Journal of the American Heart Association

ORIGINAL RESEARCH

Familial Segregation of Venous Thromboembolism in Sweden: A Nationwide Family Study of Heritability and Complex Segregation Analysis

Bengt Zöller, MD, PhD; MirNabi Pirouzifard, PhD; Peter J. Svensson, MD, PhD; Björn Holmquist, PhD; Emelie Stenman, PhD; Robert C. Elston, PhD; Yeunjoo E. Song, PhD; Jan Sundquist, MD, PhD; Kristina Sundquist, MD, PhD

BACKGROUND: This is the first nationwide segregation analysis that aimed to determine whether familial venous thromboembolism (VTE) is attributable to inheritance and/or shared environment, and the possible mode of inheritance.

METHODS AND RESULTS: The Swedish Multi-Generation Register was linked to the Swedish patient register for the period 1964 to 2015. Three generational families of Swedish-born individuals were identified. Heritability was examined using Falconer regression. Complex segregation analysis was conducted using the Statistical Analysis for Genetic Epidemiology software (version 6.4, 64-bit Linux). Among the 4 301 174 relatives from 450 558 pedigrees, 177 865 (52% women) individuals were affected with VTE. VTE occurred in 2 or more affected relatives in 61 217 (13.6%) of the pedigrees. Heritability showed age and sex dependence with higher heritability for men and young individuals. In 18 933 pedigrees, VTE occurred only in the first generation and was not inherited. Segregation analysis was performed in the remaining 42 284 pedigrees with inherited VTE and included 939 192 individuals. Prevalence constraints were imposed in the models to allow for the selection of the pedigrees analyzed. The sporadic nongenetic model could be discarded. The major-type-only model, with a correlation structure compatible with some polygenic effects, was the preferred model. Among the Mendelian models, the mixed codominant (plus polygenic) model was preferred.

CONCLUSIONS: This nationwide segregation analysis of VTE supports a genetic cause of the familial aggregation of VTE. Heritability was higher for men and younger individuals, suggesting a Carter effect, in agreement with a multifactorial threshold inheritance.

Key Words: epidemiology ■ genetics ■ heritability ■ segregation analysis ■ venous thromboembolism

Venous thromboembolism (VTE) affects ≈1 to 2 per 1000 individuals per year.[1] The existence of familial thrombophilia (eg, clustering of VTE) was recognized at the beginning of the 20th century.[2] VTE is considered to be a complex disorder influenced by several genetic and environmental factors.[3–5] Inherited deficiencies of the natural anticoagulants antithrombin, protein C, and proteins S, as well as activated protein C resistance because of the FV Leiden (rs6025) and the prothrombin 20210A (rs1799963) polymorphisms, have been associated with familial thrombophilia.[3–6] The 5 major identified genetic risk factors for VTE together account for only about 30% of the family history of VTE.[7] Recently, a whole-exome study by Desch et al reported rare variants in the *STAB2* gene to be associated with VTE.[8] Genome-wide association studies

Correspondence to: Bengt Zöller, MD, PhD, Center for Primary Health Care Research, CRC, Building 28, Floor 11, Jan Waldenströms gata 35, Skåne University Hospital, S-205 02 Malmö, Sweden. E-mail: bengt.zoller@med.lu.se

Supplementary Material for this article is available at https://www.ahajournals.org/doi/suppl/10.1161/JAHA.120.020323

For Sources of Funding and Disclosures, see page 11.

JAHA is available at: www.ahajournals.org/journal/jaha

CLINICAL PERSPECTIVE

What Is New?

- This is the first nationwide complex segregation analysis that aimed to determine whether familial venous thromboembolism is attributable to inheritance and/or shared environment and the possible mode of inheritance.
- Genes are important for the familial clustering of venous thromboembolism in Sweden. A nongenetic model was discarded.
- A single major-type-only model with the possibility of some polygenic effects was the preferred model.

What Are the Clinical Implications?

- The study supports the importance of genes in the familial clustering of venous thromboembolism in Sweden.
- Genetic studies of families with familial aggregation of venous thromboembolism might be worthwhile.

Nonstandard Abbreviations and Acronyms

AIC	Akaike information criterion

have identified common, but weaker, alleles with an overall estimated heritability of 26%.[5,9–11] Thus, as for many other complex disorders, the source of the missing heritability of VTE remains to be determined.[12] The missing heritability of VTE might be genetic, epigenetic, or nongenetic (attributable to familial environmental effects).[12] Shared familial environmental factors might be related to socioeconomic factors such as income and education but also health behaviors such as alcohol consumption, drug abuse, smoking, lack of exercise, and eating habits (eg, high intake of saturated fat and low intake of fish, fruit, and vegetables).[13]

Heritability of VTE, that is, the proportion of the variance attributable to genetic effects on VTE, has been estimated to be 50% to 60% in a twin study and 2 extended family studies.[14–16] However, the twin study was small, and the heritability was lower and nonsignificant for women.[14] A Swedish sibling and half-sibling study estimated heritability of liability to be 47% for men and 40% for women (ie, lower heritability for women compared with men), which concurs with the Danish twin study.[17] The 2 extended family studies were mainly performed in selected families with idiopathic thrombosis or thrombophilia referred to specialist clinics.[15,16] Thus, limited data exist about the heritability of VTE in the general population.

Only the study by Heit et al has performed segregation analysis of families with VTE.[16] Heit et al performed segregation analysis in 16 650 relatives of 751 probands (60% women) with objectively diagnosed VTE without cancer. Patients who were referred to the Mayo Clinic Special Coagulation Laboratory/Clinic for a clinical suspicion of thrombophilia were approached for study inclusion. The probands completed questionnaires that provided a family pedigree, ethnic ancestry, and indicated all other family members who were affected with VTE. All patients were non-Hispanic White Americans. Heit et al found that a multifactorial non-Mendelian inheritance model should be favored as the cause for VTE, whereas a model postulating a purely environmental cause was rejected.[16]

The present study aims to determine the contribution of genetic factors (ie, heritability) and whether familial VTE is attributable to inheritance and/or shared environment, and the possible mode of inheritance. The study is nationwide and uses the Swedish multigeneration register for establishing familial relationships and the Swedish inpatient and outpatient registers for VTE diagnosis.[18–20]

METHODS

Data Sharing Statement

All of the nationwide data used in the present study were obtained from Statistics Sweden and the National Board of Health and Welfare (https://bestalladata.socialstyrelsen.se/data-for-forskning/sekretessprovning/). No additional data are available because of Swedish regulations.

Registers Used

We used linked data from multiple Swedish nationwide registries and health care data.[18–20] Linking was achieved via the unique individual 10-digit personal identification number that is assigned at birth or immigration to all Swedish residents. To preserve confidentiality, this identification number was replaced by a serial number by Statistics Sweden (https://www.scb.se/en/). The following sources were used to create our database: Total Population Register; Multi-Generation Register, providing information on family relations; Swedish Hospital Discharge Register, containing hospitalizations for all Swedish inhabitants from 1964 to 2015; Outpatient Care Register, containing information from outpatient clinics from 2001 to 2015; and Swedish Mortality Register, containing all causes of death.[18–23]

Study Population and Pedigree Construction

The study population consisted of individuals from the Swedish population who were born in Sweden between

1876 and 2015. The multigeneration register was used to find 3 generation pedigrees with information about all paternal and maternal grandparents.[20] Only complete 3-generational pedigrees (grandparents, offspring, and grandchildren) where all 4 grandparents were born between 1876 and 1950 were selected. Pedigrees with any non–Swedish-born grandparents were excluded. Moreover, if a grandparent (first generation), offspring (second generation), or grandchild (third generation) died or emigrated before 1964, the whole pedigree was excluded. Pedigrees with twins, adoptees, >3 mates (eg, spouses or partners), and loops (eg, cousin marriage) were also excluded because of requirements of the software used. Pedigrees with outlier individuals over 110 years of age in the end of follow-up were excluded from the study. All pedigrees with size <7 (ie, incomplete 3-generational pedigrees) were excluded. Minimum size included 4 grandparents, 1 offspring with spouse, and 1 grandchild. To get better coverage in the Swedish patient register, pedigrees were excluded if offspring (second generation) were born after 1970 (spouses were allowed to be born after 1970).

Any duplicate pedigree was excluded, but individuals were allowed to occur in >1 pedigree. All individuals within all pedigrees were linked to the Swedish Patient Register (inpatients and outpatients) for information on VTE diagnosis (https://www.socialstyrelsen.se/statistik-och-data/register/alla-register/patientregistret/).[19]

Ascertainment of VTE Cases

Cases of VTE (ie, affection status) were classified according to the World Health Organization *International Classification of Diseases, Seventh Revision* through *Tenth Revision* (*ICD-7*, *ICD-8*, *ICD-9*, *ICD-10*). All types of VTE diagnoses were counted.[24,25] VTE was defined using the hospital discharge register, main and supplementary diagnoses (between 1964 and 2015) by the following *ICD* codes: *ICD-10*: I636, I676, I80, I81, I82, I26, O222, O223, O225, O229, O870, O871, O873, O879, O882, O082, O087; *ICD-9*: 437G, 451, 452, 453, 415B, 416W, 671C, 671D, 671E, 671F, 671X, 639G, 673C; *ICD-8*: 321, 450, 451, 452, 453, 671, 673.9; *ICD-7*: 463, 464, 465, 466, 58300, 33440, 33450, 682, 684. It was also defined in the Outpatient Care Register (between 2001 and 2015) by the same *ICD-10* codes as above. Thus, VTE included all different manifestations of venous thrombosis and pulmonary embolism.[21,24,25]

The Swedish Hospital Discharge Register has nearly 90% overall validity or positive predictive values.[19] In a Swedish study of men with VTE, hospital records were available for 304 cases (1970–1998).[26] A total of 289 out of 304 (95%) cases with diagnosed VTE were judged to be diagnosed correctly.[26] Only 12 (3.9%) cases were not diagnosed with an objectively verified method but were treated with oral anticoagulation because of strong clinical probability. In total, 277 (91%) cases were objectively diagnosed with methods such as phlebography, ultrasound, computed tomography scan, or pulmonary scintigraphy.[26]

In the present study, 177 865 individuals were affected by VTE. The first VTE event occurred between 1964 and 1969 for 548 (0.31%) individuals, between 1970 and 1979 for 10 258 (5.77%) individuals, between 1980 and 1989 for 27 376 (15.39%) individuals, between 1990 and 1999 for 37 919 (21.32%) individuals, between 2000 and 2010 for 56 391 (31.70%) individuals, and between 2010 and 2015 for 45 373 (25.51%) individuals.

Statistical Analysis

A nested case-control approach was used to determine familial odds ratios (ORs). We used a matching method (1:5) for sex, educational attainment, county of birth (ie, geographic location), and birth year by drawing a sample of unique VTE-affected grandchildren or children as cases and matched control groups of VTE-unaffected grandchildren or children. An individual may be included as a case both in the second generation or third generation nested case-control studies. However, no duplicate individuals were included in the same generation case-control study. ORs were determined with conditional logistic regression for affected grandchildren (third generation) with at least 1 affected biological parent (second generation). In the same way, ORs were determined with conditional logistic regression for affected offspring (second generation) with at least 1 affected biological parent (first generation).

Formally, heritability is defined as a ratio of variances, that is, the proportion of total variance that is associated with variation in additive genetic factors. According to classic quantitative genetics, the heritability of a binary trait (or disease) could be estimated by Falconer regression by presuming a liability threshold model of the disease in which everyone has a liability to develop the disease but only individuals above a threshold value do so.[27–29] To evaluate heritability for VTE, Falconer regression was used. It is based on the liability of the threshold to obtain heritability in children of the parents. The method and its application are described in detail by Falconer and MacKay.[28,29] We used χ^2 to compare proportions and Wald tests in logistic regression. Throughout all analyses, a 2-sided test resulting in a $P<0.05$ was considered significant. Statistical analysis was performed with SAS software version 9.3 (SAS Institute) and R version 3.6.0 (R Foundation for Statistical Computing).

To explore the pattern of familial inheritance of VTE, complex segregation analysis was performed using the maximum likelihood method to assess the parameters in each of the hypothesis-based mathematical models examined.[30–35] The program SEGREG, within

the 64-bit Linux software package Statistical Analysis for Genetic Epidemiology 6.4, was used for the complex segregation analysis.[30–35] A short summary is depicted here, but a more comprehensive description of the procedures exists elsewhere.[33–35] Moreover, a detailed protocol of the procedure used in the present study has been described by Sun.[33]

Complex segregation analysis is an essential tool in genetics and applies statistical methods (without DNA analysis) to define if the variation of a disease/phenotype (VTE), continuous or binary, is compatible with segregation of a single locus, and to establish the mode of inheritance. For each model, it is assumed that the presence (or absence) of the presumed disease allele affects susceptibility to VTE. The regressive multivariate logistic model for binary trait, as described by Karunaratne and Elston, was applied.[32] This approach enables the inclusion of covariates in the fitted models. The fitted models assume that, conditional on the phenotype and the major type of any individual who belongs to 2 nuclear families, the likelihood for those 2 nuclear families are independent. Therefore, the susceptibility that any pedigree member has to a particular phenotype (VTE) is the same for all members who have the same values of any covariates in the model.

The segregation model for a binary trait assumes that susceptibility to a disease/phenotype (VTE), γ, defined as the probability that an individual has been affected with disease, depends on an unobserved latent factor called type, labeled as *u*, which can take 1 of the 3 values AA, AB, or BB.[30–35] If the segregation is Mendelian (γ is then interpreted as the penetrance), the type *u* represents the presumed genotype that underlies the distribution of the observed disease/phenotype. Types are characterized by the type frequencies and transmission parameters. Assuming Hardy-Weinberg equilibrium, the type frequencies in the population could be defined by the parameter q_A, the allele frequency of allele A. Transmission parameters, recorded as τ_{AA}, τ_{AB}, and τ_{BB}, refer to the probabilities that parents of types AA, AB, and BB transmit the allele A to offspring. For binary traits only, the logit of the susceptibility of the major type AA, AB, and BB are β_{AA}, β_{AB}, and β_{BB}.

Assuming that mating is random, the transmissions from each parent are independent.[30–35] For a Mendelian locus, Hardy-Weinberg equilibrium is assumed, and τ_{AA}, τ_{AB}, and τ_{BB} are equal to 1, 0.5, and 0, respectively. If the disease segregation is caused solely by a random environmental factor, there is no transmission between generations, and the 3 unobserved types are equally transmitted ($\tau_{AA}=\tau_{AB}=\tau_{BB}$). In addition, a general transmission model allows the transmission probabilities τ_{AA}, τ_{AB}, and τ_{BB} to take on any arbitrary values between 0 and 1. A more restricted general transmission model assumes homogeneity of the phenotypic distribution across generations and must satisfy 2 conditions: the type frequencies must follow Hardy-Weinberg equilibrium proportions and τ_{AB} must be equal to a specific function of the frequency of A, τ_{AA}, and τ_{BB}; thus, only 2 of the transmission probabilities can be freely estimated.

For a binary trait such as VTE, the susceptibility (γ) is determined by the cumulative logistic function

$$\gamma = \frac{e^{\theta_u(i)}}{1 + e^{\theta_u(i)}},$$

where conditional on type *u*, the logit of the susceptibility for the i-th individual, θ_u, can depend on both major type *u* and mean centered covariate values x_{i1}, x_{i2}, …, x_{ip}:

$$\theta_u = \beta_u + \xi_1 x_{i1} + \xi_2 x_{i2} + \cdots + \xi_p x_{ip},$$

where β_u is the intercept corresponding to type *u* and ξ_1, ξ_2, …, ξ_p are the covariate regression coefficients. This is comparable to estimating covariate coefficients in a linear model at the same time with other parameters in the segregation analysis of a quantitative trait.

The segregation models can be described according to the number of baseline parameters for susceptibility in the model.[30–35] One susceptibility-type model (no segregation, $\beta_{AA}=\beta_{AB}=\beta_{BB}$) includes the sporadic model, with or without familial correlation (association), in the absence of transmission of a major gene. Two susceptibility-type models include Mendelian dominant and recessive models ($\beta_{AA}=\beta_{AB}$, or $\beta_{AB}=\beta_{BB}$). Three susceptibility-type models (β_{AA}, β_{AB}, β_{BB}) involve several transmission models. The general model, in particular, allows the transmission possibilities τ_{AA}, τ_{AB}, and τ_{BB} to take on any arbitrary values between 0 and 1.

The most parsimonious model is identified by using Akaike information criterion (AIC), which is defined as (–2 ln[L]+2[number of parameters estimated]). The model with a smaller AIC fits the data better. For a fixed number of susceptibility types, the SEGREG output allows for *P* values based on the appropriate asymptotic distribution of the likelihood ratio criterion, either a χ^2 distribution or a mixture of χ^2 distributions, as well as AIC values, which can be compared across results for different numbers of susceptibility types.

To account for the selection of pedigrees with inherited VTE, prevalence constraints determined from the whole national population were imposed on all models. Sex- and age-specific prevalence constraints determined in the whole sample (ie, in all of the 450 558 pedigrees comprising 4 301 174 individuals) were used. The age-prevalence constraints were imposed as generation-specific constraints (ie, specific for grandparent, offspring, and grandchildren).

The segregation models that can be tested in SEGREG are as follows[30–35]:

1. Nongenetic models are the sporadic model and the random environmental transmission model.[33] The nongenetic sporadic model assumes no intergenerational transmission of the type (ie, the phenotype has 1 distribution), and no major type or multifactorial component is transmitted. The random environmental transmission model assumes that the trait segregation is caused solely by a random environmental factor, and there is no transmission from generation to generation ($\tau_{AA}=\tau_{AB}=\tau_{BB}=q_A$). As indicated by Sun,[33] this model may be tested only for continuous variables and not binary traits and was therefore not tested in the present study.
2. The polygenic transmission model assumes that the phenotype is defined by additive polygenic inheritance, so the phenotype has 1 distribution, and familial correlations (=associations) can explain the familial aggregation of the trait.[33] If, within nuclear families, the estimated sibling correlation is the same as the parent–offspring correlation, and there is no spouse correlation, the polygenic variance can account for such residual familial correlation.
3. The polygenic–environmental model assumes that only a nontransmittable random environmental factor and a polygenic/multifactorial effect influence the trait.
4. The pure major locus transmission models assume major locus transmission in a Mendelian mode without multifactorial/polygenic inheritance (pure Mendelian models or with a multifactorial/polygenic component [mixed models]).[33] The Mendelian models are codominant, dominant, recessive, or additive, with or without polygenic component.
5. The general transmission models are the major type models transmitted with arbitrary probabilities between 0 and 1, with (general model) or without (major type only) polygenic/multifactorial effects.[33] The general model is the unrestricted full model, which subsumes all the other models.

Ethical Considerations

We secured ethical approval for this study from the Regional Ethical Review Board in Lund. Informed consent was waived as a requirement by the ethics committee.

RESULTS

Descriptive Findings of Swedish Families With VTE

After inclusions and all exclusions, a total of 450 558 pedigrees comprising 4 301 174 individuals were identified: grandparents (first generation), offspring (second generation), and grandchildren (third generation). Table 1 shows the descriptive statistics for these individuals. Some individuals could occur in >1 pedigree and/or generation. In Table 1, individuals are shown only once (ie, in the oldest generation they occur in). For comparison, in Table S1, individuals are counted each time they occur in a pedigree (n=7 568 818). In Table 1, during the study period (1964–2015), a total of 177 865 individuals were affected with VTE (4.1%), with a median age of 68 years (range, 0–104 years) at first VTE (Table 1). A little more than half of VTE cases were women (n=92 484, 52.0%). Thus, overall as well as in the grandparent and grandchildren generations, there were more women than men affected by VTE (Table 1 and Table S1). In the second generation (offspring), more men than women were affected by VTE. However, the proportion of women was slightly higher in the second generation in Table S1 compared with Table 1.

The median age at the end of follow-up (2015, death, or emigration) was 81 years (range, 20–110 years) for grandparents, 63 years (range, 1–83 years) for offspring, and 39 years (range, 0–65 years) for grandchildren (Table 1). VTE was found in 116 998

Table 1. Descriptive Statistics of the Study Sample (Population of Swedish-Born Individuals, N=4 301 174) With and Without Venous Thromboembolism (VTE) During the Study Period 1964 to 2015

	VTE cases	No VTE
Group, n (%)		
All	177 865 (100)	4 123 309 (100)
Grandparents	116 998 (65.8)	1 180 800 (28.6)
Offspring	47 968 (27.0)	1 271 962 (30.9)
Grandchildren	12 899 (7.3)	167 0547 (40.5)
Age at end of follow-up, y, median [q1–q3] (minimum–maximum)*		
All	75 [64–83] (0–110)	52 [32–71] (0–110)
Grandparents	81 [74–87] (20–110)	79 [72–85] (18–110)
Offspring	63 [55–69] (1–83)	58 [51–66] (0–83)
Grandchildren	39 [32–45] (0–65)	28 [21–38] (0–66)
Age at first VTE, y, median [q1–q3] (minimum–maximum)		
All	68 [55–77] (0–104)	…
Grandparents	74 [66–81] (18–104)	…
Offspring	55 [45–63] (0–83)	…
Grandchildren	31 [24–38] (0–64)	…
Women, n (%)		
All	92 484 (52.0)	2 006 238 (48.7)
Grandparents	62 298 (53.3)	587 499 (49.8)
Offspring	22 355 (46.6)	614 386 (48.3)
Grandchildren	7831 (60.7)	804 353 (48.2)

Each individual occurs only once. Some individuals may occur in >1 pedigree and in >1 generation. In this table they are included in the oldest generation they occur in. q1–q3 indicates interquartile range (q1, first quartile and q3, third quartile).

*End of follow-up is either age at end of study in 2015, age at death, or age at emigration depending on which occurred first.

(9.0%) of the 1 297 798 grandparents, 47 968 (3.6%) of the 1 319 930 offspring, and 12 899 (0.8%) of the 1 683 446 grandchildren. Median age at first VTE diagnosis (minimum–maximum) for the grandparents was 74 years (range, 8–104 years), for offspring 55 years (range, 0–83 years), and for the grandchildren 31 years (range, 0–64 years).

Table 2 shows pedigree data according to number of affected members in the pedigrees. No pedigree had >10 affected members. In total, 61 217 (13.6%) of all 450 558 pedigrees exhibited VTE in at least 2 affected individuals. Individuals in Table 2 could be included more than once if they were included in several pedigrees. Therefore, the number of affected individuals does not add up with Table 1. In 18 933 pedigrees, only grandparents (ie, first generation) were affected, and VTE was not inherited in these pedigrees. Thus, in the remaining 42 284 pedigrees VTE was inherited.

Nested Case-Control Study and Heritability of VTE

The results of the nested 1:5 matched case-control study are presented in Table 3. In Table 3, ORs and heritability with SE determined with the Falconer method are presented for offspring (second generation) and grandchildren (third generation). ORs and heritability were higher for the third generation compared with the second generation. ORs and heritability were also slightly higher for men compared with women, both in the second and third generations. The estimated heritability±SE for VTE in offspring (second generation) was 33.8±0.7% compared with 39.0±1.3% for grandchildren (third generation). In the second generation, heritability for men (36.8±1.0%) was higher than for women (30.4±1.0%). Also, in the third generation, heritability was higher for men (43.9±2.0%) than for women (35.7±1.7%).

Segregation Analysis of VTE in Pedigrees With Inherited VTE

For segregation analysis, pedigrees with affected members only in the first generation were excluded (18 933 pedigrees), because in these pedigrees VTE was not inherited. The remaining 42 284 pedigrees with inherited VTE (ie, at least 2 affected pedigree members not limited to the first generation and comprising 939 192 individuals) were selected for segregation analysis (Table 4). Descriptive data according to relative pairs, affection status (VTE yes/no), concordant affected relative pairs (both relatives affected in a pair), discordant relative pairs (1 affected and 1 unaffected relative), and sex are presented in Table 4.

First, the effect of sex and age as covariates on the probability of being affected with VTE was evaluated in a 1-susceptibility-type model that assumes no segregation (sporadic model). The results are presented in Table 5. Using the covariates age and sex gave the lowest Akaike value and was also significantly better than the other models (Table 5). Both age and sex were therefor included as covariates.

Because we selected for analysis families with inherited VTE, we imposed a prevalence constraint determined in the whole study population. Age- (ie, generation) and sex-specific constraints were imposed on all models (Table 6) with results shown in Tables 7 and 8. To determine whether support existed for familial correlation (association), a sporadic model without familial correlation was compared with sporadic models with 3 types of familial correlations, father-offspring (γ_{fo}), mother-offspring (γ_{mo}), and sib-sib (γ_{ss}) correlations. Table 7 presents the effect of incorporating different familial associations in these nonsegregating models. Among all models, model A had a higher AIC value than models B to D, providing support for the

Table 2. Pedigree Structure According to Number of Venous Thromboembolism–Affected Family Members in the Pedigree

No. affected in a pedigree	No. of pedigrees (%)	Pedigree size		
		Minimum	Maximum	Mean (SD)
0	247 263 (54.9)	7	76	16 (5)
1	142 078 (31.5)	7	73	17 (6)
2	45 907 (10.2)	7	73	18 (7)
3	11 748 (2.6)	7	65	19 (7)
4	2745 (0.6)	7	60	21 (8)
5	609 (0.1)	8	67	23 (8)
6	143 (0.0)	10	51	26 (9)
7	41 (0.0)	13	52	28 (9)
8	11 (0.0)	23	53	38 (11)
9	9 (0.0)	21	54	43 (12)
10	4 (0.0)	47	54	50 (3)
All pedigrees	450 558 (100)	7	76	17 (6)

An individual may occur in >1 pedigree. The number of venous thromboembolism cases is therefore not the same as in Table 1.

Table 3. Nested Case-Control Study

	Cases	Controls		
	Affected/nonaffected parents	Affected/nonaffected parents	OR (95% CI)	Falconer regression, heritability±SE, %
Second generation, offspring				
All	14 762/41 972	48 898/234 772	1.70 (1.66–1.73)	33.8±0.7
Men	7905/21 882	25 701/123 234	1.74 (1.69–1.79)	36.8±1.0
Women	6857/20 090	23 197/111 538	1.65 (1.60–1.70)	30.4±1.0
Third generation, grandchildren				
All	2855/13 919	7541/76 329	2.11 (2.01–2.21)	39.0±1.3
Men	1263/5436	3134/30 361	2.30 (2.14–2.47)	43.9±2.0
Women	1592/8483	4407/45 968	1.99 (1.86–2.11)	35.7±1.7

Heritability was estimated using Falconer regression for offspring (second generation) and grandchildren (third generation) based on a nested 1:5 matched case-control study. Cases and controls were matched 1:5 for sex, educational attainment, county of birth, and birth year. An individual included in >1 pedigree may occur in both the second generation and the third generation. Therefore, the number of cases is more than presented in Table 1, which only includes each individual once. OR indicates odds ratio.

existence of familial correlations (polygenic/multifactorial component) in the data. Only minor differences in AIC values were observed for models B to D. Model C, with equal transmission for parent-offspring and different sib-sib correlations (ie, generation-specific transmission), was selected.

Table 8 presents the estimated parameters for different models of the mode inheritance of VTE susceptibility. Hypotheses are assessed by the likelihood ratio test, when the null hypothesis is not on a boundary of the less-restricted model, and the negative of twice the difference in natural logarithms for hierarchical models follows a χ^2 distribution, with the number of degrees of freedom equal to the difference in the number of independent parameters estimated. All models, except the major-type-only model (model 12), including the sporadic nongenetic model (model 1) were rejected against the general unrestricted model 13. However, the AIC supported the major-type-only model (model 12)[33] with the lowest Akaike value compared with all other models. The worst model with highest Akaike value was the sporadic nongenetic model (model 1). Thus, all genetic models had higher Akaike values compared with the nongenetic model.

The unrestricted general model (model 13), which subsumes all the other models, was the second preferred model. A polygenic-environmental model (model 3) was the third preferred model, followed by

Table 4. Pedigrees (n=42 284) With Inherited or Familial VTE Used for Segregation Analysis

	Pedigree size, mean (SD) (range), 22.21 (8.70) (7–84)		
Relative pairs	Count, n	Concordant affected pairs, n	Discordant pairs, n
Parents/offspring	1 244 150	27 876	264 707
Sibling/sibling	563 770	6631	99 872
Sister/sister	132 637	1425	24 057
Brother/brother	148 430	1974	25 198
Brother/sister	282 703	3232	50 617
Grandparents	963 466	9553	269 052
Avuncular	1 172 790	4517	222 343
Half-sibling	65 195	78	5525
Cousin	1 255 925	1073	87 212
	VTE and sex		
	All, n	Affected, n	Affected, %
All individuals	939 192	102 752	10.9
Men	476 784	49 205	10.3
Women	462 408	53 547	11.6

Pedigrees with affected members only in the first generation were not included because VTE was not inherited in these pedigrees. Concordant affected pairs indicates both relatives in a pair are affected. Discordant pairs indicates 1 relative is affected and 1 relative is unaffected in the pair. VTE indicates venous thromboembolism.

Table 5. Testing Whether Covariates Sex and Age Should Be Included in a Segregation Model Using the Multivariate Logistic Model

Parameters	Covariate(s) in the model			
	No covariate	Sex	Age	Sex and age
β	−2.84	−2.84	−2.98	−2.98
ξsex		0.30		0.11
ξage			0.06	0.06
−2ln(L)	2 172 220	2 171 320	2 064 310	2 064 140
*Df**	1	2	2	3
AIC	2 172 222	2 171 324	2 064 314	**2 064 146**
χ^2 ($\triangle df$)†	108 080 (2)	107 180 (1)	170 (1)	…
P value	<0.001	<0.001	<0.001	…

Sex status is a covariate of the (logit) type susceptibility of a binary trait. A sporadic and nonsegregating model is used for testing whether sex, age, or sex and age should be included in further modeling. Bold indicates that this model had lowest AIC. AIC indicates Akaike information criterion.

*Number of functionally independent parameters estimated.

†χ^2 test where $\triangle df$=difference in degree of freedom between compared models.

the Mendelian mixed codominant (plus polygenic) model (model 8). Thus, among the Mendelian models, the mixed codominant (plus polygenic) model 8 was preferred with the lowest Akaike value (Table 8). This mixed Mendelian model 8 was preferred also over an additive polygenic model 2 (ie, lower Akaike value).

DISCUSSION

The present study is the first segregation study showing genetic susceptibility for VTE based on data from a nationwide general population. The preferred model was the major type only,[33] but with a correlation structure compatible with some polygenic inheritance. To be comparable with other studies we used the Falconer model, which estimates the liability to heritability, rather than the heritability of the disease itself. In agreement with previous studies, we found a higher heritability of VTE for men compared with women.[14,17] This concurs with the Carter effect, because the VTE prevalence was slightly higher for women, at least among grandchildren.[36] Heritability was also shown to be age dependent, with the highest heritability among the youngest generation (grandchildren). Only 1 previous segregation analysis of families with VTE has been published before.[16] In that study, Heit et al could discard a nongenetic model,[16] whereas a general model fitted best. In the present nationwide study in agreement with Heit et al, segregation analysis could discard a sporadic nongenetic model. However, in the present study, the major-type-only model, with correlation structure compatible with some polygenic inheritance instead of the unrestricted general model,[16] fitted the data best. In Sweden, the factor V Leiden mutation and the prothrombin 20210A mutation associated with VTE are present in around 10% and 2% of the population, respectively.[37,38] Protein S deficiency is also indicated to be rather common, with a prevalence of 1% to 2%.[39] However, the prevalence of protein C deficiency and antithrombin deficiency is unknown in Sweden. Nevertheless, the prevalence of major thrombophilic genes in the Swedish population is relatively high and fits with the major-type-only model (Table 7), for example, for Swedish families with inherited protein S deficiency.[40,41] The importance of genes for the familial clustering of VTE suggests that it might be worthwhile to search for novel genes in pedigrees with inherited VTE. Recently, a Swedish study has found a protective effect of nonsynonymous variants in procoagulant genes.[42] Rare *STAB2* variants have also been implicated in VTE.[8,43] Therefore, it is of high interest to

Table 6. Prevalence Constraints of VTE Imposed in All Models

Generation	Women	Men
Grandparents		
VTE affected	62 298 (9.6%)	54 700 (8.4%)
All	649 797 (100%)	648 001 (100%)
Age, y*	79	77
Offspring		
VTE affected	22 355 (3.5%)	25 613 (3.7%)
All	636 741 (100%)	683 189 (100%)
Age, y*	58	59
Grandchildren		
VTE affected	7831 (0.96%)	5068 (0.58%)
All	812 184 (100%)	871 262 (100%)
Age, y*	29	29

All 450 558 pedigrees comprising 4 301 174 individuals were used to determine generation (age) and sex-specific prevalence. VTE indicates venous thromboembolism.

*Mean age at end follow-up.

Table 7. Parameter Estimated in a Sporadic Model and Nonsegregating Model (Model A), Incorporating Various Familial Associations (Model B to D) by Multivariate Logistic Model (γ_{fm}=0)

Parameters	Familial association components			
	Model A, $\gamma_{fo}=\gamma_{mo}=\gamma_{ss}=0$	Model B, $\gamma_{fo}=\gamma_{mo}=\gamma_{ss}$	Model C, $\gamma_{fo}=\gamma_{mo}$, γ_{ss}	Model D, γ_{fo}, γ_{mo}, γ_{ss}
β	−2.93	−2.92	−2.92	−2.92
Sex	0.08	0.08	0.08	0.08
Age	0.03	0.03	0.03	0.03
Generation	−0.48	−0.40	−0.41	−0.41
γ_{fo}	[0]*	0.63	0.64	0.62
γ_{mo}	[0]*	0.63	0.64	0.65
γ_{ss}	[0]*	0.63	0.65	0.65
−2ln(L)	1 967 670	1 963 110	1 963 110	1 963 106
df	4	5	6	7
AIC	1 967 678	1 963 120	1 963 122	1 963 120
χ^2 ($\triangle df$)	4 564 (3)	4 (2)	4 (1)	…
P value	<0.001	0.135	0.046	…

Familial association is the same as familial correlation. To determine support for familial association (polygenic/multifactorial component), we compared the sporadic model without familial association together with sporadic models with 3 types of familial associations. The table presents the effect of incorporating different familial associations in these nonsegregating models. γ_{fm} correlations set to 0 assuming absence of assortative mating or consanguineous mating. Familial correlations; father-offspring (γ_{fo}), mother-offspring (γ_{mo}), and sib-sib (γ_{ss}) correlations. AIC indicates Akaike information criterion; $\triangle df$, difference in degrees of freedom between compared models.

*Parameters were fixed at the values indicated.

determine whether unknown strong major thrombophilic variants exist in the Swedish population.

Among the Mendelian models, the lowest Akaike value was observed for the Mendelian codominant plus polygenic model. In codominant inheritance, 2 different alleles of a gene are expressed, and each influence the genetic trait. Speculatively, this finding could fit with the recent finding of a protective effect of nonsynonymous variants in procoagulant genes, including the *F5* gene.[42] In the study by Heit et al, the dominant mixed model was preferred among the Mendelian models.[16] It is possible that this difference is because of the high prevalence (10%) of Factor V Leiden in Sweden.[37,39]

Limitations

The use of a nationwide sample minimizes the risk of selection bias when proper allowance is made for how the sample of nonsegregating pedigrees is selected for analysis. However, as always in variance partition studies, the interpretation of the results is constrained by time (1964–2015) and geographical location (Sweden). Inclusion and exclusion criteria were set to secure that all VTE events in a pedigree could be registered in the Swedish patient registers (1964–2015). This is a limitation because families with fatal pulmonary embolism before 1964 are not included. It is possible that some pedigrees with major genetic factors causing early death were missed. Thus, it is possible that the influence of major genes is underestimated. Moreover, families with better health than the average might be overrepresented, which is reflected by the high median age at end of follow-up in the oldest generation. However, this secures a correct ascertainment of VTE cases in included pedigrees, though some cases may not be registered because the registers are only nationwide from 1987. However, asking a proband about the family history of a disease is an important source of self-report and recall bias, probably exceeding the study limitations of the present study design.[44]

Another limitation of this study is that outpatient data were only available from 2001. However, outpatient treatment before this date was not as common as today. Moreover, because VTE is a recurrent disease, these outpatients are likely to be affected later and found in the register in 2001 and after.[24] Severe cases with VTE are more likely to be hospitalized than less severe cases and it might be an advantage when studying inherited disorders to focus on the most severe phenotypes.[45] In addition, the large number of comparisons is a point worthy of consideration. Although some researchers support correction for multiple comparisons, others argue that such corrections may lead to an underestimation of effects.[46]

Another limitation is the risk for misclassification of VTE cases diagnosed between 1964 and 1969, because VTE was more often based on clinical judgment rather than objective methods during this period. However, only 0.31% (n=548) of all cases in the present study were diagnosed during this period. Still, it is

Table 8. Parameter Estimates From Segregation Analysis of Venous Thromboembolism

Hypothesis	Model no.	Allele frequency	Transmission probabilities			Susceptibilities			Residual association		Covariate			*df*	−2 ln L	AIC	χ^2 ($\triangle df$)	*P* value
		q_A	τ_{AA}	τ_{AB}	τ_{BB}	β_{AA}	β_{AB}	β_{BB}	$\gamma_{fo}=\gamma_{mo}$	γ_{ss}	Sex	Age	Generation					
Sporadic	1	[0]	…	…	…	−2.83	[=β_{AA}]	[=β_{AA}]	…	…	0.08	0.03	−0.48	4	1 967 670	1 967 678	72 630 (6)	<0.001
Polygenic	2	[0]	…	…	…	−2.92	[=β_{AA}]	[=β_{AA}]	0.64	0.65	0.08	0.03	−0.41	6	1 963 110	1 963 122	68 070 (4)	<0.001
Polygenic-environmental	3	0.12	[0.12]	[0.12]	[0.12]	−426.40	−129.97	−2.67	1.35	8.44	0.08	0.03	−036	9	1 933 906	1 933 924	38 866 (1)	<0.001
Mendelian codominant	4	0.21	1.00	0.50	0.00	−3.89	−1.72	−35.70	…	…	0.11	0.04	−0.43	7	1 962 146	1 962 160	67 106 (3)	<0.001
Mendelian dominant	5	0.24	1.00	0.50	0.00	−2.01	[=β_{AA}]	−19.64	…	…	0.10	0.04	−0.40	6	1 962 712	1 962 724	67 672 (4)	<0.001
Mendelian recessive	6	0.76	1.00	0.50	0.00	−43.92	−2.01	[=β_{AB}]	…	…	0.10	0.04	−0.40	6	1 962 712	1 962 724	67 672 (4)	<0.001
Mendelian additive	7	0.68	1.00	0.50	0.00	−2.09	−18.04	−33.98	…	…	0.10	0.04	−0.41	6	1 964 018	1 964 030	68 978 (4)	<0.001
Mendelian plus polygenic: codominant	8	0.66	1.00	0.50	0.00	−141.91	−2.24	−3.03	0.78	1.20	0.09	0.04	−0.36	9	1 956 138	1 956 156	61 098 (1)	<0.001
Mendelian plus polygenic: dominant	9	0.26	1.00	0.50	0.00	−84.47	[=β_{AA}]	−2.32	0.41	0.79	0.09	0.04	−0.39	8	1 962 036	1 962 052	66 996 (2)	<0.001
Mendelian plus polygenic: recessive	10	0.61	1.00	0.50	0.00	−145.21	−2.45	[=β_{AB}]	0.67	1.26	0.09	0.03	−0.40	8	1 958 102	1 958 118	63 062 (2)	<0.001
Mendelian plus polygenic: additive	11	0.74	1.00	0.50	0.00	−2.32	−111.02	−219.72	0.41	0.79	0.09	0.04	−0.39	8	1 962 036	1 962 052	66 996 (2)	<0.001
Major type only	12	1.00*	0.04	0.90	1.00*	−2.65	−3.26	−0.64	…	…	0.10	0.01	−0.94	10	**1 895 040**	**1 895 062**	…	…
General model	13	0.74	0.00*	0.00*	0.35	-76×10^8	−2.35	−1.85	0.90	4.10	0.09	0.04	−0.46	12	1 896 710	1 896 734	1 670 (2)	<0.001

There were 42 284 pedigrees with inherited venous thromboembolism (ie, at least 2 individuals affected by venous thromboembolism). The lowest AIC in each model was observed for the major-type-only model. The differences were significant ($P<0.001$). Bold indicates that this model had the lowest −2 ln L and AIC values, respectively. Familial correlations; father-offspring (γ_{fo}), mother-offspring (γ_{mo}), and sib-sib (γ_{ss}) correlations. AA, AB, BB: the type, or genotype if the segregation is Mendelian. AIC indicates Akaike information criterion; $\triangle df$, difference in degree of freedom between compared models; q_A, the frequency of A.

*Independent fixed at bound.

important to recognize that this limitation is present in all family studies that include cases from early periods.

Another limitation is that we were not able to control for several VTE risk factors that has changed in frequency during the study period such as overweight/obesity, height, smoking, arterial cardiovascular disorders, hormone replacement therapy, and oral contraceptives use. Moreover, because of limitations of the software (ie, Statistical Analysis for Genetic Epidemiology), pedigrees with twins, loops, and >3 mates were excluded. It is possible that the unexpected low frequency of affected women in the second generation is related to some of these uncontrolled factors (Table 1 and Table S1).

Strengths

The large study size of the present study is a major advantage. Another strength is the use of validated hospital discharge data.[19,26] The validity is 90% to 95% for cardiovascular disorders such as VTE, stroke, and myocardial infarction.[19,26] Another advantage is that the use of objective clinical hospital diagnoses allows for the elimination of any recall bias. Recall and self-report bias are common problems in many family studies of VTE.[44] Use of the multigeneration register is therefore an advantage.[20] Swedish registers such as the Swedish Total Population Register and the Swedish Hospital Discharge Register are highly complete.[18–23]

A key strength is that the results of the present study are in line with findings from genetic studies based on genomic data[4–6,9–11] and the only previous segregation study of VTE by Heit et al.[16] Compared with the only previous complex segregation study by Heit et al, who used the Pedigree Analysis Package, we were able to statistically compare the models because of the large sample size.[16] Moreover, the program SEGREG within the 64-bit Linux software package Statistical Analysis for Genetic Epidemiology 6.4, which was used in the present study, allowed for adjustment and test of suitable covariates and also for selection of the most appropriate familial correlations. A further strength is that we did not exclude pedigrees with probands with cancer as in the Heit study.[16] In the Heit study, family history was based on self-report by the proband, and probands were all referred to a highly specialized coagulation clinic at the Mayo clinic. The lower heritability in the present study might be attributable to these differences in study design and sample size. A strength is that age- and sex-specific prevalence constraints were imposed on all models.

CONCLUSIONS

The heritability of VTE in the general population in Sweden exhibits a Carter effect, with higher heritability among young individuals and men. In addition, segregation analysis discards a pure nongenetic cause of VTE. The findings from the Swedish pedigrees with inherited VTE are in line with a single major-type-only model transmitted with the possibility of some polygenic effects. The present study suggests that further genetic studies of families with familial aggregation of VTE might be worthwhile.

ARTICLE INFORMATION

Received December 7, 2020; accepted November 15, 2021.

Affiliations

Center for Primary Health Care Research, Lund University/Region Skåne, Malmö, Sweden (B.Z., M.P., E.S., J.S., K.S.); Department of Coagulation Disorders, Skåne University Hospital, Lund University, Malmö, Sweden (P.J.S.); Department of Statistics, Lund University, Lund, Sweden (B.H.); and Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, OH (R.C.E., Y.E.S.).

Acknowledgments

The authors thank P. Reilly, science editor at Center for Primary Health Care Research, Department of Clinical Sciences, Lund University and Region Skåne, Malmö, Sweden, for his comments on the text. The registers used in the present study are maintained by Statistics Sweden and the National Board of Health and Welfare.

Author Contributions: All authors contributed to the conception and design of the study. Drs J. Sundquist and K. Sundquist contributed to the acquisition of data. All authors contributed to the analysis and interpretation of data. Dr Zöller drafted the article, and all authors revised it critically and approved the final version. Drs Pirouzifard and Zöller had full access to all of the data (including statistical reports and tables) and take responsibility for the integrity of the data and the accuracy of their analysis.

Sources of Funding

This work was supported by grants awarded to Drs Zöller and K. Sundquist by ALF (Avtal om Läkarutbildning och Forskning) funding from Region Skåne; grants awarded to Drs Zöller, K. Sundquist, and J. Sundquist by the Swedish Research Council; a grant to Dr Zöller from Sparbanken Skåne; and a grant awarded to Dr K. Sundquist by the Swedish Heart-Lung Foundation.

Disclosures

None.

Supplementary Material

Table S1

REFERENCES

1. Heit JA. Epidemiology of venous thromboembolism. *Nat Rev Cardiol*. 2015;12:464–474. doi: 10.1038/nrcardio.2015.83
2. Zöller B, Li X, Ohlsson H, Ji J, Sundquist J, Sundquist K. Family history of venous thromboembolism as a risk factor and genetic research tool. *Thromb Haemost*. 2015;114:890–900. doi: 10.1160/TH15-04-0306
3. Rosendaal FR. Venous thrombosis: a multicausal disease. *Lancet*. 1999;353:1167–1173. doi: 10.1016/S0140-6736(98)10266-0
4. Zöller B, García de Frutos P, Hillarp A, Dahlbäck B. Thrombophilia as a multigenic disease. *Haematologica*. 1999;84:59–70.
5. Zöller B, Svensson PJ, Dahlbäck B, Lind-Hallden C, Hallden C, Elf J. Genetic risk factors for venous thromboembolism. *Expert Rev Hematol*. 2020;13:971–981. doi: 10.1080/17474086.2020.1804354
6. Mannucci PM, Franchini M. Classic thrombophilic gene variants. *Thromb Haemost*. 2015;114:885–889. doi: 10.1160/TH15-02-0141
7. Bezemer ID, van der Meer FJM, Eikenboom JC, Rosendaal FR, Doggen CJ. The value of family history as a risk indicator for venous thrombosis. *Arch Intern Med*. 2009;169:610–615. doi: 10.1001/archinternmed.2008.589
8. Desch KC, Ozel AB, Halvorsen M, Jacobi PM, Golden K, Underwood M, Germain M, Tregouet D-A, Reitsma PH, Kearon C, et al. Whole-exome

sequencing identifies rare variants in STAB2 associated with venous thromboembolic disease. *Blood*. 2020;136:533–541. doi: 10.1182/blood.2019004161
9. Morange PE, Suchon P, Trégouët DA. Genetics of venous thrombosis: update in 2015. *Thromb Haemost*. 2015;114:910–919. doi: 10.1160/TH15-05-0410
10. Lindström S, Wang LU, Smith EN, Gordon W, van Hylckama Vlieg A, de Andrade M, Brody JA, Pattee JW, Haessler J, Brumpton BM, et al. Genomic and transcriptomic association studies identify 16 novel susceptibility loci for venous thromboembolism. *Blood*. 2019;134:1645–1657. doi: 10.1182/blood.2019000435
11. Klarin D, Busenkell E, Judy R, Lynch J, Levin M, Haessler J, Aragam K, Chaffin M, Haas M, Lindström S, et al. Genome-wide association analysis of venous thromboembolism identifies new risk loci and genetic overlap with arterial vascular disease. *Nat Genet*. 2019;51:1574–1579.
12. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, et al. Finding the missing heritability of complex diseases. *Nature*. 2009;461:747–753. doi: 10.1038/nature08494
13. Lawlor DA, Mishra GD. *Family Matters. Deigning, Analyzing and Understanding Family Based Studies in Life Course Epidemiology*. Oxford University Press; 2009.
14. Larsen TB, Sørensen HT, Skytthe A, Johnsen SP, Vaupel JW, Christensen K. Major genetic susceptibility for venous thromboembolism in men: a study of Danish twins. *Epidemiology*. 2003;14:328–332. doi: 10.1097/01.EDE.0000060457.51194.BC
15. Souto JC, Almasy L, Borrell M, Blanco-Vaca F, Mateo J, Soria JM, Coll I, Felices R, Stone W, Fontcuberta J, et al. Genetic susceptibility to thrombosis and its relationship to physiological risk factors: the GAIT study. Genetic Analysis of Idiopathic Thrombophilia. *Am J Hum Genet*. 2000;67:1452–1459. doi: 10.1086/316903
16. Heit JA, Phelps MA, Ward SA, Slusser JP, Petterson TM, De Andrade M. Familial segregation of venous thromboembolism. *J Thromb Haemost*. 2004;2:731–736. doi: 10.1111/j.1538-7933.2004.00660.x
17. Zöller B, Ohlsson H, Sundquist J, Sundquist K. A sibling based design to quantify genetic and shared environmental effects of venous thromboembolism in Sweden. *Thromb Res*. 2017;149:82–87. doi: 10.1016/j.thromres.2016.10.014
18. Rosen M, Hakulinen T. Use of disease registers. In: Ahrens W, Pigeot I, eds. *Handbook of Epidemiology*. Springer-Verlag; 2005:231–252.
19. Ludvigsson JF, Andersson E, Ekbom A, Feychting M, Kim JL, Reuterwall C, Heurgren M, Olausson PO. External review and validation of the Swedish national inpatient register. *BMC Public Health*. 2011;11:450. doi: 10.1186/1471-2458-11-450
20. Ekbom A. The Swedish Multi-generation Register. *Methods Mol Biol*. 2011;675:215–220.
21. Zöller B. Nationwide family studies of cardiovascular diseases—clinical and genetic implications of family history. *EMJ Cardiol*. 2013;1:102–113.
22. Ludvigsson JF, Almqvist C, Bonamy AK, Ljung R, Michaëlsson K, Neovius M, Stephansson O, Ye W. Registers of the Swedish total population and their use in medical research. *Eur J Epidemiol*. 2016;31:125–136.
23. Ludvigsson JF, Otterblad-Olausson P, Pettersson BU, Ekbom A. The Swedish personal identity number: possibilities and pitfalls in healthcare and medical research. *Eur J Epidemiol*. 2009;24:659–667. doi: 10.1007/s10654-009-9350-y
24. Zöller B, Li X, Sundquist J, Sundquist K. Age- and gender-specific familial risks for venous thromboembolism: a nationwide epidemiological study based on hospitalizations in Sweden. *Circulation*. 2011;124:1012–1020. doi: 10.1161/CIRCULATIONAHA.110.965020
25. Zöller B, Ohlsson H, Sundquist J, Sundquist K. Familial risk of venous thromboembolism in first-, second- and third-degree relatives: a nationwide family study in Sweden. *Thromb Haemost*. 2013;109:458–463. doi: 10.1160/TH12-10-0743
26. Rosengren A, Fredén M, Hansson P-O, Wilhelmsen L, Wedel H, Eriksson H. Psychosocial factors and venous thromboembolism: a long-term follow-up study of Swedish men. *J Thromb Haemost*. 2008;6:558–564. doi: 10.1111/j.1538-7836.2007.02857.x
27. Tenesa A, Haley CS. The heritability of human disease: estimation, uses and abuses. *Nat Rev Genet*. 2013;14:139–149. doi: 10.1038/nrg3377
28. Falconer DS. The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Ann Hum Genet*. 1965;29:51–76. doi: 10.1111/j.1469-1809.1965.tb00500.x
29. Falconer Falconer DS, MacKay TFC. *Introduction to Quantitative Genetics*. Pearson Educated Limited; 1996.
30. Elston RC, Gray-McGuire C. A review of the 'Statistical Analysis for Genetic Epidemiology' (S.A.G.E.) software package. *Hum Genomics*. 2004;1:456–459.
31. S.A.G.E. Statistical Analysis for Genetic Epidemiology (S.A.G:E) Version 6.4 User Reference Manual. Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH, 2016. Available at: http://darwin.cwru.edu/sage/index.php. Accessed September 1, 2021.
32. Karunaratne PM, Elston RC. A multivariate logistic model (MLM) for analyzing binary family data. *AM J Med Genet*. 1998;76:428–437. doi: 10.1002/(SICI)1096-8628(19980413)76:5<428:AID-AJMG12>3.0.CO;2-O
33. Sun X. Segregation analysis using the unified model. In: Elston RC, ed. *Statistical Human Genetics. Methods and Protocols*. Humana Press; 2017:233–256.
34. Jiang Y, Lv Q, Rao S, Liao Z, Zhang P, Yang M, Li Q, Cao S, Lin Z, Gu J. Complex segregation analysis provides evidence for autosomal dominant transmission in the Chinese Han families with ankylosing spondylitis. *Biomed Res Int*. 2017;2017:4515701. doi: 10.1155/2017/4515701
35. Tayo BO, DiCioccio RA, Liang Y, Trevisan M, Cooper RS, Lele S, Sucheston L, Piver SM, Odunsi K. Complex segregation analysis of pedigrees from the Gilda Radner Familial Ovarian Cancer Registry reveals evidence for Mendelian dominant inheritance. *PLoS One*. 2009;4:e5939.
36. Carter CO, Evans KA. Inheritance of congenital pyloric stenosis. *J Med Genet*. 1969;6:233–254. doi: 10.1136/jmg.6.3.233
37. Svensson PJ, Zöller B, Mattiasson I, Dahlbäck B. The factor VR506Q mutation causing APC resistance is highly prevalent amongst unselected outpatients with clinically suspected deep venous thrombosis. *J Intern Med*. 1997;241:379–385. doi: 10.1046/j.1365-2796.1997.124140000.x
38. Hillarp A, Zöller B, Svensson PJ, Dahlbäck B. The 20210 A allele of the prothrombin gene is a common risk factor among Swedish outpatients with verified deep venous thrombosis. *Thromb Haemost*. 1997;78:990–992. doi: 10.1055/s-0038-1657674
39. Zöller B. Prevalence and in silico analysis of missense mutations in the PROS1 gene in the Swedish population: the SweGen dataset. *Thromb Res*. 2018;168:28–30.
40. Zöller B, Berntsdotter A, García de Frutos P, Dahlbäck B. Resistance to activated protein C as an additional genetic risk factor in hereditary deficiency of protein S. *Blood*. 1995;85:3518–3523.
41. Zöller B, García de Frutos P, Dahlbäck B. Evaluation of the relationship between protein S and C4b-binding protein isoforms in hereditary protein S deficiency demonstrating type I and type III deficiencies to be phenotypic variants of the same genetic disease. *Blood*. 1995;85:3524–3531. doi: 10.1182/blood.V85.12.3524.bloodjournal85123524
42. Manderstedt E, Lind-Halldén C, Svensson P, Zöller B, Halldén C. Next-generation sequencing of 17 genes associated with venous thromboembolism reveals a deficit of non-synonymous variants in procoagulant genes. *Thromb Haemost*. 2019;119:1441–1450. doi: 10.1055/s-0039-1693130
43. Manderstedt E, Halldén C, Lind-Halldén C, Elf J, Svensson PJ, Engström G, Melander O, Baras A, Lotta LA, Zöller B; Regeneron Genetics Center*. Thrombotic risk determined by STAB 2 variants in a population-based cohort study. *Circ Genom Precis Med*. 2021;14:e003449.
44. Weinstock MA, Brodsky GL. Bias in the assessment of family history of melanoma and its association with dysplastic nevi in a case-control study. *J Clin Epidemiol*. 1998;51:1299–1303.
45. Lander ES, Schork NJ. Genetic dissection of complex traits. *Science*. 1994;265:2037–2048. doi: 10.1126/science.8091226
46. Rothman KJ. No adjustments are needed for multiple comparisons. *Epidemiology*. 1990;1:43–46. doi: 10.1097/00001648-199001000-00010

SUPPLEMENTAL MATERIAL

Table S1. Descriptive statistics of the study sample (population of Swedish-born, n=7,568,818) with and without venous thromboembolism (VTE) during study period 1964-2015.

	VTE cases	No VTE
Numbers (%)		
All	284515 (100)	7284303 (100)
Grandparents	165021 (58.0)	1637211 (22.5)
Offspring	87177 (30.6)	2108359 (28.9)
Grandchildren	32317 (11.4)	3538733 (48.6)
*Age at end of follow-up median [q1–q3] (min–max)		
All	74 [61–83] (0–110)	49 [31–69] (0–110)
Grandparents	81 [74–87] (20–110)	79 [72–85] (18–110)
Offspring	65 [57–71] (1–83)	60 [52–67] (0–83)
Grandchildren	42 [34–48] (0–65)	31 [23–41] (0–67)
Age at first VTE median [q1–q3] (min–max)		
All	66 [52–76] (0–104)	–
Grandparents	74 [66–81] (18–104)	–
Offspring	57 [47–65] (0–83)	–
Grandchildren	32 [25–41] (0–64)	–
Female, Sex, n (%)		
All	149257 (52.5)	3557388 (48.8)
Grandparents	88150 (53.4)	812966 (49.7)
Offspring	41474 (47.6)	1030269 (48.9)
Grandchildren	19633 (60.8)	1714153 (48.4)

Individuals could be included more than once if they occur in more than one pedigree. In Table 1 each individual appears only once.

*End of follow-up is either age at end of study 2015, age at death, or age at emigration, depending on which occurred first. [q1–q3]=Interquartile range (IQR); (min–max)=range; Q1=fist quartile; Q3=third quartile.