or smokers, 0.5% (n=2/407) were obese *MTHFR* C677T homozygous carriers and about 0.5%-0.7% were homozygous smokers for each *MTHFR* C677T and A1298C variant.

As there were no cases in our study sample of HT users who had MI or stroke events, no associations were performed with the presence of the thrombogenic variants. However, according to the literature, the *FVL* and the *FII* G20210A variants are also expected to be the main genetic risks in women under hormonal treatments[155,156], especially if the acquired thrombogenic component is present. Besides no women with the obese and CS habits acquired risk factors were carriers of a pathogenic allele for the *FVL*, among the 0.7% of women who did HT treatments and were heterozygous for the *FII* G20210A variant, 0.25% (n=1/407) were also obese while other 0.25% (n=1/407) were smokers. Regarding the most prevalent inherited risks in Portugal, several fractions of the female population may be at increased thrombogenic risk by carrying at least one thrombogenic risk allele of these variants and presenting the obesity and CS risks: about 6 to 10% of the female participants were carriers of at least one respective risk allele for each *PAI-1* 4G/5G, *MTHFR* C677T or *MTHFR* A1298C variants; whereas among those fractions, up to 3% were simultaneously obese, while from those who smoke its frequency does not exceed 1%. Additionally, there were 3 obese women among our population who used HT and had CS habits (3 acquired thrombogenic risks); none of them is a wild-type for *PAI-1* 4G/5G variant: 2 presented the *PAI-1* 4G/5G heterozygous genotype (one was simultaneously heterozygous for the *MTHFR* A1298C) and another one was homozygous for the *PAI-1* 4G/5G (who was also heterozygous for the *MTHFR* A1298C). So, these women may present a substantial higher synergistic risk for thrombosis incidence due to the concomitant association with CS habits and the *PAI-1* 4G/5G and *MTHFR* variants.

4.3 Pregnancy Thrombotic Risks and the Incidence of Spontaneous Abortion

Pregnancy constitutes an acquired hypercoagulability state for women[93]. During pregnancy, women undergo several hemostatic changes, that include not only the increase of procoagulants and decrease of the natural anticoagulants, but also reduced fibrinolytic activity. This "acquired hypercoagulable state" aims to help maintaining the placental function and to minimize further blood losses at delivery. Yet, such changes may predispose to maternal thrombosis and placental vascular complications[25]. The risk of VTE in pregnancy is believed to be increase by 5- to 10-fold[147], however the individual risk of VTE in pregnancy is influenced by women related-factors, such as age older than 35 years and obesity[93]. The risk can be boosted up to 34-fold in the presence of hereditary thrombophilias[113]. Maternal hereditary thrombophilia is indeed a risk factor for SA as the additional hemostasis disturbances, caused by the inherited genetic defects, compromise the placental system[114].

Pregnancy is expected to be the most common predisposing and circumstantial risk factor for thrombotic episodes in *FVL* or *FII* G20210A carriers[154]. According to Kujovich *et al.* (2004), the *FVL* variant was previously found to be present in 20–46% of women with pregnancy-associated VTE and it is associated with a 5- to 16-fold increased thrombotic risk during pregnancy and the puerperium[154]. On the other hand, but apparently not so frequent, the *FII* G20210A variant was found to be present in about 6–26% of pregnant women with VTE, being believed to confer a 3- to 15-fold higher risk[154]. Our female participants with a previous event of SA had indeed presented a higher frequency of *FVL* carriers (4.5% *vs.* 2%), yet the frequency of *FII* G20210A carriers among women who had or hadn't developed a SA (4.5 *vs.* 5.2%, respectively) was not very different and it was indeed higher among the group without SA. A previous study by Sosic *et al.* (2017) have also evaluated the presence of *FVL*, *FII* G20210A, *MTHFR* C677T and *PAI-1* 4G/5G variants among women with previous SA, finding a significantly higher frequency for all variants' risk genotypes, with the exception of *FII* G20210A[157]. Nevertheless, conflicting results regarding the role of the thrombogenic variants in SA episodes exist. While significant associations were observed for both heterozygous and homozygous *FVL* carriers and for *FII* G20210A heterozygosity[104,150], Baumann *et al.* (2013) indicates that the maternal *FVL* or *FII* G20210A variants do not have a significant influence in the pathophysiology of recurrent SA among Caucasian women[114]. In the Baumann *et al.* study the homozygous genotype of the *MTHFR* C677T revealed to be more prevalent instead[114], which was not observed in the present study. In fact, the

MTHFR A1298C homozygous frequency was the one that revealed to be higher among women who had previous suffered a SA.

The MTHFR enzyme is extremely important during pregnancy, not only to control the Hcy levels, but also because it maintains the methyl pool (for control of gene expression by DNA methylation) during implantation and invasion of the embryo[158]. The C677T and A1298C variants have already been described to promote a HHcy condition which can cause endothelial damage and induce platelet aggregation. Such damage is particularly harmful in pregnant women, as it may cause the obstruction of the placental vessels, resulting in recurrent abortions. Thus, the maintenance of pregnancy may be compromised if women display these *MTHFR* genetic variants[159,160]. However, in the present study, only the *MTHFR* A1298C homozygous frequency was found to be higher among women with SA (15.9% *vs.* 9.2%, p=0.037). In fact, Nair *et al.* (2013) had previously suggested a significant risk of pregnancy loss associated with the *MTHFR* A1298C variant[158]; whereas, recently, such risk associated with the *MTHFR* A1298C homozygous genotype was reported by Madduru *et al.* (2019) to be 2.5-fold higher[161]. In concordance with our findings, a previous study by Yang *et al.* (2016) had also established a significant association between the *MTHFR* A1298C variant and recurrent pregnancy loss, but not with the *MTHFR* C677T[162].

The *MTHFR* C677T homozygous condition is stated to severely alters pre-implantation development[96], being also associated with HHcy, which in turn is significantly associated with abortion[150]. Therefore, it is surprising that women who had suffered a SA did not had a higher frequency of the *MTHFR* C677T harmful genotypes. Moreover, when it was analyzed the distribution of the *MTHFR* risk alleles among women who had suffered a SA, those who were carriers of the two risk alleles of the *MTHFR* C677T variant were a small minority, whereas a quarter of the women who had SA were carriers of only one risk allele for the same variant (25%, n=11/44) and another quarter did not carry any risk allele, either for the C677T or the A1298C variants (25%, n=11/44). Overall, the group of women who had previously suffered a SA presented a lower frequency of *MTHFR* risk allele carriers than women who had never suffered a SA (75% *vs.* 88.8%). These data seem to suggest that the incidence of SA among our female participants was under influence of other factors than just the presence of *MTHFR* risk alleles by themselves. Also, the Hcy levels in normal pregnancy are *per se* elevated, compared with nonpregnant women[158], and so it can be hypothesized that, although *MTHFR* variants may exacerbate a possible HHcy condition, even without the influence of the genetic component, pregnant women may be at increased risk. In concordance with our results, a previous review article by Robertson *et al.* (2006) had stated that pregnant women homozygous for the *MTHFR* C677T variant did not indeed present a

significant association with VTE (there was no raised risk), in contrast to the non-pregnant women; it was also suggested that this lack of association was possibly due to folic acid supplementation during pregnancy, which reduces Hcy levels, consequently lowering the risk[150]. Therefore, pregnant carriers of this thrombogenic variant can, at least to some extent, control its harmful effects by folic acid supplementation and so, prevent possible adverse pregnancy outcomes as SA. Such explanation can indeed justify why there was a higher frequency of the *MTHFR* C677T variant among women without SA, that is, why these women did not suffer an abortion. However, when the CS risk factor was considered, the *MTHFR* C677T homozygous genotype showed a higher frequency (18.2% *vs.* 3%) among women smokers who had suffered SA when compared to the subgroup of non-smokers, suggesting a possible synergy of CS with the *MTHFR* variant. That is, the presence of these risk factors can possibly exceed the defensive mechanisms conferred by folic acid supplementation. According to our results, it was also found that the group of women who displayed at least one risk allele of the *MTHFR* variants and had also CS habits presented a higher frequency of SA, compared to non-smoker women.

Women who had previously suffered a SA also presented an increased frequency of the *PAI-1* 4G/5G heterozygous genotype (56.8% *vs.* 47%), compared to women who did not. On the other hand, its homozygous frequency was found to be similar in both groups and even slightly lower among women who had a SA (22.7% *vs.* 24.3%). Despite it appears to be quite proved that higher PAI-1 levels significantly increase the risk of abortion (as well as other obstetric complications)[54,90], PAI-1 levels are however found to be increased during pregnancy[25,89]. This is part of the hemostatic changes that women undergo during the "acquired hypercoagulable state" of pregnancy, in which they acquire reduced fibrinolytic activity, a consequence (also, but not only) of the higher PAI-1 concentration[25]. Therefore, in addition to the normal increase of this fibrinolysis inhibitor, the presence of the *PAI-1* 4G risk allele can play a critical role in SA pathology by causing even higher levels of circulating PAI-1, during the already hypercoagulable state conferred by pregnancy itself. Therefore the overexpression of PAI-1 (and the resulting impaired fibrinolysis) may lead to a compromised and insufficient trophoblast invasion, resulting in an abnormal placental formation[90]. Huang *et al.* (2017) had indeed demonstrated that women who carry the homozygous (4G/4G) genotype had a significantly elevated predisposition to recurrent SA and that this variant conferred an increased risk of recurrent SA for Caucasians[54]. The role of *PAI-1* 4G/5G is however controversial, with some studies reporting significantly higher homozygous genotypic frequency in women with abortion, while others showed no difference regarding this variant prevalence in women who had a SA and those who did not[90]. On the present

study, the homozygous (4G/4G) genotype of *PAI-1* 4G/5G was indeed not found to be increased among participants with abortion. Nevertheless, as previously determined, this variant is quite prevalent among the population, which can justify why women without SA had similar homozygous frequencies compared to the group with SA. In other perspective, the wild-type frequency in the SA group was decreased which indeed shows that women with SA presented a higher frequency of 4G allele carriers.

It is also worth noting that women with AT deficiency are expected to be at particularly high risk for pregnancy-associated thrombosis[104]. However, none of the women participants who had been pregnant (independently or not if they had suffered a SA) were carriers of the *AT* Cambridge II variant and so, although this variant can possible cause adverse pregnancy outcomes, due to its rareness, no associations could be performed in the present study since there were not enough participants to comprise substantial carriers of the *AT* Cambridge II. Finally, in addition to the higher frequency of *MTHFR* C677T homozygous carriers among women smokers who developed SA, it was also found a higher frequency of the *FVL* (9.1% *vs.* 3%) and the *PAI-1 4G/5G* (72.7% *vs.* 51.5%) heterozygous genotype as well as of the *MTHFR* A1298C (27.3% *vs.* 12.1%) homozygous genotype, which were already found to be more present among women who had SA (even without considering CS habits). Regardless the fact that women with these inherited risk factors may have already increased propensity to the incidence of SA, due to all the previously described harmful effects of CS, the possible synergistic effect among those factors cannot be negligible.

4.4 Study Limitations

The e_COR study was only designed for the mainland Portuguese population, not including the Azores and Madeira archipelagos. This way, it is possible that the genotyping profile determined for mainland Portugal may differ from the Portuguese islands. Also, despite the large sample of participants, the inner aim of e_COR was to estimate the prevalence of CVDs risk factors in the population and, so, one main limitation of the present project is the lack of some relevant clinical and personal data regarding thrombosis (such as, the occurrence of venous thrombotic events and some parameters of thrombotic risk influence), which does not allow to establish an individual thrombotic risk. Moreover, due to the rareness of the *AT* Cambridge II variant, the number of participants in this study is not enough to establish associations between these genetic risk factor with other thrombogenic parameters. Also, as this was a study performed for the general population, only a small number of participants from our

sample were diagnosed with thrombotic events, which also limited the strength of the associations performed. Therefore, our highlights of the possible associations established between the genetic variants and the acquired thrombogenic risks in thrombosis onset may act as a template for further studies to validate our results and determine more significant associations in the future.

5. Conclusion and Future Perspectives

The present study provided updated information on the prevalence of six important genetic risk factors for thrombosis in Portugal. The results showed a high frequency of the *PAI-1* 4G/5G, the *MTHFR* C677T and the *MTHFR* A1298C variants among the mainland Portuguese population, whereas the *FVL* and the *FII* G20210A variants were substantial less frequent. The *AT* Cambridge II was found to be almost absent in the general population. Moreover, our results showed that most mainland Portuguese were carriers of at least two different genetic variants and that the population presented a higher frequency of individuals with acquired thrombotic risks. Obesity was found to be the most prevalent (and therefore the most worrisome) but cigarette smoking habits as well as hormonal contraception were also quite frequent among individuals. The presence of the genetic risks *per se*, the population frequencies of individuals with both genetic factors and acquired risk behaviors, as well as the possible associations of such risk factors with thrombosis incidence, needs attention. Overall, our findings suggested that carriers of the *FVL*, *FII* G20210A and *MTHFR* C677T variants, as well as the *PAI-1* 4G/5G homozygous carriers could be at higher risk of developing arterial thrombotic events, especially those who smoke. In the particular case of women, the carriers of *FVL* and *FII* G20210A variants who were also under oral contraception must be at increased risk, but the presence of the *PAI-1* 4G/5G or of the *MTHFR* variants cannot be neglected. Pregnant carriers of the *MTHFR* A1298C variant can also be at high-risk of SA, an event that can also be under influence of *FVL* and *PAI-1* 4G/5G variants. The inheritance of the *MTHFR* C677T homozygous condition can also contribute to the SA pathology among pregnant smokers. Since the interaction of all these risk factors can culminate in thrombotic incidence, awareness about thrombophilia can therefore be critical for prevention, either by elimination of acquired risk factors (lifestyle changes) or to optimize prophylaxis, according to the predicted risk and the clinical settings. Therefore, screening for these thrombogenic variants may be important to consider, whenever a specific phase of life deserves screening (for example, women planning pregnancy or starting contraception).

To our knowledge, this is the first study to report an estimated prevalence of the *FV G1691A* (*FV Leiden*), *FII G20210A*, *AT Cambridge II*, *PAI-1 4G/5G, MTHFR C677T* and *MTHFR A1298C* genetic risk factors for thrombophilia in the mainland Portuguese population. However, it would also be interesting to extent the prevalence of the genetic variants to the Portuguese islands and therefore, not only to estimate their prevalence in the overall Portuguese population, but also to evaluate possible differences regarding

the thrombogenic profile of both archipelagos and mainland individuals. Additionally to what we already determined in the general population, evaluating the frequency of the thrombogenic variants in individuals at already high-risk of CVDs (such as those diagnosed with hypercholesterolemia or dyslipidemias) can also be a next step to establish associations between the risk of developing a thrombotic event under such pathological conditions. Moreover, due to the rareness of the *AT* Cambridge II variant (the main cause of AT deficiency) in the population sample, it would perhaps be useful to also evaluate the levels of AT in the participants; that is, to evaluate the frequency of those who present AT deficiency, regardless the underlying cause. In line with these, another future approach can be determining the participants with low levels of other main anticoagulants, i.e. PC and PS deficiencies. These data would also be important at a clinical level to strength the identification of individuals displaying higher thrombotic risk. Finally, larger scale investigations should also be conducted to determine the clinical implications and contribution (real risk) of each of the thrombogenic variants, with and without the presence of the acquired risks, which may be useful to possibly create a *score* system for decision making, regarding the individuals deserving screening.

6. References

1. Raskob G, Angchaisuksiri P, Blanco A, et al. (2014) Thrombosis: A Major Contributor to Global Disease Burden. *Semin. Thromb. Hemost.* 40, 724–735.
2. Chockalingam A, Thakur J & Varma S (2017) Evolution of noncommunicable diseases: Past, present, and future. *Int. J. Noncommunicable Dis.* 2, 1.
3. Slama S, Kim H-J, Roglic G, et al. (2017) Care of non-communicable diseases in emergencies. *Lancet* 389, 326–330.
4. Chen S, Kuhn M, Prettner K, et al. (2018) The macroeconomic burden of noncommunicable diseases in the United States: Estimates and projections. *PLoS One* 13, e0206702 [Husain MJ, editor].
5. Prüss-Ustün A, van Deventer E, Mudu P, et al. (2019) Environmental risks and non-communicable diseases. *BMJ* 364, l265.
6. Ogoina D & Onyemelukwe GC (2009) The role of infections in the emergence of non-communicable diseases (NCDs): Compelling needs for novel strategies in the developing world. *J. Infect. Public Health* 2, 14–29.
7. World Health Organization (WHO) (2018) Noncommunicable diseases. https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases (accessed August 2019).
8. Piepoli MF, Hoes AW, Agewall S, et al. (2016) 2016 European Guidelines on cardiovascular disease prevention in clinical practice. *Eur. Heart J.* 37, 2315–2381.
9. Fiatal S & Ádány R (2018) Application of Single-Nucleotide Polymorphism-Related Risk Estimates in Identification of Increased Genetic Susceptibility to Cardiovascular Diseases: A Literature Review. *Front. Public Heal.* 5.
10. Asghar M, Kabita S, Kalla L, et al. (2013) Prevalence of *MTHFR*, *Factor V*, *ACE* and *APOE* gene polymorphisms among Muslims of Manipur, India. *Ann. Hum. Biol.* 40, 83–87.
11. World Health Organization (WHO) (2017) Cardiovascular diseases (CVDs). https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) (accessed October 2019).
12. Souto JC & Soria JM (2012) Predicting individual risk of venous thrombosis. *Blood* 120, 500–501.
13. Bhakuni T, Sharma A, Rashid Q, et al. (2015) Antithrombin III Deficiency in Indian Patients with Deep Vein Thrombosis: Identification of First India Based *AT* Variants Including a Novel Point Mutation (T280A) that Leads to Aggregation. *PLoS One* 10, e0121889 [El-Maarri O, editor].
14. Smith SA, Travers RJ & Morrissey JH (2015) How it all starts: Initiation of the clotting cascade. *Crit. Rev. Biochem. Mol. Biol.* 50, 326–336. Informa Healthcare USA, Inc.
15. Soria JM, Morange P, Vila J, et al. (2014) Multilocus Genetic Risk Scores for Venous Thromboembolism Risk Assessment. *J. Am. Heart Assoc.* 3, 1–10.
16. Araújo F, Santos A, Araújo V, et al. (1999) Genetic Risk Factors in Acute Coronary Disease. *Pathophysiol. Haemost. Thromb.* 29, 212–218.
17. Mi Y, Yan S, Lu Y, et al. (2016) Venous thromboembolism has the same risk factors as atherosclerosis. *Medicine (Baltimore).* 95, e4495.
18. Crous-Bou M, Harrington L & Kabrhel C (2016) Environmental and Genetic Risk Factors Associated with Venous Thromboembolism. *Semin. Thromb. Hemost.* 42, 808–820.
19. Prandoni P (2009) Venous and arterial thrombosis: Two aspects of the same disease? *Eur. J. Intern. Med.* 20, 660–661.
20. Jackson SP (2011) Arterial thrombosis—insidious, unpredictable and deadly. *Nat. Med.* 17, 1423–1436. Nature Publishing Group.

21. Yeung J, Li W & Holinstat M (2018) Platelet Signaling and Disease: Targeted Therapy for Thrombosis and Other Related Diseases. *Pharmacol. Rev.* 70, 526–548 [Isom LL, editor].
22. Ginsburg D (2005) Genetic Risk Factors for Arterial Thrombosis and Inflammation. *Hematology* 2005, 442–444.
23. Himabindu G, Rajasekhar D, Latheef K, et al. (2012) *Factor V Leiden* mutation is not a predisposing factor for acute coronary syndromes. *Indian Heart J.* 64, 570–575. Elsevier Ltd.
24. Siegerink B, Maino A, Algra A, et al. (2015) Hypercoagulability and the risk of myocardial infarction and ischemic stroke in young women. *J. Thromb. Haemost.* 13, 1568–1575.
25. Previtali E, Bucciarelli P, Passamonti SM, et al. (2011) Risk factors for venous and arterial thrombosis. *Blood Transfus.* 9, 120–38.
26. Nagareddy P & Smyth SS (2013) Inflammation and thrombosis in cardiovascular disease. *Curr. Opin. Hematol.* 20, 457–463.
27. Sambola A, Fuster V & Badimon JJ (2003) Role of coronary risk factors in blood thrombogenicity and acute coronary syndromes. *Rev. Esp. Cardiol.* 56, 1001–9.
28. Widimský P, Kožnar B, Vaško P, et al. (2013) Acute myocardial infarction and acute stroke: what are the differences? Focus on reperfusion therapy. *Cor Vasa* 55, e111–e116.
29. Parpugga TK, Tatarunas V, Skipskis V, et al. (2015) The Effect of PAI-1 4G/5G Polymorphism and Clinical Factors on Coronary Artery Occlusion in Myocardial Infarction. *Dis. Markers* 2015, 1–8. Hindawi Publishing Corporation.
30. Gregson J, Kaptoge S, Bolton T, et al. (2019) Cardiovascular Risk Factors Associated With Venous Thromboembolism. *JAMA Cardiol.* 4, 163.
31. Anderson FA (2003) Risk Factors for Venous Thromboembolism. *Circulation* 107, 9I--16.
32. Cunha MLR, Meijers JCM, Rosendaal FR, et al. (2017) Whole exome sequencing in thrombophilic pedigrees to identify genetic risk factors for venous thromboembolism. *PLoS One* 12, e0187699 [Garcia de Frutos P, editor].
33. Brito MB, Nobre F & Vieira CS (2011) Contracepção hormonal e sistema cardiovascular. *Arq. Bras. Cardiol.* 96, e81–e89.
34. Jadaon MMM (2011) Epidemiology of Activated Protein C Resistance and *Factor V Leiden* Mutation in the Mediterranean Region. *Mediterr. J. Hematol. Infect. Dis.* 3, e2011037.
35. Ogedegbe HO (2002) An Overview of Hemostasis. *Lab. Med.* 33, 948–953.
36. Machado AI & Lima J (2008) Trombofilias e Contracepção. *Acta Obs. e Ginecológica Port.* 2, 84–95.
37. Smith JW, Day TK & Mackin A (2005) Diagnosing bleeding disorders. *Compendium*, 828–843.
38. Furie B & Furie BC (2008) Mechanisms of Thrombus Formation. *N. Engl. J. Med.* 359, 938–949.
39. Lippi G, Franchini M, Montagnana M, et al. (2012) Inherited disorders of blood coagulation. *Ann. Med.* 44, 405–418.
40. Jin NZ & Gopinath SCB (2016) Potential blood clotting factors and anticoagulants. *Biomed. Pharmacother.* 84, 356–365. Elsevier Masson SAS.
41. Repetto O & De Re V (2017) Coagulation and fibrinolysis in gastric cancer. *Ann. N. Y. Acad. Sci.* 1404, 27–48.
42. Patton LL & DDS (2008) Bleeding and Clotting Disorders. In *Burket's Oral Med. Diagnosis Treat.*, 11th ed., pp. 411–414. Hamilton: BC Decker Inc.
43. Sacher RA (1999) Thrombophilia: a genetic predisposition to thrombosis. *Trans. Am. Clin. Climatol. Assoc.* 110, 51–60; discussion 60-1.
44. Yoshioka FKN, Araújo AG, Tavella MH, et al. (2006) Prevalence of hereditary risk factors for thrombophilia in Belém, Brazilian Amazon. *Genet. Mol. Biol.* 29, 38–40.

45. David D, Ribeiro S, Ferrão L, et al. (2004) Molecular basis of inherited antithrombin deficiency in Portuguese families: Identification of genetic alterations and screening for additional thrombotic risk factors. *Am. J. Hematol.* 76, 163–171.
46. Müller F, Mutch NJ, Schenk WA, et al. (2009) Platelet Polyphosphates Are Proinflammatory and Procoagulant Mediators In Vivo. *Cell* 139, 1143–1156.
47. Gailani D & Renné T (2007) The intrinsic pathway of coagulation: a target for treating thromboembolic disease? *J. Thromb. Haemost.* 5, 1106–1112.
48. Kottke-Marchant K (2002) Genetic polymorphisms associated with venous and arterial thrombosis: an overview. *Arch. Pathol. Lab. Med.* 126, 295–304.
49. Ali N, Ayyub M & Khan SA (2014) Protein C, protein S, *Factor V Leiden* mutation and antithrombin deficiency as a cause of hereditary thrombophilia in patients of venous thromboembolism and cerebrovascular accident. *Pakistan J. Med. Sci.* 30, 1323–1326.
50. Angelopoulou K, Nicolaides A & Constantinou Deltas C (2000) Prevalence of Genetic Mutations That Predispose to Thrombophilia in a Greek Cypriot Population. *Clin. Appl. Thromb.* 6, 104–107.
51. Feddersen S & Nybo M (2014) Discordant diagnoses obtained by different approaches in antithrombin mutation analysis. *Clin. Biochem.* 47, 1337–1340. The Canadian Society of Clinical Chemists.
52. Zeng W, Tang L, Jian X-R, et al. (2015) Genetic analysis should be included in clinical practice when screening for antithrombin deficiency. *Thromb. Haemost.* 113, 262–271.
53. Giurgea G-A, Brunner-Ziegler S, Jilma B, et al. (2017) *Plasminogen activator inhibitor-1* 4G/5G genotype and residual venous occlusion following acute unprovoked deep vein thrombosis of the lower limb: A prospective cohort study. *Thromb. Res.* 153, 71–75. Elsevier Ltd.
54. Huang Z, Tang W, Liang Z, et al. (2017) *Plasminogen Activator Inhibitor-1* Polymorphism Confers a Genetic Contribution to the Risk of Recurrent Spontaneous Abortion: An Updated Meta-Analysis. *Reprod. Sci.* 24, 1551–1560.
55. Moussaoui S, Saussoy P, Ambroise J, et al. (2017) Genetic Risk Factors of Venous Thromboembolism in the East Algerian Population. *Clin. Appl. Thromb.* 23, 105–115.
56. Schmidt C (2018) The fight against non-communicable disease in emerging economies. *Nature* 562, S65–S67.
57. Yapijakis C, Serefoglou Z, Nixon AM, et al. (2012) Prevalence of thrombosis-related DNA polymorphisms in a healthy Greek population. *In Vivo* 26, 1095–101.
58. Zhang G, Tang Y, Tang M, et al. (2010) Antithrombin Cambridge II(A384S) mutation frequency and antithrombin activity levels in 120 of deep venous thrombosis and 150 of cerebral infarction patients in a single center in Southern China. *Blood Coagul. Fibrinolysis* 21, 588–591.
59. Corral J, Hernandez-Espinosa D, Soria JM, et al. (2007) Antithrombin Cambridge II (A384S): an underestimated genetic risk factor for venous thrombosis. *Blood* 109, 4258–4263.
60. Bezemer ID (2008) Gene Variants Associated With Deep Vein Thrombosis. *JAMA* 299, 1306.
61. Abdi AA & Osman A (2017) Prevalence of common hereditary risk factors for thrombophilia in Somalia and identification of a novel *Gln544Arg* mutation in coagulation factor V. *J. Thromb. Thrombolysis* 44, 536–543. Springer US.
62. Poort SR, Rosendaal FR, Reitsma PH, et al. (1996) A common genetic variation in the 3'-untranslated region of the *prothrombin* gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis. *Blood* 88, 3698–703.
63. Meyer MR, Witt DM, Delate T, et al. (2015) Thrombophilia testing patterns amongst patients with acute venous thromboembolism. *Thromb. Res.* 136, 1160–1164. Elsevier Ltd.
64. Al-Allawi NAS, Badi AIA, Goran MA, et al. (2015) The Contributions of Thrombophilic Mutations to

Genetic Susceptibility to Deep Venous Thrombosis in Iraqi Patients. *Genet. Test. Mol. Biomarkers* 19, 500–504.

65. Szpecht D, Gadzinowski J, Seremak-Mrozikiewicz A, et al. (2017) The role of *FV* 1691G>A, *FII* 20210G>A mutations and *MTHFR* 677C>T; 1298A>C and 103G>T *FXIII* gene polymorphisms in pathogenesis of intraventricular hemorrhage in infants born before 32 weeks of gestation. *Child's Nerv. Syst.* 33, 1201–1208.
66. Lucotte G & Mercier G (2001) Population Genetics of *Factor V Leiden* in Europe. *Blood Cells, Mol. Dis.* 27, 362–367.
67. Kujovich JL (1993) Prothrombin-Related Thrombophilia. *GeneReviews®*. http://www.ncbi.nlm.nih.gov/pubmed/20301327 (accessed June 2018).
68. Li C, Ren H, Chen H, et al. (2017) *Prothrombin* G20210A (rs1799963) polymorphism increases myocardial infarction risk in an age-related manner: A systematic review and meta-analysis. *Sci. Rep.* 7, 13550. Springer US.
69. AlSheef M, Almohaya M, Mian A, et al. (2019) Homozygous *prothrombin* gene mutation and ischemic cerebrovascular disease: A case report and review of literature. *J. Appl. Hematol.* 10, 70.
70. Hellfritzsch M & Grove EL (2015) Life-Threatening Contraceptive-Related Pulmonary Embolism in a 14-Year-Old Girl with Hereditary Thrombophilia. *Am. J. Case Rep.* 16, 667–669.
71. Wang D, Cui G, Hu S, et al. (2016) Subtypes of *SERPINC1* mutations and the thrombotic phenotype of inherited antithrombin deficient individuals in Chinese Han population. *Blood Cells, Mol. Dis.* 62, 38–41. Elsevier B.V.
72. Mulder R, Croles FN, Mulder AB, et al. (2017) *SERPINC1* gene mutations in antithrombin deficiency. *Br. J. Haematol.* 178, 279–285.
73. Luxembourg B, Delev D, Geisen C, et al. (2011) Molecular basis of antithrombin deficiency. *Thromb. Haemost.* 105, 635–646.
74. Corral J & Vicente V (2015) Puzzling questions on antithrombin: Diagnostic limitations and real incidence in venous and arterial thrombosis. *Thromb. Res.* 135, 1047–1048. Elsevier Ltd.
75. Águila S, Martínez-Martínez I, Collado M, et al. (2013) Compound heterozygosity involving Antithrombin Cambridge II (p.Ala416Ser) in antithrombin deficiency. *Thromb. Haemost.* 109, 556–558.
76. Deshpande R, Kulkarni B, Ghosh K, et al. (2016) A common missense variant in exon 5 of *antithrombin* gene (*SERPINC1*) in Indian patients with thrombosis. *Thromb. Res.* 143, 1–2. Elsevier Ltd.
77. Picard V, Presot I, Scarabin P-Y, et al. (2007) Antithrombin Cambridge II (A384S): prevalence in patients of the Paris Thrombosis Study (PATHROS). *Blood* 110, 2777–2778.
78. Barba P, Santamaría A, Tirado I, et al. (2008) Antithrombin Cambridge II mutation as a risk factor to develop cerebral venous thrombosis. *Thromb. Haemost.* 99, 443–444.
79. Sanchez C, Alessi MC, Saut N, et al. (2008) Relation between the antithrombin Cambridge II mutation, the risk of venous thrombosis, and the endogenous thrombin generation. *J. Thromb. Haemost.* 6, 1975–1977.
80. Perry DJ, Daly ME, Tait RC, et al. (1998) Antithrombin Cambridge II (Ala384Ser): Clinical, Functional and Haplotype Analysis of 18 Families. *Thromb. Haemost.* 79, 249–253.
81. Wan Ab Rahman WS, Abdullah WZ, Hassan MN, et al. (2017) Familial antithrombin III deficiency in a Malay patient with massive thrombosis. *Malays. J. Pathol.* 39, 197–200.
82. Prabhudesai A, Shetty S, Ghosh K, et al. (2017) Investigation of *Plasminogen Activator Inhibitor-1* (*PAI-1*) 4G/5G promoter polymorphism in Indian venous thrombosis patients: A case-control study. *Eur. J. Haematol.* 99, 249–254.
83. Vuckovic BA, Djeric MJ, Tomic B V., et al. (2018) Influence of decreased fibrinolytic activity and

plasminogen activator inhibitor-1 4G/5G polymorphism on the risk of venous thrombosis. *Blood Coagul. Fibrinolysis* 29, 19–24.

84. Huq MA, Takeyama N, Harada M, et al. (2012) *4G/5G* Polymorphism of the *Plasminogen Activator Inhibitor-1* Gene Is Associated with Multiple Organ Dysfunction in Critically Ill Patients. *Acta Haematol.* 127, 72–80.
85. Song C, Burgess S, Eicher JD, et al. (2017) Causal Effect of Plasminogen Activator Inhibitor Type 1 on Coronary Heart Disease. *J. Am. Heart Assoc.* 6, e004918.
86. Liu Y, Cheng J, Guo X, et al. (2018) The roles of *PAI-1* gene polymorphisms in atherosclerotic diseases: A systematic review and meta-analysis involving 149,908 subjects. *Gene* 673, 167–173. Elsevier.
87. Akhter MS, Biswas A, Abdullah SM, et al. (2017) The Role of *PAI-1* 4G/5G Promoter Polymorphism and Its Levels in the Development of Ischemic Stroke in Young Indian Population. *Clin. Appl. Thromb.* 23, 1071–1076.
88. Ruiz-Quezada S, Vázquez-Del Mercado M, Parra-Rojas I, et al. (2004) Genotype and allele frequency of *PAI-1* promoter polymorphism in healthy subjects from the west of Mexico. Association with biochemical and hematological parameters. *Ann. Génétique* 47, 155–162.
89. Kurzawińska G, Barlik M, Drews K, et al. (2016) Coexistence of *ACE* (I/D) and *PAI-1*(4G/5G) gene variants in recurrent miscarriage in Polish population. *Polish Gynaecol.* 87, 271–276.
90. Đorđević V, Gvozdenov M, Pruner I, et al. (2013) The Prevalence of *PAI-1* 4G/5G Polymorphism in Women with Fetal Loss – First Data For A Serbian Population. *J. Med. Biochem.* 33, 203–207.
91. Sogutlu Sari E, Yazici A, Eser B, et al. (2014) The prevalence of 4G/5G polymorphism of *plasminogen activator inhibitor-1 (PAI -1)* gene in central serous chorioretinopathy and its association with plasma PAI-1 levels. *Cutan. Ocul. Toxicol.* 33, 270–274.
92. Price BR, Wilcock DM & Weekman EM (2018) Hyperhomocysteinemia as a Risk Factor for Vascular Contributions to Cognitive Impairment and Dementia. *Front. Aging Neurosci.* 10.
93. Anderson JAM & Weitz JI (2010) Hypercoagulable States. *Clin. Chest Med.* 31, 659–673. Elsevier Ltd.
94. Tinelli C, Di Pino A, Ficulle E, et al. (2019) Hyperhomocysteinemia as a Risk Factor and Potential Nutraceutical Target for Certain Pathologies. *Front. Nutr.* 6, 1–13.
95. Clément A, Menezo Y, Cohen M, et al. (2018) 5-Methyltetrahydrofolate reduces blood homocysteine level significantly in C677T *methyltetrahydrofolate reductase* single-nucleotide polymorphism carriers consulting for infertility. *Reprod. Biomed. Online*, 8–11.
96. Servy EJ, Jacquesson-Fournols L, Cohen M, et al. (2018) MTHFR isoform carriers. 5-MTHF (5-methyl tetrahydrofolate) vs folic acid: a key to pregnancy outcome: a case series. *J. Assist. Reprod. Genet.* 35, 1431–1435. Journal of Assisted Reproduction and Genetics.
97. Nefic H, MackicDjurovic M & Eminovic I (2018) The Frequency of the 677C>T and 1298A>C Polymorphisms in the *Methylenetetrahydrofolate Reductase* (*MTHFR*) Gene in the Population. *Med. Arch.* 72, 164.
98. Seremak-Mrozikiewicz A, Barlik M, Borowczak P, et al. (2013) The frequency of 677C>T polymorphism of *MTHFR* gene in the Polish population. *Arch. Perinat. Med.* 19, 12–18.
99. Osunkalu VO, Taiwo IA, Makwe CC, et al. (2019) Methylenetetrahydrofolate Reductase Enzyme Level and Antioxidant Activity in Women with Gestational Hypertension and Pre-eclampsia in Lagos, Nigeria. *J. Obstet. Gynecol. India* 69, 317–324. Springer India.
100. Wolski H, Kocięcka M, Mrozikiewicz A, et al. (2015) Coexistence of the 677C>T and 1298A>C *MTHFR* polymorphisms and its significance in the population of Polish women. *Polish Gynaecol.* 86, 742–747.
101. Merviel P, Cabry R, Lourdel E, et al. (2017) Comparison of two preventive treatments for patients with recurrent miscarriages carrying a C677T *methylenetetrahydrofolate reductase* mutation: 5-year

experience. *J. Int. Med. Res.* 45, 1720–1730.

102. Komsa-Penkova R, Golemanov G, Tsankov B, et al. (2017) Rs5918ITGB3 Polymorphism, Smoking, and BMI as Risk Factors for Early Onset and Recurrence of DVT in Young Women. *Clin. Appl. Thromb.* 23, 585–595.
103. Pinjala RK, Reddy LRC, Nihar RP, et al. (2012) Thrombophilia – How Far and How Much to Investigate? *Indian J. Surg.* 74, 157–162.
104. Gerhardt A, Scharf RE, Greer IA, et al. (2016) Hereditary risk factors for thrombophilia and probability of venous thromboembolism during pregnancy and the puerperium. *Blood* 128, 2343–2349.
105. Souto JC, Almasy L, Borrell M, et al. (2000) Genetic susceptibility to thrombosis and its relationship to physiological risk factors: the GAIT study. Genetic Analysis of Idiopathic Thrombophilia. *Am. J. Hum. Genet.* 67, 1452–9.
106. Stewart M & Black K (2015) Choosing a combined oral contraceptive pill. *Aust. Prescr.* 38, 6–11.
107. Beach SRH, Lei MK, Ong ML, et al. (2017) MTHFR methylation moderates the impact of smoking on DNA methylation at AHRR for African American young adults. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* 174, 608–618.
108. Delgado G, Siekmeier R, Grammer TB, et al. (2014) Alterations in the Coagulation System of Active Smokers from the Ludwigshafen Risk and Cardiovascular Health (LURIC) Study. In *Adv. Exp. Med. Biol.*, vol. 857, pp. 9–14.
109. Elkhalifa A (2018) Effects of cigarette smoking on coagulation screening tests and platelet counts in a Sudanese male adults population. *Saudi Med. J.* 39, 897–901.
110. Cheng Y-J, Liu Z-H, Yao F-J, et al. (2013) Current and Former Smoking and Risk for Venous Thromboembolism: A Systematic Review and Meta-Analysis. *PLoS Med.* 10, e1001515 [Lowe G, editor].
111. Yang D, Iyer S, Gardener H, et al. (2015) Cigarette Smoking and Carotid Plaque Echodensity in the Northern Manhattan Study. *Cerebrovasc. Dis.* 40, 136–143.
112. Stefanski A-L, Specker C, Fischer-Betz R, et al. (2018) Maternal Thrombophilia and Recurrent Miscarriage – Is There Evidence That Heparin is Indicated as Prophylaxis against Recurrence? *Geburtshilfe Frauenheilkd.* 78, 274–282.
113. Croles FN, Nasserinejad K, Duvekot JJ, et al. (2017) Pregnancy, thrombophilia, and the risk of a first venous thrombosis: systematic review and bayesian meta-analysis. *BMJ* 359, j4452.
114. Baumann K, Beuter-Winkler P, Hackethal A, et al. (2013) Maternal *Factor V Leiden* and *Prothrombin* Mutations Do Not Seem to Contribute to the Occurrence of Two or More Than Two Consecutive Miscarriages in Caucasian patients. *Am. J. Reprod. Immunol.* 70, 518–521.
115. Lobo RA & Romão F (2011) Hormonas sexuais femininas e trombose venosa profunda. *Angiol. e Cir. Vasc.* 7, 208–214.
116. Wu O, Robertson L, Twaddle S, et al. (2005) Screening for thrombophilia in high-risk situations: a meta-analysis and cost-effectiveness analysis. *Br. J. Haematol.* 131, 80–90.
117. Sacco S, Merki-Feld GS, Ægidius KL, et al. (2017) Hormonal contraceptives and risk of ischemic stroke in women with migraine: a consensus statement from the European Headache Federation (EHF) and the European Society of Contraception and Reproductive Health (ESC). *J. Headache Pain* 18, 108. The Journal of Headache and Pain.
118. Ceviz T, Karakus N, Basol N, et al. (2017) Evaluation of *Protein C* Gene Polymorphism in Patients with Diagnosis of Pulmonary embolism in Turkish Population. *J. Clin. Exp. Investig.* 8, 8–4.
119. Chora JR, Alves AC, Medeiros AM, et al. (2017) Lysosomal acid lipase deficiency: A hidden disease among cohorts of familial hypercholesterolemia? *J. Clin. Lipidol.* 11, 477-484.e2.

120. Branco CC, Pereirinha T, Cabral R, et al. (2009) Thrombotic genetic risk factors and warfarin pharmacogenetic variants in São Miguel's healthy population (Azores). *Thromb. J.* 7, 9.
121. Lahiri DK & Numberger JI (1991) A rapid non-enzymatic method for the preparation of HMW DNA from blood for RFLP studies. *Nucleic Acids Res.* 19, 5444–5444.
122. Applied Biosystems (2008) TaqMan® OpenArray™ Genotyping System. Product Bulletin. California, USA:https://www.ehu.eus/documents/2458249/2712771/Plataforma+TaqMan+OpenArrayeq.pdf (accessed September 2019).
123. Applied Biosystems (2010) TaqMan® OpenArray® Genotyping Getting Started Guide. http://tools.thermofisher.com/content/sfs/manuals/cms_058198.pdf (accessed September 2019).
124. Kowles R (2001) Population Genetics. In *Solving Probl. Genet.*, pp. 219–258. New York, NY: Springer New York.
125. Relethford JH (2012) Hardy-Weinberg Equilibrium. In *Hum. Popul. Genet.*, pp. 23–48. Hoboken, NJ, USA: John Wiley & Sons, Inc.
126. Strachan T & Read A (2011) Factors Affecting Gene Frequencies. In *Hum. Mol. Genet.*, 4th ed, pp. 84–88. New York, NY: Garland Science.
127. Lewis R (2009) Population Genetics. In *Hum. Genet. Concepts Appl.*, 9th ed., pp. 265–306. United States of America: McGraw−Hill Companies.
128. Maia R, Cortesao E, Geraldes C, et al. (2004) 4g/5g Prevalence in 223 Patients with Thrombosis and in 162 Healthy Controls, from the Centre Region of Portugal. *Blood* 104, 4057–4057.
129. Wilcken B, Bamforth F, Li Z, et al. (2003) Geographical and ethnic variation of the 677C>T allele of *5,10 methylenetetrahydrofolate reductase* (*MTHFR*): findings from over 7000 newborns from 16 areas world wide. *J. Med. Genet.* 40, 619–25.
130. Castro R, Rivera I, Ravasco P, et al. (2003) *5,10-methylenetetrahydrofolate reductase* 677C→T and 1298A→C mutations are genetic determinants of elevated homocysteine. *QJM - Mon. J. Assoc. Physicians* 96, 297–303.
131. van Ommen CH & Nowak-Göttl U (2017) Inherited Thrombophilia in Pediatric Venous Thromboembolic Disease: Why and Who to Test. *Front. Pediatr.* 5, 1–6.
132. Stoeva NY & Koleva VS (2016) Homozygous Carrier of *Prothrombin* G20210A Mutation with Massive Pulmonary Embolism and His Family: Gender Differences of Susceptibility to Mutation. *Folia Med. (Plovdiv).* 58, 64–66.
133. Bosler D, Mattson J & Crisan D (2006) Phenotypic Heterogeneity in Patients with Homozygous *Prothrombin* 20210AA Genotype. *J. Mol. Diagnostics* 8, 420–425.
134. Rosendaal FR, Doggen CJM, Zivelin A, et al. (1998) Geographic Distribution of the 20210 G to A *Prothrombin* Variant. *Thromb. Haemost.* 79, 706–708.
135. Jadaon MM (2011) Epidemiology of *Prothrombin* G20210A Mutation in the Mediterranean Region. *Mediterr. J. Hematol. Infect. Dis.* 3, e2011054.
136. Mansilha A, Araújo F, Severo M, et al. (2006) Combined *Factor V Leiden* (R506Q) and *prothrombin* G20210A genotyping in young patients presenting with deep venous thrombosis. *Phlebol. J. Venous Dis.* 21, 24–27.
137. Slatkin M (2004) A Population-Genetic Test of Founder Effects and Implications for Ashkenazi Jewish Diseases. *Am. J. Hum. Genet.* 75, 282–293.
138. Chiasakul T, De Jesus E, Tong J, et al. (2018) Inherited Thrombophilia and the Risk of Arterial Ischemic Stroke: A Systematic Review and Meta-Analysis. *Blood* 132, 2518–2518.
139. Anderson FA (2003) Risk Factors for Venous Thromboembolism. *Circulation* 107, 9I--16.
140. Alfirevic Z, Simundic A-M, Nikolac N, et al. (2010) Frequency of *factor II* G20210A, factor V Leiden,

MTHFR C677T and *PAI-1* 5G/4G polymorphism in patients with venous thromboembolism: Croatian case control study. *Biochem. Medica* 20, 229–235.

141. Siasios I, Fotiadou A, Tsezou A, et al. (2017) A deadly combination of the point mutation A1298C in the *MTHFR* gene , and the 4G/5G polymorphism in the MCA infarct. A case presentation and review of the literature. *SM J Surg*, 1–6.
142. Zhang M-J, Hu Z-C, Yin Y-W, et al. (2014) A Meta-Analysis of the Relationship between *MTHFR* Gene A1298C Polymorphism and the Risk of Adult Stroke. *Cerebrovasc. Dis.* 38, 425–432.
143. Nasiri M, Roostaei A & Ehsanian Z (2014) Association of *Methylenetetrahydrofolate Reductase* (*MTHFR*) Gene C677T and A1298C Polymorphisms with Myocardial Infarction From North of Fars Province. *Res. Mol. Med.* 2, 36–40.
144. Zidovetzki R, Chen P, Fisher M, et al. (1999) Nicotine Increases Plasminogen Activator Inhibitor-1 Production by Human Brain Endothelial Cells via Protein Kinase C–Associated Pathway. *Stroke* 30, 651–655.
145. El-Aziz TAA & Rezk NA (2015) Relation of *PAI-1* and *TPA* Genes Polymorphisms to Acute Myocardial Infarction and its Outcomes in Egyptian Patients. *Cell Biochem. Biophys.* 71, 227–234.
146. Corral J, Hernandez-Espinosa D, Soria JM, et al. (2007) Antithrombin Cambridge II (A384S): an underestimated genetic risk factor for venous thrombosis. *Blood* 109, 4258–4263.
147. Simone B, De Stefano V, Leoncini E, et al. (2013) 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.* 28, 621–647.
148. Hayat D, Mouloud Y & Abdelali B (2018) Frequency and coexistence of the C677T and A1298C polymorphisms of the *MTHFR* gene in Aures region of Algeria. *Biodiversitas J. Biol. Divers.* 19, 1169–1175.
149. Cesari M, Pahor M & Incalzi RA (2010) Plasminogen activator inhibitor-1 (PAI-1): a key factor linking fibrinolysis and age-related subclinical and clinical conditions. *Cardiovasc. Ther.* 28, e72-91.
150. Robertson L, Wu O, Langhorne P, et al. (2006) Thrombophilia in pregnancy: a systematic review. *Br. J. Haematol.* 132, 171–196.
151. Munshi R, Panchal F, Kulkarni V, et al. (2019) *Methylenetetrahydrofolate reductase* polymorphism in healthy volunteers and its correlation with homocysteine levels in patients with thrombosis. *Indian J. Pharmacol.* 51, 248.
152. Bassuk SS & Manson JE (2015) Oral contraceptives and menopausal hormone therapy: relative and attributable risks of cardiovascular disease, cancer, and other health outcomes. *Ann. Epidemiol.* 25, 193–200.
153. Melo A, Reis R, Ferriani R, et al. (2017) Hormonal contraception in women with polycystic ovary syndrome: choices, challenges, and noncontraceptive benefits. *Open Access J. Contracept.* Volume 8, 13–23.
154. Kujovich JL (2004) Hormones and pregnancy: thromboembolic risks for women. *Br. J. Haematol.* 126, 443–454.
155. Rott H (2014) Prevention and treatment of venous thromboembolism during HRT: current perspectives. *Int. J. Gen. Med.* 7, 433–40.
156. Straczek C, Oger E, Yon de Jonage-Canonico MB, et al. (2005) Prothrombotic Mutations, Hormone Therapy, and Venous Thromboembolism Among Postmenopausal Women. *Circulation* 112, 3495–3500.
157. Sosic GM, Sretenovic S, Radivojevic D, et al. (2017) The Impact of the Gene Variants *FV Leiden*, *FII*

G20210A, *MTHFR* C677T and *PAI-1* 4G/5G on Pregnancy Loss in Women from Central Serbia. *Serbian J. Exp. Clin. Res.* 82, 1471–8.

158. Nair RR, Khanna A, Singh R, et al. (2013) Association of maternal and fetal *MTHFR* A1298C polymorphism with the risk of pregnancy loss: a study of an Indian population and a meta-analysis. *Fertil. Steril.* 99, 1311-1318.e4. Elsevier Inc.
159. Dell'Edera D, L'Episcopia A, Simone F, et al. (2018) *Methylenetetrahydrofolate reductase* gene C677T and A1298C polymorphisms and susceptibility to recurrent pregnancy loss. *Biomed. reports* 8, 172–175.
160. Hwang KR, Choi YM, Kim JJ, et al. (2017) *Methylenetetrahydrofolate Reductase* Polymorphisms and Risk of Recurrent Pregnancy Loss: a Case-Control Study. *J. Korean Med. Sci.* 32, 2029.
161. Madduru D, Supriya K, Dirsipam K, et al. (2019) Relation between *MTHFR* (*Methylenetetrahydrofolate Reductase*) gene functional polymorphisms and magnitude of Recurrent Pregnancy Losses (RPL) among women with low socio-economic status in Asian Indians. *Genet. Mol. Res.* 18.
162. Yang Y, Luo Y, Yuan J, et al. (2016) Association between maternal, fetal and paternal *MTHFR* gene C677T and A1298C polymorphisms and risk of recurrent pregnancy loss: a comprehensive evaluation. *Arch. Gynecol. Obstet.* 293, 1197–1211. Springer Berlin Heidelberg.