



Genetic association study of fatal pulmonary embolism

Lisa Meißner¹ · Peter Schürmann² · Thilo Dörk² · Lars Hagemeier¹ · Michael Klintschar¹ 

Received: 16 June 2020 / Accepted: 7 October 2020 / Published online: 30 October 2020
© The Author(s) 2020

Abstract

Pulmonary embolism (PE) is a complex multi-factorial disease and represents one manifestation of venous thromboembolism (VTE). Most commonly PE constitutes a complication of VTE's other clinical presentation deep vein thrombosis (DVT). The majority of studies concerning risk factors do not distinguish between PE and DVT. The risk factors are often estimated to be alike, but the prevalence and the risk associated with the major genetic factor Factor V Leiden differ between the two disease states. We have investigated the association of 22 SNPs with PE in 185 PE case and 375 healthy control subjects. At $p = 0.05$, eight SNPs presented with nominally significant evidence of association (EOA), although no significantly different genotype distributions remained between cases and controls after Bonferroni correction. Three of these variants (rs1800790, rs3813948, rs6025) showed EOA in the main analysis, and five variants (rs169713, rs1801131, rs4524, rs5985 and rs8176592) demonstrated EOAs in subgroups. Genomic variation modulating Factor V, Factor XIII, Beta fibrinogen (FGB), TFPI or HIVEP1 should be worth to be followed in subsequent studies. The findings of this study support the view that PE represents a complex disease with many factors contributing relatively small effects. Larger sample sizes will be required to reliably detect these small effects.

Keywords Pulmonary embolism · Venous thromboembolism · Deep vein thrombosis · Sudden death

Introduction

Pulmonary embolism (PE) is one manifestation of venous thromboembolism (VTE) and a potentially lethal complication of VTE's other clinical presentation deep vein thrombosis (DVT). Approximately 65% of VTE patients display only DVT, whereas around 30% of VTE patients also manifest PE [1].

PE diagnosis is challenging due to its variable and often non-specific clinical presentation. Silent PE develops in around 50% of patients with DVT and autopsy reports demonstrated that only 30–45% of all PE cases were diagnosed prior to death. The annual incidence of PE is estimated to be

about 50–100 per 100,000 individuals and increases with age [2, 3]. Males are more often affected than females. However, women above 75 years display an incidence of more than 500 cases per 100,000 individuals [4]. In forensic practice, pulmonary embolism is relatively common, either as sudden unexpected death or in the context of potential maltreatment. In our own praxis, we observe pulmonary embolism in ca. 4% of all medicolegal autopsies.

Pulmonary embolism is a complex multi-factorial disease which is caused by external as well as multiple genetic factors. However, most studies on risk factors were conducted analysing associations with VTE and not PE in particular [5]. External risk factors can be divided into permanent, patient-associated risk factors and, temporary circumstances like immobilisation [2]. In 20–50% of the cases, PE occurs in absence of these risk factors [6, 7]. Moreover, twin and family studies suggested a major genetic constituent risk to be related to VTE by demonstrating an increased risk for individuals with affected siblings and a strong heritability [8–10]. The genetic predisposition to elevated risk of VTE is described by the term thrombophilia, which is often associated with gene variations of the physiological coagulation cascade [11]. Minor changes in the balance between the system of coagulation and fibrinolysis might cause thrombus formation. These changes can occur through mutations in genes

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00414-020-02441-7>) contains supplementary material, which is available to authorized users.

✉ Michael Klintschar
Klintschar.Michael@mh-hannover.de

¹ Institute of Legal Medicine, Hannover Medical School, Carl-Neuberg-Str.1, 30625 Hannover, Germany

² Gynaecology Research Unit, Department of Obstetrics and Gynaecology, Hannover Medical School, Carl-Neuberg-Str.1, 30625 Hannover, Germany

coding for coagulation factors and inhibitors, which lead to altered levels or functional losses of associated proteins [12].

Until the beginning of 2000, only few gene loci were robustly identified to be associated with VTE. Most of these variations represent strong genetic risk factors increasing the risk of VTE in heterozygous carriers almost 10-fold, such as deficiencies of antithrombin or protein C [5, 13]. However, these genotypes occur in less than 1% of the population and thus represent only few patients suffering from VTE [14]. Subsequently, new loci, presenting weaker risk factors but higher frequencies, were identified through studies of candidate genes [15] and genome-wide association studies (GWAS) [14, 16–18].

Since PE is usually regarded as a complication of DVT, their inherited risk factors are often regarded to be similar, but the prevalence and the risk associated with some risk factors such as Factor V Leiden or activated protein C resistance differ between DVT and PE [19, 20]. The Factor V Leiden paradox states that the rs6025 polymorphism increases the risk of DVT much more than the risk of PE [20]. These findings emphasize the necessity of considering DVT and PE as different phenotypes in VTE genetic studies. Only one study investigating a broad panel of genetic VTE risk factors in patients with fatal PE has been reported so far [21]. Brandimarti et al. conducted a study in an Italian population investigating associations between 12 thrombophilic risk factors and VTE in 44 individuals who died from PE. Strong associations between PE and the polymorphisms C677T (MTHFR), Pai-1 4G/5G and ITGB3 were found [21].

The present study aims to replicate the results of the aforementioned study using a considerably larger sample size. Furthermore, this study intends to evaluate the association of additional SNPs which were robustly associated with VTE or DVT in previous studies in subjects who died from PE.

Material and methods

Investigated subjects

Patients who died from PE were identified by a review of the autopsy reports of the Department of Forensic Medicine at Hannover Medical School and its remote station in Oldenburg. The case group comprised samples from 185 Caucasian adults (mean age 62 years) for whom PE was diagnosed post mortem. The group comprises samples from 90 male and 92 female subjects and three individuals without documented gender. For four samples, no information was given about the age. Concerning VTE risk factors, 90% of all case samples presented at least one external risk factor. Of all samples, 51% exhibit at least one of the risk factors surgery, tumour or obesity. Blood samples on filter paper or cotton swabs were obtained during autopsy of the PE cases and stored with the respective file at room temperature.

The control group was composed of 375 samples from healthy Caucasian individuals (mean age 36 years) which had been voluntary blood donors at the Transfusion Department of Hannover Medical School. DNA extracts were obtained from the Hannover Breast Cancer Study and Hannover Prostate Cancer Study of the Gynaecology Research Unit (Hannover Medical School). The group comprised samples from 189 female and 186 male blood donors.

DNA extraction

DNA was extracted using the QIAmp®DNA Mini Kit (QIAGEN, Hilden, Germany). In order to achieve from the filter paper the recommended DNA concentration of 10 ng/ml for SNPtype genotyping on the Fluidigm Dynamic Array integrated fluidic circuits (IFCs), pilot experiments using different amounts of elution buffer were conducted. Results were compared utilizing Thermo Scientific NanoDrop 8000. Best reproducible DNA yields were reached eluting with 50 µl and three times repeated elutions.

Variant selection

A total of 27 candidate gene variants were selected based on previous literature. Twelve SNPs used in the study conducted by Brandimarti et al. [21] were chosen in order to examine whether their results can be replicated in our larger sample. The National Center for Biotechnology Information (NCBI) PubMed database was searched for papers assessing other relevant VTE and PE associated SNPs or SNPs affecting the coagulation of blood and their prevalence in the general population as well as in VTE-exposed populations. Used search terms were genetics, risk factors, thrombophilia, venous thrombosis, pulmonary embolism and the polymorphisms themselves (date of research: May–July 2018). Due to power calculations, only SNPs exhibiting a minor allele frequency (MAF) higher than 1% were chosen. For 26 of 27 initially chosen SNPs, SNPtype genotyping assays using the Fluidigm® approach could be successfully designed. These SNPs were tested for proper clustering utilizing samples from healthy controls for genetic studies of the Gynaecology Research Unit (Hannover Medical School). SNPs displaying poor clustering in the wet lab testing were excluded from further analyses. Altogether 22 SNPs, including nine of those SNPs tested by Brandimarti et al. [21], could be successfully genotyped. Table 1 displays the SNPs used in this study.

Genotyping

Amplification and genotyping were performed using allele-specific SNPtype assays on 192.24 Dynamic Array IFCs using the BioMark EP1 Real-Time PCR System (Fluidigm Corp., South San Francisco, CA, USA) as previously

Table 1 Single nucleotide polymorphisms associated with VTE or hereditary thrombophilia

Gene	refSNP accession no.	Effect allele	Reference publication
<i>MTHFR</i>	rs1801133	T	Brandimarti et al. [21]
<i>MTHFR</i>	rs1801131	C	Brandimarti et al. [21]
<i>F5</i>	rs6025	A	Brandimarti et al. [21]
<i>F2</i>	rs1799963	A	Brandimarti et al. [21]
<i>FGB</i>	rs1800790	A	Brandimarti et al. [21]
<i>F13A</i>	rs5985	A	Brandimarti et al. [21]
<i>ITGB3</i>	rs5918	C	Brandimarti et al. [21]
<i>TFPI</i>	rs8176592	C	Brandimarti et al. [21]
<i>TFPI</i>	rs10931292	C	Brandimarti et al. [21]
<i>F5</i>	rs4524	G	Germain et al. [22]
<i>FGG</i>	rs2066865	T	Germain et al. [16]
<i>FGG</i>	rs6536024	C	Tang et al. [18]
<i>ABO</i>	rs8176719	Deletion	Heit et al. [17]
<i>ABO</i>	rs2519093(v2)	A	Heit et al. [17]
<i>ABO</i>	rs529565(v2)	C	Germain et al. [22]
<i>F11</i>	rs4253399	G	Tang et al. [18]
<i>F11</i>	rs4253417	C	Germain et al. [22]; Sennblad et al. [23]
<i>F11</i>	Rs2036914	C	Austin et al. [24]; De Haan et al. [25]
<i>SLC44A2</i>	rs2288904	T	Germain et al. [22]
<i>HIVEP1</i>	rs169713	C	Morange et al. [26]
<i>C4BPB</i>	rs3813948	G	Buil et al. [27]
<i>TC2N</i>	rs1884841	T	Morange et al. [28]

described [29]. In short, assays composed of FAM or HEX fluorescence-labelled primer, detecting respectively one allele of each SNP, and an untagged locus-specific reverse primer were designed by Fluidigm for each SNP [30] (online supplementary material: Table 1).

Specific target amplification (STA) as a preamplification step was performed for all case and control samples as well as two non-template controls per array in a 96-well thermocycler using 10× SNPtype STA Primer Pool, 2× Multiplex PCR Master Mix (Qiagen, Hilden, Germany) and genomic DNA. Thermocycling conditions were 15 s at 95 °C followed by 14 cycles of 15 s at 95 °C and 4 min at 60 °C. The STA products were then diluted 1:100 in DNA suspension buffer, a low-EDTA TE buffer (10 mM Tris, 0.1 mM EDTA; pH 8.0), and stored at −20 °C.

The Sample Pre-Mix per 192 samples was prepared by combining 540 µl Biotium 2x Fast Probe Master Mix, 54 µl 20x SNPtype Sample Loading Reagent, 18 µl 60x SNPtype Reagent, 7 µl ROX and 12 µl PCR-certified water. After vortexing and centrifugation, the sample mix for each sample was prepared by adding 2.6 µl Sample Pre-Mix to 2 µl of the diluted STA product. For each SNP, the 10x Assay mix was prepared by 2 µl 2× Assay Loading Reagent (Fluidigm), 1.2 µl PCR-certified water, and 1 µl SNPtype assay mix (Fluidigm).

The Dynamic Array IFCs were each loaded with 190 times 3 µl prepared sample pre-mix, two NTCs and the 10x assays. Loading of the array IFC as well as mixing and thermocycling followed the manufacturer's instructions. Mixing of the samples and assays was conducted with the IFC Controller RX (Fluidigm). Thermocycling was performed on the BioMark EP1 Real-Time PCR System and run according to the default protocol.

In total, five Dynamic Array IFCs were run. For quality control, 16.5% of all samples were run in duplicates.

Data analysis

Resulting data from the array runs was displayed, automatically analysed and manually checked using the data analysis software Fluidigm SNP Genotyping Analysis [30]. Samples with more than 4 “no calls” or less than 90% concordance with their duplicates were eliminated from the study ($n = 26$). SNPtype assays were checked for call rates below 90% or poor clustering (none excluded). Moreover, Hardy-Weinberg equilibrium (HWE) was examined using the χ^2 tests. Assays displaying significant deviations were also omitted ($n = 2$) (rs2227589, rs710446).

Statistical analyses were conducted using logistic regression with STATA® 12.0 on distinct results for 22 SNPs, overall and within subgroups. Genotype frequencies from cases and controls were compared and tested for significant differences using a univariate linear regression analysis with case-control status as the outcome variable under an additive risk model. In addition, a stratified analysis according to known environmental risk factors of the case samples was conducted. For this purpose, samples were stratified in eight different subcategories in order to filter out those case samples of patients whose death is more likely related to environmental risk factors than genetic risk factors and thus increase the proportion of case samples that died because of a stronger genetic influence. These categories included age over 50, sex, body-mass index (BMI) above 30, recent surgery and tumours. We tested all cases without known risk factors against all controls, as well as all cases without tumour, operation and BMI below 30 against all controls. Results were expressed as odds ratios (OR) with 95% confidence intervals (CI) and two-sided *p*-values. In case of zero fields for a particular SNP, we tested for significant differences using the Fisher's exact test.

In order to avoid increased false positivity rates due to multiple comparisons [31], this study also utilized the Bonferroni correction for all *p* values. The significance level after multiple testing was considered 0.002 [32].

Mantel-Haenszel pooled OR after a fixed-effect meta-analysis was conducted for the combined data set of Brandimarti et al. [21] and our study results using the metan command in STAT12.0 with an increment of 0.1 to account for zero fields. As the meta-analysis comprised nine different variants, the level of statistical significance was adjusted accordingly ($\alpha = 0.006$).

Results

Genotyping achieved call rates between 95.7% (rs8176592) and 99.1% (rs3813948) as well as an overall call rate of 98.1% for the 560 samples that were genotyped for 22 SNPs.

At a significance level of 0.05, eight of 22 SNPs presented nominally significant results in either the main or the eight subgroup analyses before the correction for multiple testing. After utilizing the multiple-comparison Bonferroni correction for the main analysis ($p = 0.0017$), no significantly different genotype distributions remained between cases and controls. However, $p \leq 0.05$ was regarded in this study as an evidence of association (EOA) suggesting noteworthiness for the investigation of causes for PE. The results for those variants demonstrating $p < 0.05$ in the logistic regression analyses are depicted in Table 2. Three of 22 SNPs demonstrated EOAs in the main analysis (rs1800790, rs3813948, rs6025). The five additional polymorphisms rs169713, rs1801131, rs4524, rs5985 and rs8176592 presented EOA in subgroup analyses.

Table 2 Selected results from the logistic regression for the main- and subgroup analyses according to risk factors

Analyse	Gene	Marker	OR	<i>p</i> value	95% confidence interval	
Main analysis						
All cases vs. all controls						
	<i>FGB</i>	rs1800790	0.7078	0.028	0.52	0.96
	<i>C4BPB</i>	rs3813948	0.4592	0.016	0.24	0.87
	<i>F5</i>	rs6025	2.1312	0.010	1.2	3.79
Subgroup analyses						
Male cases vs. all controls						
	<i>HIVEP1</i>	rs169713	1.5650	0.042	1.02	2.41
	<i>F5</i>	rs4524	0.6323	0.035	0.41	0.97
Female cases vs. all controls						
	<i>FGB</i>	rs1800790	0.5906	0.027	0.37	0.94
	<i>C4BPB</i>	rs3813948	0.3305	0.028	0.12	0.89
	<i>F5</i>	rs6025	2.7021	0.013	1.23	5.93
Cases with PE at age < 50 years vs. all controls						
	<i>F13A</i>	rs5985	0.4676	0.009	0.27	0.83
	<i>FGG</i>	rs6536024	0.5552	0.013	0.35	0.88
Cases with BMI > 30 vs. all controls						
	<i>TFPI</i>	rs10931292	0.4020	0.019	0.19	0.86
Cases without known risk factors vs. all controls						
	<i>MTHFR</i>	rs1801131	0.3456	0.050	0.12	1.00
	<i>TFPI</i>	rs8176592	0.3298	0.039	0.12	0.95
All cases without surgery vs. all controls						
	<i>C4BPB</i>	rs3813948	0.3906	0.012	0.19	0.81
	<i>F5</i>	rs6025	2.2277	0.009	1.22	4.07
All cases without tumour vs. all controls						
	<i>FGB</i>	rs1800790	0.7081	0.035	0.51	0.98
	<i>C4BPB</i>	rs3813948	0.4704	0.024	0.24	0.91
	<i>F5</i>	rs6025	2.2821	0.006	1.27	4.09
All cases without operation, tumour and with BMI below 30 kg/m ² vs. all controls						
	<i>HIVEP1</i>	rs169713	1.7373	0.006	1.18	2.57
	<i>FGB</i>	rs1800790	0.6049	0.022	0.39	0.92
	<i>TC2N</i>	rs1884841	0.6326	0.012	0.44	0.91
	<i>F5</i>	rs6025	2.6133	0.007	1.3	5.26

Bonferroni correction was used to correct for multiple comparisons. After correction, no SNPs remained nominally significant. $\alpha = 0.05$ was used as an evidence of association. For rs1799963, no results were obtained in the subgroups “cases with PE before 50 years vs. all controls” and “cases without known risk factors vs. all controls” because none of the residual case samples showed deviations from the wild-type allele and the odds ratios could not be calculated with zero fields

Values ranged between $p = 0.006$ (rs169713 in “all cases without surgery, tumour or BMI below 30 kg/m² vs. all controls”; rs6025 in “all cases without tumour vs. all controls”) and $p = 0.05$ (rs1801131 in “cases without known risk factors vs. all controls”). EOAs were found for rs169713 in two subgroups, for rs1800790 in four, for rs1801131 in one, for rs3813948 in four, for rs4524 in one, for rs8176592 in one and for rs6025 in

five subgroups (Table 2). When being compared to literature [33], the MAF of our control group presented deviations ranging from 0.043 (rs1801133) to 11.2% (rs169713) (online supplementary material: Table 1).

A meta-analysis of our results with previously published results from Brandimarti et al. [21] showed for some previously proposed SNPs opposing or strongly deviating effects, with significant heterogeneity between studies for rs8176592 (p_{het} 0.028) and rs5918 (p_{het} 0.001). For rs1801133 and rs10931292, both studies showed similar tendencies though the meta-analysis did not reach statistical significance (Table 3). From our own study, *FGB* (rs1800790, p_{meta} = 0.068) and *F5* (rs6025, p_{met} = 0.064) remained with marginal evidence of association.

Discussion and conclusion

Clinical medicine as well as forensic post mortem investigations would benefit from well-defined genetic risk factors for PE. However, almost all studies on the association of gene loci with VTE up to now do not differentiate between PE and DVT [24, 27, 34]. There is however a big difference in consequences for a patient between DVT and the potentially deleterious complication of PE. Furthermore, it is well known that, fortunately, only a minority of patients with increased DVT risk, e.g. from Factor V Leiden, succumb from PE, a phenomenon described as Factor V Leiden paradox. On the other hand, studies on risk factors for PE are mostly focused on external risk factors and not on candidate genes from studies on VTE, which results in an insufficient knowledge concerning their influence on PE (e.g. [35]).

The only study that focuses on genetic variants in deaths from PE we are aware of is that by Brandimarti et al. [21] that found significant differences between a group of persons deceased from PE and a control group for SNPs in *MTHFR**C6777T (rs1801133), *SERPINC1**4G/5G (rs397832688) and *ITGB3**T196C (rs5918). However, the groups tested in this study were small: 44 persons who died from PE and 102 controls. While rs397832688 was not amenable to our SNPtype assay design, our study of 375 controls

and 185 cases could not reproduce a significant difference in genotype frequencies for the two SNPs *MTHFR**C6777T (rs1801133) and *ITGB3**T196C (rs5918), respectively, between PE cases and controls reported by Brandimarti et al. [21]. These differences in results may be explained by the small number of PE cases and controls in the initial study [21]. Small sample sizes may lead to overestimations of the particular effect size [36]. The overestimation of the effect size can, in turn, result in overestimations of the power of replication studies, which might have affected the a priori calculated power for a few variants of the present study [37]. For instance, the statistical power for rs5918 in our study design was determined a priori on the basis of the genotyping results of Brandimarti et al. [21] as being close to 100% to detect a significant difference at $p < 0.05$ with the given risk estimate and study size. However, the statistical power for the rs5918 polymorphism based on the truly obtained genotype frequencies amounts to 16%. The post hoc calculated power for the other SNPs tested in our study ranges between 1 (rs1801133) and 62% (rs1800790). This emphasizes that, according to Ziegler et al. [38], extensive sample sizes are necessary for a reliable detection of the small effects which are expected for the contributing factors in complex diseases like PE. Most of the variants that we analysed showed only small effects (see OR and their 95% CI in Table 2). In comparison to Brandimarti et al. [21], we used approximately the 4-fold number of cases as well as controls, but for a reliable determination and replication of the small effect sizes, the study size still would need to be strongly increased.

In our study, rs1800790, rs3813948 and rs6025 demonstrated EOA for the main logistic regression analysis. The Factor V Leiden mutation (rs6025) displayed similar effects in the subgroup analyses indicating a rather small effect of stratification. According to the FV Leiden paradox, the risk of DVT is substantially increased in carriers of the *FV* polymorphisms, whereas the risk for PE is only mildly increased [39]. The rs6025 was not found in any of the 44 cases studied by Brandimarti et al.; however, Van Stralen et al. [39] demonstrated in carriers of the Factor V Leiden mutation an

Table 3 Meta-analysis with data from Brandimarti et al. [21]

Variant ID	Brandimarti et al. [21]	This study	Meta OR (95% CI)	p_{meta}	p_{het}
rs1801133	0.879 (0.522; 1.479)	0.910 (0.698; 1.186)	0.903 (0.713; 1.144)	0.398	0.907
rs1801131	0.517 (0.294; 0.911)	0.993 (0.756; 1.304)	0.872 (0.684; 1.113)	0.273	0.042
rs6025	0.032 (0.000; 16.266)	2.114 (1.196; 3.734)	1.660 (0.971; 2.836)	0.064	0.134
rs1799963	0.032 (0.000; 16.266)	0.867 (0.330; 2.274)	0.600 (0.241; 1.490)	0.271	0.234
rs1800790	2.398 (0.676; 8.501)	0.712 (0.524; 0.966)	0.758 (0.564; 1.018)	0.068	0.068
rs5985	1.988 (0.590; 6.694)	1.042 (0.793; 1.369)	1.073 (0.822; 1.401)	0.603	0.309
rs5918	9.571 (2.600; 35.238)	0.897 (0.626; 1.285)	1.107 (0.794; 1.545)	0.548	0.001
rs8176592	1.701 (1.013; 2.857)	0.883 (0.673; 1.159)	1.014 (0.798; 1.289)	0.907	0.028
rs10931292	0.619 (0.257; 1.489)	0.791 (0.534; 1.170)	0.758 (0.530; 1.083)	0.128	0.617

approximately 8-fold increased risk of DVT (OR 7.7, 95% CI 3.9–15.3), whereas the risk of PE was small (OR 1.4, 95% CI 0.7–2.7). Our study is consistent with an about 2-fold increased risk of PE in carriers of the mutation which supports a role of this mutation for PE but is well in line with the results from the aforementioned study.

The *FGB* polymorphism 455 G/A (rs1800790) has been associated with higher levels of fibrinogen, which in turn suggests an association of the variant with a hypercoagulable state as well as thrombosis [40]. Studies could not agree whether the polymorphism exerts a protecting role or none at all. Whereas some studies [40, 41] detected a lower VTE risk to be associated with the polymorphism, another [42] found no association between the mutation and PE. Brandimarti et al. [21] could not demonstrate significant differences, but our study demonstrated an EOA for the polymorphism in the main analysis and three subgroup analyses (see Table 2). The computed ORs indicate a protective effect of the rare allele.

The protein C4BPB binds to protein S and thus presents a main element of the coagulation and fibrinolysis cascade. A previous study demonstrated a slightly increased risk associated with the polymorphism rs3813948 located in the *C4BPB* gene [27]. However, the results from our study indicate the opposite effect, demonstrating a rather protective effect of the polymorphism. A possible explanation for the opposing results is that Buil et al. [27] analysed the association with VTE without differentiation in DVT and PE cases, whereas our study examined those polymorphisms only in PE deaths. In this case, the results would suggest strongly differing risks associated with the polymorphism for PE and DVT, which need to be examined in further studies.

The polymorphisms at rs169713, rs1801131, rs4524, rs5985 and rs8176592 presented EOA in the subgroup analyses that we performed in order to maximize the effect of the genetic variant by enriching the susceptible group through exclusion of externally influenced cases, as proposed [38].

Franco et al. [43] demonstrated a protective role of the rare SNP allele for the Factor XIII p.Val34Leu polymorphism (rs5985). The subjects who did not harbour the rare allele presented a 6-fold increased risk of developing DVT. Our study results suggest the rare allele of this polymorphism to exert a protective role not only for DVT but possibly also for PE.

Lincz et al. [44] examined the activity of the TFPI and frequencies of *TFPI* polymorphisms in patients presenting Factor V Leiden and a manifestation of VTE (rs8176592). They found a strongly increased VTE risk in carriers of the polymorphism in the presence of Factor V Leiden. However, most other studies demonstrated a small protective effect of the polymorphism on the VTE risk by increasing the TFPI plasma level [32]. In the present study, we replicated the protective effect of the T33C mutation. Individuals without environmental risk factors, which do not exhibit the cysteine variant, displayed an about 3-fold increased risk of dying of PE.

As the first VTE susceptibility locus outside the traditional coagulation cascade, the rs169713 polymorphism of the *HIVEP1* locus was significantly associated with VTE risk in a study by Morange et al. [26]. *HIVEP1* affects the transcriptional regulation of target genes associated with inflammatory processes [5]. By increasing the damage on endothelial cells and enhancing hypercoagulability, inflammation is considered to act as a contributor to the VTE risk [45]. The results of our study support the aforementioned result, by demonstrating notable differences in two subgroup analyses. Moreover, our results suggest that the rare allele is associated with a slightly higher risk of PE than VTE in general.

Associations of the A1298C (rs1801131) and C677T (rs1801133) variants of the *MTHFR* gene and an increased VTE risk were found by Liu et al. [46]. Brandimarti et al. [21] examined these two polymorphisms but their results did not demonstrate significant differences of the polymorphisms in cases and controls. Furthermore, a study on a Chinese population [47] demonstrated significant results for the C677T variant of the *MTHFR* gene and VTE. Our study demonstrated evidence for differences in genotype frequencies of cases and controls for the A1298C variant within the subgroup analysis of cases without known risk factors. The C-allele yielded a protective effect on the PE risk. Our results suggest that subjects without known risk factors that do not exhibit the C-allele may display an almost 3-fold increased risk of developing PE.

The rs4524 polymorphism of the Factor V gene was associated in a meta-analysis of more than 65,000 individuals with a slightly increased risk for VTE [22]. Moreover, the effect allele was repeatedly associated with an elevated risk of DVT in three Dutch studies conducted by Bezemer et al. [48]. Our results suggest a rather decreased risk, if any, of PE for carriers of the effect allele in male individuals. These results might be in line with the aforementioned Factor V Leiden paradox. Nevertheless, the paradox was only observed as well as analysed for the rs6025 polymorphism [39] and needs to be investigated for the rs4524 mutation.

The results of our study, the largest in terms of sample size ever conducted solely on deaths from PE, may indicate that the risk of death from PE is—at least in part—determined by a number of gene variants. As the evaluation of the impact of these variants is far from conclusively resolved, further analyses in even larger samples would be rewarding. We followed Ziegler et al. [38] who suggested that trying to simplify the respective disease might improve the outcome of association studies. The authors suggest reducing the diseases' complexity by choosing subgroups of case individuals which are more homogenous. We already demonstrated that for some SNPs, EOAs were only achieved in subgroup analyses. Further increase of the case group would also lead to increased subgroups and thus more distinct results. Moreover, it would be useful to select preferably non-anonymous age-matched controls, who did not develop the disease despite strong external exposure to risk factors.

In summary, our relatively large case-control genetic association study provided supportive evidence for genetic differences at eight candidate risk loci between cases with death from PE and controls, whereas it has refuted some others. Genomic variation modulating Factor V, Factor XIII, FGB, TFPI or HIVEP1 should be worth to be followed in subsequent studies. The results demonstrate evidence for an association; however, more investigations are necessary to prove a causal relationship. Therefore, sample sizes could further be increased and the number of independent tests could be minimized by focusing on the genetic variants demonstrating EOAs. From the forensic point of view, as already mentioned by Brandimarti et al. [21], genetic testing might be especially discriminatory in cases of medical malpractice.

Acknowledgements Open Access funding enabled and organized by Projekt DEAL.

Compliance with ethical standards This study was approved by the local ethics committee at Hannover Medical School.

The authors would like to thank the Exome Aggregation Consortium and the groups that provided exome variant data for comparison. A full list of contributing groups can be found at <http://exac.broadinstitute.org/about>.

The authors would also like to thank the Genome Aggregation Database (gnomAD) and the groups that provided exome and genome variant data to this resource. A full list of contributing groups can be found at <https://gnomad.broadinstitute.org/about>.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- White RH (2003) The epidemiology of venous thromboembolism. *Circulation* 107:14–18. <https://doi.org/10.1161/01.CIR.0000078468.11849.66>
- Bělohávek J, Dytrych V, Linhart A (2013) Pulmonary embolism, part I: epidemiology, risk factors and risk stratification, pathophysiology, clinical presentation, diagnosis and nonthrombotic pulmonary embolism. *Exp Clin Cardiol* 18:129–138
- Torbicki A, Perrier A, Konstantinides S, Agnelli G, Galiè N, Pruszczyk P, Bengel F, Brady A, Ferreira D, Janssens U, Klepetko W, Mayer E, Remy-Jardin M, Bassand J-P (2008) Guidelines on the diagnosis and management of acute pulmonary embolism: the Task Force for the Diagnosis and Management of Acute Pulmonary Embolism of the European Society of Cardiology (ESC). *Eur Heart J* 29:2276–2315. <https://doi.org/10.1093/eurheartj/ehn310>
- Silverstein MD, Heit JA, Mohr DN, Petterson TM, O'Fallon WM, Melton LJ (1998) Trends in the incidence of deep vein thrombosis and pulmonary embolism: a 25-year population-based study. *Arch Intern Med* 158:585–593
- Morange P-E, Trégouët D-A (2013) Current knowledge on the genetics of incident venous thrombosis. *J Thromb Haemost* 11(Suppl 1):111–121. <https://doi.org/10.1111/jth.12233>
- Ferrari E, Baudouy M, Cerboni P, Tibi T, Guigner A, Leonetti J, Bory M, Morand P (1997) Clinical epidemiology of venous thromboembolic disease: results of a French Multicentre Registry. *Eur Heart J* 18:685–691. <https://doi.org/10.1093/oxfordjournals.eurheartj.a015316>
- Goldhaber SZ, Visani L, de Rosa M (1999) Acute pulmonary embolism: clinical outcomes in the International Cooperative Pulmonary Embolism Registry (ICOPER). *Lancet* 353:1386–1389
- Larsen TB, Sørensen HT, Skytthe A, Johnsen SP, Vaupel JW, Christensen K (2003) Major genetic susceptibility for venous thromboembolism in men: a study of Danish twins. *Epidemiology* 14:328–332
- Souto JC, Almasy L, Borrell M, Garí M, Martínez E, Mateo J, Stone WH, Blangero J, Fontcuberta J (2000) Genetic determinants of hemostasis phenotypes in Spanish families. *Circulation* 101:1546–1551
- Zöller B, Li X, Sundquist J, Sundquist K (2011) Age- and gender-specific familial risks for venous thromboembolism: a nationwide epidemiological study based on hospitalizations in Sweden. *Circulation* 124:1012–1020. <https://doi.org/10.1161/CIRCULATIONAHA.110.965020>
- Nicolaides AN, Breddin HK, Carpenter P, Coccheri S, Conard J, de Stefano V, Elkoofy N, Gerotziakas G, Guermazi S, Haas S, Hull R, Kalodiki E, Kristof V, Michiels JJ, Myers K, Pineo G, Prandoni P, Romeo G, Samama MM, Simonian S, Xenophonthos S (2005) Thrombophilia and venous thromboembolism. International consensus statement. Guidelines according to scientific evidence. *Int Angiol* 24:1–26
- Auro K, Alanne M, Kristiansson K, Silander K, Kuulasmaa K, Salomaa V, Peltonen L, Perola M (2007) Combined effects of thrombosis pathway gene variants predict cardiovascular events. *PLoS Genet* 3:e120. <https://doi.org/10.1371/journal.pgen.0030120>
- Rosendaal FR (1999) Venous thrombosis: a multicausal disease. *Lancet* 353:1167–1173
- Trégouët D-A, Heath S, Saut N, Biron-Andreani C, Schved J-F, Pernod G, Galan P, Drouet L, Zelenika D, Juhán-Vague I, Alessi M-C, Tiret L, Lathrop M, Emmerich J, Morange P-E (2009) Common susceptibility alleles are unlikely to contribute as strongly as the FV and ABO loci to VTE risk: results from a GWAS approach. *Blood* 113:5298–5303. <https://doi.org/10.1182/blood-2008-11-190389>
- Uitte de Willige S, de Visser MCH, Houwing-Duistermaat JJ, Rosendaal FR, Vos HL, Bertina RM (2005) Genetic variation in the fibrinogen gamma gene increases the risk for deep venous thrombosis by reducing plasma fibrinogen gamma' levels. *Blood* 106:4176–4183. <https://doi.org/10.1182/blood-2005-05-2180>
- Germain M, Saut N, Greliche N, Dina C, Lambert J-C, Perret C, Cohen W, Oudot-Mellakh T, Antoni G, Alessi M-C, Zelenika D, Cambien F, Tiret L, Bertrand M, Dupuy A-M, Letenneur L, Lathrop M, Emmerich J, Amouyel P, Trégouët D-A, Morange P-E (2011) Genetics of venous thrombosis: insights from a new genome wide association study. *PLoS One* 6:e25581. <https://doi.org/10.1371/journal.pone.0025581>
- Heit JA, Armasu SM, Asmann YW, Cunningham JM, Matsumoto ME, Petterson TM, de Andrade M (2012) A genome-wide association study of venous thromboembolism identifies risk variants in chromosomes 1q24.2 and 9q. *J Thromb Haemost* 10:1521–1531. <https://doi.org/10.1111/j.1538-7836.2012.04810.x>
- Tang W, Teichert M, Chasman DI, Heit JA, Morange P-E, Li G, Pankratz N, Leebeek FW, Paré G, de Andrade M, Tzourio C, Psaty BM, Basu S, Rutter R, Rose L, Armasu SM, Lumley T, Heckbert SR, Uitterlinden AG, Lathrop M, Rice KM, Cushman M, Hofman A, Lambert J-C, Glazer NL, Pankow JS, Witteman JC, Amouyel P, Bis JC, Bovill EG, Kong X, Tracy RP, Boerwinkle E, Rotter JI, Trégouët D-A, Loth DW, Stricker BHC, Ridker PM, Folsom AR,

- Smith NL (2013) A genome-wide association study for venous thromboembolism: the extended cohorts for heart and aging research in genomic epidemiology (CHARGE) consortium. *Genet Epidemiol* 37:512–521. <https://doi.org/10.1002/gepi.21731>
19. Desmarais S, de Moerloose P, Reber G, Minazio P, Bounameaux H, Perrier A (1996) Resistance to activated protein C in an unselected population of patients with pulmonary embolism. *Lancet* 347:1374–1375. [https://doi.org/10.1016/S0140-6736\(96\)91013-2](https://doi.org/10.1016/S0140-6736(96)91013-2)
 20. van Langevelde K, Flinterman LE, van Hylckama VA, Rosendaal FR, Cannegieter SC (2012) Broadening the factor V Leiden paradox: pulmonary embolism and deep-vein thrombosis as 2 sides of the spectrum. *Blood* 120:933–946. <https://doi.org/10.1182/blood-2012-02-407551>
 21. Brandimarti F, Alessandrini F, Pesaresi M, Catalani C, de Angelis L, Galeazzi R, Giovagnetti S, Gesuita R, Righi E, Giorgetti R, Tagliabracchi A (2017) Investigation on genetic thrombophilic factors in FFPE autopsy tissue from subjects who died from pulmonary embolism. *Int J Legal Med* 131:447–458. <https://doi.org/10.1007/s00414-016-1508-z>
 22. Germain M, Chasman DI, de Haan H, Tang W, Lindström S, Weng LC, de Andrade M, de Visser MC, Wiggins KL, Suchon P, Saut N, Smadja DM, le Gal G, van Hylckama Vlieg A, di Narzo A, Hao K, Nelson CP, Rocanin-Arjo A, Folkersen L, Monajemi R, Rose LM, Brody JA, Slagboom E, Aïssi D, Gagnon F, Deleuze JF, Deloukas P, Tzourio C, Dartigues JF, Berr C, Taylor KD, Civelek M, Eriksson P, Cardiogenics Consortium, Psaty BM, Houwing-Duitermaat J, Goodall AH, Cambien F, Kraft P, Amouyel P, Samani NJ, Basu S, Ridker PM, Rosendaal FR, Kabrhel C, Folsom AR, Heit J, Reitsma PH, Trégouët DA, Smith NL, Morange PE (2015) Meta-analysis of 65,734 individuals identifies TSPAN15 and SLC44A2 as two susceptibility loci for venous thromboembolism. *Am J Hum Genet* 96(4):532–542. <https://doi.org/10.1016/j.ajhg.2015.01.019>
 23. Sennblad B, Basu S, Mazur J, Suchon P, Martinez-Perez A, van Hylckama Vlieg A, Truong V, Li Y, Gådin JR, Tang W, Grossman V, de Haan HG, Handin N, Silveira A, Souto JC, Franco-Cereceda A, Morange P-E, Gagnon F, Soria JM, Eriksson P, Hamsten A, Maegdefessel L, Rosendaal FR, Wild P, Folsom AR, Trégouët D-A, Sabater-Lleal M (2017) Genome-wide association study with additional genetic and post-transcriptional analyses reveals novel regulators of plasma factor XI levels. *Hum Mol Genet* 26:637–649. <https://doi.org/10.1093/hmg/ddw401>
 24. Austin H, de Staercke C, Lally C, Bezemer ID, Rosendaal FR, Hooper WC (2011) New gene variants associated with venous thrombosis: a replication study in White and Black Americans. *J Thromb Haemost* 9: 489–495. <https://doi.org/10.1111/j.1538-7836.2011.04185.x>
 25. De Haan HG, Bezemer ID, Doggen CJM, Le Cessie S, Reitsma PH, Arellano AR (2012) Multiple SNP testing improves risk prediction of first venous thrombosis. *Blood* 120(3):656–663. <https://doi.org/10.1182/blood-2011-12-397752>
 26. Morange PE, Bezemer I, Saut N, Bare L, Burgos G, Brocheton J, Durand H, Biron-Andreani C, Schved JF, Pernod G, Galan P, Drouet L, Zelenika D, Germain M, Nicaud V, Heath S, Ninio E, Delluc A, Münzel T, Zeller T, Brand-Hermann SM, Alessi MC, Tired L, Lathrop M, Cambien F, Blankenberg S, Emmerich J, Trégouët DA, Rosendaal FR (2010) A follow-up study of a genome-wide association scan identifies a susceptibility locus for venous thrombosis on chromosome 6p24.1. *Am J Hum Genet* 86:592–595
 27. Buil A, Trégouët D-A, Souto JC, Saut N, Germain M, Rotival M, Tired L, Cambien F, Lathrop M, Zeller T, Alessi M-C, Rodriguez de Cordoba S, Münzel T, Wild P, Fontcuberta J, Gagnon F, Emmerich J, Almasy L, Blankenberg S, Soria J-M, Morange P-E (2010) C4BPB/C4BPA is a new susceptibility locus for venous thrombosis with unknown protein S-independent mechanism: results from genome-wide association and gene expression analyses followed by case-control studies. *Blood* 115:4644–4650. <https://doi.org/10.1182/blood-2010-01-263038>
 28. Morange P-E, Saut N, Antoni G, Emmerich J, Trégouët D-A (2011) Impact on venous thrombosis risk of newly discovered gene variants associated with FVIII and VWF plasma levels. *J Thromb Haemost* 9:229–231. <https://doi.org/10.1111/j.1538-7836.2010.04082.x>
 29. Hafke A, Schürmann P, Rothämel T, Dörk T, Klintschar M (2019) Evidence for an association of interferon gene variants with sudden infant death syndrome. *Int J Legal Med* 133(3):863–869. <https://doi.org/10.1007/s00414-018-1974-6>
 30. Fluidigm Corporation (2016) SNP genotyping. Users Guide: PN 68000098 M2
 31. Hong EP, Park JW (2012) Sample size and statistical power calculation in genetic association studies. *Genomics Inform* 10:117–122. <https://doi.org/10.5808/GL.2012.10.2.117>
 32. Butler JM (2010) Fundamentals of forensic DNA typing. Academic Press/Elsevier, Amsterdam
 33. Lek M, Karczewski KJ, Minikel EV et al (2016) Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536(7616):285–291. <https://doi.org/10.1038/nature19057>
 34. Ameziane N, Seguin C, Borgel D, Fumeron F, Moatti D, Alhenc-Gelas M, Grandchamp B, Aiach M, Emmerich J, de Prost D (2002) The -33T-C polymorphism in intron 7 of the TFPI gene influences the risk of venous thromboembolism, independently of the factor V Leiden and prothrombin mutations. *Thromb Haemost* 88:195–199
 35. Stein PD, Beemath A, Olson RE (2005) Obesity as a risk factor in venous thromboembolism. *Am J Med* 118:978–980. <https://doi.org/10.1016/j.amjmed.2005.03.012>
 36. Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, Munafò MR (2013) Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci* 14:365–376. <https://doi.org/10.1038/nrn3475>
 37. Suresh K, Chandrashekar S (2012) Sample size estimation and power analysis for clinical research studies. *J Hum Reprod Sci* 5: 7–13. <https://doi.org/10.4103/0974-1208.97779>
 38. Ziegler A, König IR, Pahlke F (2012) A statistical approach to genetic epidemiology: concepts and applications, with an e-learning platform, 4th edn. John Wiley & Sons, Weinheim
 39. van Stralen KJ, Doggen CJM, Bezemer ID, Pomp ER, Lisman T, Rosendaal FR (2008) Mechanisms of the factor V Leiden paradox. *Arterioscler Thromb Vasc Biol* 28:1872–1877. <https://doi.org/10.1161/ATVBAHA.108.169524>
 40. Cushman M, Corneli A, Folsom AR, Wang L, Tsai MY, Polak J, Tang Z (2007) Associations of the beta-fibrinogen Hae III and factor XIII Val34Leu gene variants with venous thrombosis. *Thromb Res* 121: 339–345. <https://doi.org/10.1016/j.thromres.2007.05.009>
 41. Harrington DJ, Malefora A, Schmeleva V, Kapustin S, Papayan L, Blinov M, Harrington P, Mitchell M, Savidge GF (2003) Genetic variations observed in arterial and venous thromboembolism—relevance for therapy, risk prevention and prognosis. *Clin Chem Lab Med* 41:496–500
 42. Camilleri RS, Cohen H (2005) No association between pulmonary embolism or deep vein thrombosis and the -455G/A β -fibrinogen gene polymorphism. *Blood Coagul Fibrinolysis* 16:193–198. <https://doi.org/10.1097/01.mbc.0000164428.13417.e2>
 43. Franco RF, Reitsma PH, Lourenço D, Maffei FH, Morelli V, Tavella MH, Araújo AG, Piccinato CE, Zago MA (1999) Factor XIII Val34Leu is a genetic factor involved in the etiology of venous thrombosis. *Thromb Haemost* 81:676–679
 44. Lincz LF, Adams MJ, Scorgie FE, Thom J, Baker RI, Seldon M (2007) Polymorphisms of the tissue factor pathway inhibitor gene are associated with venous thromboembolism in the antiphospholipid syndrome and carriers of factor V Leiden. *Blood Coagul Fibrinolysis* 18:559–564. <https://doi.org/10.1097/MBC.0b013e3281ee977>

45. Branchford BR, Carpenter SL (2018) The role of inflammation in venous thromboembolism. *Front Pediatr* 6:142
46. Liu F, Silva D, Malone MV, Seetharaman K (2017) MTHFR A1298C and C677T polymorphisms are associated with increased risk of venous thromboembolism: a retrospective chart review study. *Acta Haematol* 138:208–215. <https://doi.org/10.1159/000480447>
47. Li Z, Yadav U, Mahemuti A, Tang B-P, Upur H (2015) Association of MTHFR genetic polymorphisms with venous thromboembolism in Uyghur population in Xinjiang, China. *Int J Clin Exp Med* 8: 17703–17711
48. Bezemer ID, Bare LA, Arellano AR, Reitsma PH, Rosendaal FR (2010) Updated analysis of gene variants associated with deep vein thrombosis. *JAMA* 303:421–422. <https://doi.org/10.1001/jama.2010.57>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.