Association of Genetic Variations With Nonfatal Venous Thrombosis in Postmenopausal Women

Nicholas L. Smith, PhD, MPH
Lucia A. Hindorff, PhD
Susan R. Heckbert, MD, PhD
Rozenn N. Lemaitre, PhD, MPH
Kristin D. Marciante, PhD
Kenneth Rice, PhD
Thomas Lumley, PhD
Joshua C. Bis, MS
Kerri L. Wiggins, MS, RD
Bruce M. Psaty, MD, PhD

Deep vein thrombosis and pulmonary embolism are significant contributors to morbidity and mortality in adult women and occur at annual rates of 3 to 5 per 1000 person-years among women older than 55 years.\(^1,2\) Despite improved prophylaxis in high-risk patients, the incidence of venous thrombosis (VT) has not decreased.\(^1\)

Causes of VT primarily involve modifications in the coagulation, anticoagulation, fibrinolysis, and antifibrinolytic pathways regulating hemostasis. Although the roles of pathway proteins and enzymes that activate or inhibit fibrin production and lysis are well characterized, the underlying contribution of genetic variation in these constituents to risk of venous thrombosis (VT) has not been fully investigated.\(^1\)

Objective To describe the association of common genetic variation in 24 coagulation, anticoagulation, fibrinolysis, and antifibrinolysis candidate genes with risk of incident nonfatal VT in postmenopausal women.

Design, Setting, and Participants Population-based case-control study conducted in a large integrated health care system in Washington State. Participants were perimenopausal and postmenopausal women aged 30 to 89 years who sustained a first VT event between January 1995 and December 2002 (n=349) and 1680 controls matched on age, hypertension status, and calendar year (n=1680).

Main Outcome Measure Risk of venous thrombosis associated with global variation within a gene as measured by common haplotypes and with individual haplotypes and single nucleotide polymorphisms (SNPs). Significance of the associations was assessed by a .20 threshold of the false-discovery rate \(q\) value, which accounts for multiple testing.

Results Only the tissue factor pathway inhibitor gene demonstrated global association with risk \((q=.13)\). Five significant SNP associations were identified across 3 of the candidate genes (factors V, XI, and protein C) in SNP analyses. Two associations have been previously reported. Another 22 variants across 15 genes had \(P\) values less than .05 but \(q\) values between .20 and .35. Five of these confirm previously reported associations (fibrinogen genes and protein C), 2 were inconsistent with earlier reports (thrombomodulin and plasminogen activator inhibitor 1), and 15 were new discoveries.

Conclusions After accounting for multiple testing, 5 SNPs associated with VT risk were identified, 3 of which have not been previously reported. Replication of these novel associations in other populations is necessary to corroborate these findings and identify which genetic factors may influence VT risk in postmenopausal women.

JAMA. 2007;297:489-498 www.jama.com

Author Affiliations: Departments of Epidemiology (Drs Smith, Hindorff, Heckbert, and Psaty and Mr Bis), Medicine (Drs Lemaitre, Marciante, and Psaty and Ms Wiggins), Biostatistics (Drs Rice and Lumley), and Health Services (Dr Psaty), University of Washington, Seattle; and Departments of Clinical Epidemiology and Hematology, Leiden University Medical Center, Leiden, the Netherlands (Dr Rosendaal).

Context Although the roles of clotting proteins and enzymes that activate or inhibit fibrin production and lysis are well characterized, the underlying contribution of genetic variation in these constituents to risk of venous thrombosis (VT) has not been fully investigated.

Objective To describe the association of common genetic variation in 24 coagulation, anticoagulation, fibrinolysis, and antifibrinolysis candidate genes with risk of incident nonfatal VT in postmenopausal women.

Design, Setting, and Participants Population-based case-control study conducted in a large integrated health care system in Washington State. Participants were perimenopausal and postmenopausal women aged 30 to 89 years who sustained a first VT event between January 1995 and December 2002 (n=349) and 1680 controls matched on age, hypertension status, and calendar year (n=1680).

Main Outcome Measure Risk of venous thrombosis associated with global variation within a gene as measured by common haplotypes and with individual haplotypes and single nucleotide polymorphisms (SNPs). Significance of the associations was assessed by a .20 threshold of the false-discovery rate \(q\) value, which accounts for multiple testing.

Results Only the tissue factor pathway inhibitor gene demonstrated global association with risk \((q=.13)\). Five significant SNP associations were identified across 3 of the candidate genes (factors V, XI, and protein C) in SNP analyses. Two associations have been previously reported. Another 22 variants across 15 genes had \(P\) values less than .05 but \(q\) values between .20 and .35. Five of these confirm previously reported associations (fibrinogen genes and protein C), 2 were inconsistent with earlier reports (thrombomodulin and plasminogen activator inhibitor 1), and 15 were new discoveries.

Conclusions After accounting for multiple testing, 5 SNPs associated with VT risk were identified, 3 of which have not been previously reported. Replication of these novel associations in other populations is necessary to corroborate these findings and identify which genetic factors may influence VT risk in postmenopausal women.

JAMA. 2007;297:489-498 www.jama.com

Author Affiliations: Departments of Epidemiology (Drs Smith, Hindorff, Heckbert, and Psaty and Mr Bis), Medicine (Drs Lemaitre, Marciante, and Psaty and Ms Wiggins), Biostatistics (Drs Rice and Lumley), and Health Services (Dr Psaty), University of Washington, Seattle; and Departments of Clinical Epidemiology and Hematology, Leiden University Medical Center, Leiden, the Netherlands (Dr Rosendaal).

Context Although the roles of clotting proteins and enzymes that activate or inhibit fibrin production and lysis are well characterized, the underlying contribution of genetic variation in these constituents to risk of venous thrombosis (VT) has not been fully investigated.

Objective To describe the association of common genetic variation in 24 coagulation, anticoagulation, fibrinolysis, and antifibrinolysis candidate genes with risk of incident nonfatal VT in postmenopausal women.

Design, Setting, and Participants Population-based case-control study conducted in a large integrated health care system in Washington State. Participants were perimenopausal and postmenopausal women aged 30 to 89 years who sustained a first VT event between January 1995 and December 2002 (n=349) and 1680 controls matched on age, hypertension status, and calendar year (n=1680).

Main Outcome Measure Risk of venous thrombosis associated with global variation within a gene as measured by common haplotypes and with individual haplotypes and single nucleotide polymorphisms (SNPs). Significance of the associations was assessed by a .20 threshold of the false-discovery rate \(q\) value, which accounts for multiple testing.

Results Only the tissue factor pathway inhibitor gene demonstrated global association with risk \((q=.13)\). Five significant SNP associations were identified across 3 of the candidate genes (factors V, XI, and protein C) in SNP analyses. Two associations have been previously reported. Another 22 variants across 15 genes had \(P\) values less than .05 but \(q\) values between .20 and .35. Five of these confirm previously reported associations (fibrinogen genes and protein C), 2 were inconsistent with earlier reports (thrombomodulin and plasminogen activator inhibitor 1), and 15 were new discoveries.

Conclusions After accounting for multiple testing, 5 SNPs associated with VT risk were identified, 3 of which have not been previously reported. Replication of these novel associations in other populations is necessary to corroborate these findings and identify which genetic factors may influence VT risk in postmenopausal women.

JAMA. 2007;297:489-498 www.jama.com

Author Affiliations: Departments of Epidemiology (Drs Smith, Hindorff, Heckbert, and Psaty and Mr Bis), Medicine (Drs Lemaitre, Marciante, and Psaty and Ms Wiggins), Biostatistics (Drs Rice and Lumley), and Health Services (Dr Psaty), University of Washington, Seattle; and Departments of Clinical Epidemiology and Hematology, Leiden University Medical Center, Leiden, the Netherlands (Dr Rosendaal).

Corresponding Author: Nicholas L. Smith, PhD, MPH, Cardiovascular Health Research Unit, University of Washington, 1730 Minor Ave, Suite 1360, Seattle, WA 98101 (nlsmith@u.washington.edu).

©2007 American Medical Association. All rights reserved.

(Reprinted) JAMA, February 7, 2007—Vol 297, No. 5 489
4G/5G variant in plasminogen activator inhibitor 1 (PAI-1), haplotype 2 (HR2) of factor V, promoter variants in protein C, and variants in thrombomodulin and the 3 fibrinogen genes, have shown weak and inconsistent associations across populations or have not been adequately replicated in other populations.8–15

The aim of this study was to describe the association of common genetic variation in 24 clotting-related candidate genes with risk of first-time VT in postmenopausal women. The approach was 2-tiered: using a case-control design, the global association of haplotypes within a gene with VT risk was first investigated; then the association of individual haplotypes and single nucleotide polymorphisms (SNPs) with risk was assessed.

METHODS

Setting and Design

The setting for this observational study was Group Health, a large integrated health care system in western Washington State. This study is part of an ongoing, population-based, case-control study of myocardial infarction, stroke, atrial fibrillation, and VT.16,17 The analyses presented are part of a larger study examining the association of hormone use with VT risk in postmenopausal women. This study was approved by the Group Health human subjects review committee, and all study participants provided written informed consent.

Study Population

Participants were perimenopausal and postmenopausal women aged 30 to 89 years who were members of Group Health. Case participants were members who experienced a first incident deep vein thrombosis or pulmonary embolism event between January 1, 1995, and December 31, 2002, and who were alive at the time of study recruitment. The date of VT diagnosis served as an index date. Control participants were a random sample of members comprising a pool shared by several case-control studies conducted at Group Health.16,17 The control group was frequency matched to myocardial infarction cases, the largest case group, by age (within decade), sex, treated hypertension status, and calendar year of identification. All controls included in this analysis were perimenopausal and postmenopausal women aged 30 to 89 years who had no prior history of deep vein thrombosis or pulmonary embolism. For control participants, the index date was a randomly chosen date within the calendar year from which each participant was selected as a control.

Women with a potentially eligible VT event were screened using administrative records from inpatient and outpatient care settings. In the patient setting, International Classification of Diseases, Ninth Revision codes were obtained from Group Health hospitalization records, which included hospital stays at Group Health and outside facilities. In the outpatient setting, Group Health pharmacy records were used to identify women who were dispensed a prescription for a low-molecular-weight heparin (the pathway required outpatient treatment for deep vein thrombosis). Trained medical record data abstractors reviewed the medical records of all potential cases to verify the diagnosis of VT and to determine how the diagnosis was made. Of the 349 eligible cases, 334 (95%) had a diagnosis that was confirmed with a diagnostic imaging test result and only the diagnosis of a physician.

Measures

Genetic Variation. The 24 candidate genes were selected a priori for investigation based on their role in regulating venous hemostasis. The genes code for proteins affecting coagulation (factors II, V, VII, VIII, IX, X, XI, XII, XIIIa1, and XIIIb; fibrinogen α, β, and γ; and tissue factor), anticoagulation (antithrombin, proteins C and S, endothelial protein C receptor, thrombomodulin, and tissue factor pathway inhibitor [TFPI]), fibrinolysis (plasminogen and tissue-type plasminogen activator), and antifibrinolysis (PAI-1 and thrombin activatable fibrinolysis inhibitor [TAFI]). All genes were sequenced by the Seattle Program for Genomic Applications using DNA obtained from a panel of 24 black and 23 white individuals. Genes were sequenced starting 2.5 kilobases upstream of the start of transcription and extending 1.5 kilobases beyond the end of transcription. For large genes, sequencing was limited to exons and promoter regions. Details are available at http://pga.gs.washington.edu/.

Among variants with a minor allele frequency of 5% or greater in the Program for Genomic Applications panel, LDSelect version 1.0 (University of Washington, Seattle; http://pga.gs.washington.edu/) was used to identify maximally informative sets of SNPs to describe genetic variation using linkage disequilibrium.18 This process was performed separately for sequence data from black and white individuals. From within race-stratified SNP sets, a single variant was hierarchically chosen based on the following priorities: nonsynonymous, untranslated regions (5' or 3'), promoter, intron, and synonymous. In addition, we identified nonsynonymous SNPs that were not already selected and that had a minor allele frequency of 4% or greater in either racial group.

Blood Collection and Genotyping. A blood sample was collected from all participants, and DNA was extracted from white blood cells using standard procedures. Genotyping was performed by Illumina Inc with a GoldenGate custom panel using BeadArray technology (Illumina Inc, San Diego, Calif). Illumina personnel were blinded to case-control status. The final list of SNPs in the 24 genes numbered 306, of which 280 were successfully genotyped on the Illumina platform. Among the 280 successfully genotyped SNPs, 99.94% of nucleotide pairs were successfully called. Agreement was high between Illumina results and results from a substudy of 1472 participants that used restricted fragment-length tech-
nology to perform genotyping of 2 variants (factor V Leiden and factor II G20210A) (κ = 0.99, and 1.00, respectively).

**Haplotypic Construction and SNP Selection.** Phased haplotypic were estimated for each gene using PHASE2.0 software (University of Washington, Seattle; http://www.stat.washington.edu/stephens/software.html), which computes probabilities for each haplotype pair consistent with the observed data. When ambiguous haplotypic were encountered, multiple probability-weighted haplotype pairs were created. Haplotypic that had a frequency of less than 2% in case and control populations were combined into a single “other” haplotype. The factor V gene was split into 2 sections below a point of high recombination (Program for Genomic Applications site 36090) to minimize the number of haplotypic below the 2% threshold.

Among the 280 SNPs that were genotyped in the study population, 108 were found to be either extremely rare or nonvarying, or tightly linked (r² ≥ 0.8) with another SNP. Many were SNPs prevalent in black individuals. These variants were excluded from SNP analyses, leaving 172 SNPs characterizing genetic variation across the 24 candidate genes. A table presenting reference SNP accession identifiers and Program for Genomic Applications identifiers for all SNPs is available from the authors at http://depts.washington.edu/chru.

**Demographic and Clinical Information.** Demographic and health-status information was obtained by review of the entire Group Health ambulatory medical record up to the index date. Information was collected on treated hypertension, outpatient surgical or major diagnostic procedures, and menopause status. Menopause was defined by the cessation of ovarian function that occurred naturally or following bilateral oophorectomy. If menopausal status was not explicitly stated in the record, women 55 years and older were considered postmenopausal. Self-reported information on race, which can confound genetic association studies, was collected from the participant by telephone interview using 4 race categories.

Data collection from the ambulatory medical record was supplemented with information from several Group Health databases. In particular, cancer history information was collected from a Group Health cancer registry file, based on the Surveillance Epidemiology and End Results registry, that included all diagnosed cancers except nonmelanoma skin cancers. Information on hospitalizations and bone fractures before the index date was collected from Group Health administrative files that included diagnoses from inpatient and outpatient care delivered at Group Health and outside facilities.

**Statistical Analyses**

**Haplotype and SNP Models.** For each gene, weighted logistic regression was used to estimate haplotype associations. To express haplotype uncertainty, probability weights produced by PHASE 2.0 and robust “sandwich” errors clustered by participant were used. The number of copies of each haplotype (0, 1, or 2) was used as the predictor, and the most common haplotype in the gene was used as the reference level and not included in the model. The estimates were adjusted for race and for the matching variables age, index year, and hypertension. For each haplotype, this model provides an estimate of the risk associated with each additional copy of the specified haplotype. These estimates are presented as odds ratios (ORs) with 95% confidence intervals (CIs). To test global gene-level variation within a gene, we compared the haplotype model that characterizes all common variation within the gene with a model that included no haplotype information. For each gene, a Wald test was used to compare these models; this tests the null hypothesis that all gene-wide variants have an OR equal to 1. For the factor II and factor V genes, we performed this global test twice, first including all study participants and then excluding those carrying the factor II 20210A and factor V Leiden variants. In addition, for each SNP, unweighted logistic regression models were used to estimate the association with VT risk while adjusting for race and matching variables. A single predictor coded as 0, 1, or 2 (representing the number of copies of the minor allele) was created for each SNP, giving estimates of the OR associated with each additional copy of the minor allele.

To quantify the amount of the overall risk explained by an additional copy of a specified SNP or haplotype, we calculated population-attributable risk percentages for all SNP and haplotype associations with P values less than .05. The percentages were based on relative risk estimates from models described above and on population minor allele prevalence estimates from controls.

Models were not adjusted for acquired risk factors because these do not generally confound genetic associations except through selection bias. However, these risk factors may modify the gene-VT association, so the estimated effects were compared in participants with and without known acquired risk factors (recent cancer diagnosis, hospitalization, major fracture, or outpatient surgical or major diagnostic procedures; estrogen use; or body mass index ≥ 35, calculated as weight in kilograms divided by height in meters squared) in sensitivity analyses and then tested for interactions.

**Multiple Testing and False Discovery Rate Analysis.** Twenty-five global gene association tests were performed on the 24 candidate genes (the factor V gene was split into 2 parts). The a priori expectation was that factor II 20210A and factor V Leiden would be associated with risk of VT, so factor II and the upper half of factor V (where the Leiden variant is located) were tested after removing participants with these variants. Variants in other genes that have shown weak and inconsistent associations or that have not been replicated adequately in other populations were not excluded. Even in the absence of a true association, we would...
expect a P value less than .05 in at least 1 of the 25 gene-wide associations. A total of 173 haplotypes and 170 SNPs were tested for individual associations across the 24 genes after excluding SNPs and haplotypes marked by factor V Leiden and factor II 20210A. Therefore, by chance alone, P values less than .05 would be expected in 9 of the haplotype-specific tests and in at least 9 of the SNP-specific tests.

To account for multiple statistical testing, false discovery rate (FDR) q values were calculated separately for genome-wide haplotype and SNP analyses. To calculate the q value, the P values are modified by multiplying them by the number of tests performed and then dividing them by the rank order of each P value (where rank order 1 is assigned to the smallest P value). For each test, the q value is defined as the smallest modified P value among tests with equal or larger rank. This definition ensures that the q values have the same ordering as the P values. Rather than expressing the probability of a single false-positive result among all tests, the q value estimates the proportion of results declared interesting that are actually false. For example, if 100 statistical tests are performed with an α of .05, then 5 tests on average would produce P values less than .05 by chance alone. If 10 of these 100 tests produce P values of exactly .05, then all 10 would have a q value of .50. If all 10 were declared interesting, this q value indicates that up to 50% of discoveries would be expected to be false. No conventional q value threshold has been established to separate false discoveries from true discoveries in candidate gene investigations. We used a q value threshold of .20 and so should expect up to only 20% of declared discoveries to be false. A lower threshold is often used in genome-wide association studies, which do not use prior information to select candidate genes.

**Interesting Haplotype or SNP Variants.** Measures of global association are reported for each of the 24 genes. However, given 173 haplotypes and 170 SNPs to assess, key findings must be summarized. After computing P values and q values for all haplotypes and SNPs separately, haplotypes and SNPs with P values less than .05 in the models described above were identified and listed. In genes in which the listed haplotype association was primarily attributable to a single SNP, we delisted the haplotype to avoid redundancy of information. For genes with listed haplotypes, we then reestimated the haplotype effect by fitting a model that merged all other (nonlisted) haplotypes. This changed the reference group, providing an estimate of risk relative to all other haplotypes. Haplotypes with a recalculated P value greater than .05 were also delisted. For listed haplotypes and SNPs with a minor allele frequency greater than 10%, we also evaluated risk estimates for dominant and recessive associations. These risk estimates come from a 2-parameter model, with indicator variables for heterozygotes and minor allele homozygotes.

**Power.** Based on the sample size, we estimated that for a variant with 20% prevalence, there was 81% power to detect a 50% change in risk. For a variant with 5% prevalence, there was 83% power to detect a halving or doubling of risk and less power to detect smaller risks.

All statistical analyses were conducted using SAS version 9.1 (SAS Institute Inc, Cary, NC).

### RESULTS

Three hundred forty-nine perimenopausal and postmenopausal women who experienced onset of incident VT between January 1, 1995, and December 31, 2002, and who provided a blood sample were included in analyses. From among the 531 potential case participants who were alive at the time of study recruitment, blood samples could not be obtained from 77 (15%) who were too ill or were not able to give informed consent and from 83 (16%) who refused study participation. Another 22 (4%) were excluded because blood could not be drawn or because genotyping on the sample failed or results were pending at the time of the analyses. Compared with included women, those excluded were on average older, nonwhite, and had more known risk factors for VT. Controls were 1680 perimenopausal and postmenopausal women without a history of incident VT. Characteristics of cases and controls are listed in **Table 1**. The study population was largely white, with a mean age of 66.6 years for cases and 68.0 years for controls. Traditional acquired risk factors, such as history of cancer, recent hospitalizations, and use of hormone therapy, were more common in cases than in controls.

#### Gene Associations

**Table 2** provides a summary of the 24 investigated candidate clotting-related genes and their global associa-

---

**Table 1. Characteristics of Study Participants**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Cases (n = 349)</th>
<th>Controls (n = 1680)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD), y</td>
<td>66.6 (10.6)</td>
<td>68.0 (9.4)</td>
</tr>
<tr>
<td>White race</td>
<td>331 (95)</td>
<td>1559 (93)</td>
</tr>
<tr>
<td>Body mass index, mean (SD)*</td>
<td>31.1 (7.8)</td>
<td>28.4 (6.4)</td>
</tr>
<tr>
<td>History of cancer within 5 y before index date</td>
<td>45 (13)</td>
<td>84 (5)</td>
</tr>
<tr>
<td>Hospitalization within 30 d before index date</td>
<td>58 (17)</td>
<td>7 (0.4)</td>
</tr>
<tr>
<td>Outpatient diagnostic or surgical procedure within 30 d before index date</td>
<td>17 (5)</td>
<td>5 (0.3)</td>
</tr>
<tr>
<td>Major fracture within 30 d before index date</td>
<td>5 (1)</td>
<td>1 (0.06)</td>
</tr>
<tr>
<td>Perimenopausal</td>
<td>12 (3)</td>
<td>47 (3)</td>
</tr>
<tr>
<td>Current estrogen use</td>
<td>157 (45)</td>
<td>667 (40)</td>
</tr>
</tbody>
</table>

*Calculated as weight in kilograms divided by height in meters squared.
tion with VT risk. The number of rare haplotypes was small; coverage of known haplotypes encompassed at least 90% of alleles in 19 genes and at least 80% in the remaining 5 genes. There were 3 genes (factor V upper half, protein C, and TFPI) for which the P value for the global test of haplotype associations was less than .05, but the only one with a q value less than .20 was TFPI (q = .13). When analyses were limited to the 331 cases and 1559 controls who were white, global P value and q value changes were minimal.

**Haplotype and SNP Associations**

Tables presenting the analyses conducted for each of the 24 genes are available from the authors at http://depts.washington.edu/chru. TABLE 3 is presented as an example. It provides a matrix of nucleotides that summarize the association between SNP variation and haplotypes within the factor II gene. Haplotype- and SNP-specific ORs, 95% CIs, and P values are presented. For this gene, we found that 2 of the 7 haplotypes were associated with an increased risk of VT when compared with the reference haplotype. Each high-risk haplotype was marked by a single SNP that was also associated with an increased risk of VT: haplotype 6 was marked by SNP 3696, and haplotype 7 was uniquely marked by SNP 20210A.

A total of 173 haplotypes were tested across the 24 candidate genes after removing haplotypes that were marked by the factor V Leiden and factor II G20210A variants. Twenty (12%) of the 173 tests had P values less than .05, which represents 11 more positive find-
ings than the 9 that would have been expected by chance alone at the α level of .05. These haplotypes were found in 14 of the 24 genes (factors II, V, X, and XI; the 3 fibrinogen genes; tissue factor; proteins C and S; thrombomodulin; TFPI; PAI-1; and TAFI). A total of 170 SNPs were tested across 24 candidate genes after removing the factor V Leiden and factor II G20210A variants. Twenty-one (12%) of the 170 tests had P values less than .05, which represents 12 more positive findings than the 9 that would have been expected by chance alone. Associations were found in 13 of the 24 genes (factors II, V, IX, X, and XI; all 3 fibrinogen genes; protein C; thrombomodulin; TFPI; PAI-1; and TAFI).

Table 4 lists the 29 associations with P values less than .05 that were identified in the 24 candidate genes in haplotype and SNP analyses. The minor allele frequency of variants among controls ranged from 0.3% to 48.6%. The FDR q values for the 29 associations ranged from .09 to .35, with only 5 below the .20 threshold. The population-attributable risk percentage ranged from 0.3% to 12.2%. When analyses were limited to whites, q value changes were minimal.

Table 5 compares risk among participants with and without a known acquired VT risk factor. Few idiopathic VT events occurred (26% of all events), and risk estimates for idiopathic and nonidiopathic events were not statistically different from one another.

**COMMENT**

This study assessed the associations between VT risk in postmenopausal women and common genetic variation in 24 candidate coagulation, anticoagulation, fibrinolysis, and antifibrinolysis genes. We identified 1 gene (TFPI) with a global association with VT risk and 5 SNP associations with FDR q values less than .20, of which 3 (in factors V and XI and in protein C) have not been previously reported. None of the ORs approached the nearly 4-fold increase in risk demonstrated for the factor V Leiden variant. None of the haplotype associations had an FDR q value less than .20. Differential risk by acquired risk status was not detected.

**Novel Findings, Replications, and Discrepancies With Other Studies**

The factor V exon 13 substitution K868R, the factor XI G22771A intron 14 of the 24 genes (factors II, V, X, and XI; the 3 fibrinogen genes; tissue factor; proteins C and S; thrombomodulin; TFPI; PAI-1; and TAFI).

and TAFI).

Table 3. Analysis of Associations of Factor II Haplotypes and Single Nucleotide Polymorphisms (SNPs) With Incident Venous Thrombosis

<table>
<thead>
<tr>
<th>Haplotype Identifier</th>
<th>2890</th>
<th>3696</th>
<th>4992</th>
<th>5389</th>
<th>5467 (T165M)</th>
<th>7530</th>
<th>21239</th>
<th>rs1799963 (20210)</th>
<th>OR (95% CI)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>G</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>A</td>
<td>G</td>
<td>45.4</td>
<td>48.8</td>
</tr>
<tr>
<td>2</td>
<td>G</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>G†</td>
<td>G</td>
<td>14.6</td>
<td>14.3</td>
</tr>
<tr>
<td>3</td>
<td>G</td>
<td>C</td>
<td>C†</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>G†</td>
<td>C†</td>
<td>15.3</td>
<td>13.7</td>
</tr>
<tr>
<td>4</td>
<td>G</td>
<td>C</td>
<td>C†</td>
<td>G†</td>
<td>C</td>
<td>C</td>
<td>G†</td>
<td>C†</td>
<td>7.7</td>
<td>8.5</td>
</tr>
<tr>
<td>5</td>
<td>A†</td>
<td>C</td>
<td>T</td>
<td>G†</td>
<td>C</td>
<td>C</td>
<td>G†</td>
<td>C†</td>
<td>6.7</td>
<td>7.0</td>
</tr>
<tr>
<td>Other†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.7</td>
<td>3.8</td>
</tr>
<tr>
<td>6</td>
<td>G</td>
<td>T†</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>G‡</td>
<td>G</td>
<td>3.9</td>
<td>2.3</td>
</tr>
<tr>
<td>7</td>
<td>G</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>A</td>
<td>At</td>
<td>2.6</td>
<td>1.5</td>
</tr>
<tr>
<td>Control MAF %</td>
<td>8.2</td>
<td>2.3</td>
<td>23.2</td>
<td>33.1</td>
<td>14.6</td>
<td>14.5</td>
<td>49.0</td>
<td>1.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OR (95% CI)</td>
<td>0.93</td>
<td>1.82</td>
<td>2.02</td>
<td>3.01</td>
<td>1.11</td>
<td>1.09</td>
<td>1.13</td>
<td>1.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P Value</td>
<td>.67</td>
<td>.09</td>
<td>.84</td>
<td>.83</td>
<td>.36</td>
<td>.46</td>
<td>.16</td>
<td>.03</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; MAF, minor allele frequency; OR, odds ratio.

*Comparable tables of analyses for all haplotypes and SNPs are available at http://depts.washington.edu/chru.

†Minor allele.

§Global P = .20 after removal of the carriers of the 20210A allele.

|| (Risk for each additional copy of the variant allele compared with risk associated with an additional copy of the wild-type allele.)
GENETIC VARIATIONS AND NONFATAL VENOUS THROMBOSIS IN POSTMENOPAUSAL WOMEN

Table 4. Summary of Associations Between Single Nucleotide Polymorphisms (SNPs) and Haplotypes and Risk of Incident Venous Thrombosis

<table>
<thead>
<tr>
<th>Gene</th>
<th>Variant Identifier (Location)/RS No.*</th>
<th>Substitution</th>
<th>Control MAF, %</th>
<th>OR (95% CI)</th>
<th>P Value</th>
<th>FDR q Value</th>
<th>PAR %</th>
<th>Heterozygous for Variant</th>
<th>Homozygous for Variant</th>
<th>Novelty of Association</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor II</td>
<td>rs3136520 C→T</td>
<td></td>
<td>2.3</td>
<td>1.82 (1.16-2.86)</td>
<td>.009 .22</td>
<td>1.9</td>
<td>NP</td>
<td>NP</td>
<td>New</td>
<td></td>
</tr>
<tr>
<td>Factor V</td>
<td>rs3753305 C→G</td>
<td></td>
<td>42.5</td>
<td>1.20 (1.01-1.42)</td>
<td>.04 .32</td>
<td>7.8</td>
<td>1.31 (0.99-1.73)</td>
<td>1.41 (0.99-2.00)</td>
<td>New</td>
<td></td>
</tr>
<tr>
<td>Factor XI</td>
<td>rs3822057 C→A</td>
<td></td>
<td>1.5</td>
<td>1.88 (1.05-3.26)</td>
<td>.03 NA 1.3</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>Rpl</td>
<td></td>
</tr>
<tr>
<td>TFPI</td>
<td>rs2192824 G→T</td>
<td></td>
<td>2.3</td>
<td>3.75 (2.56-5.15)</td>
<td>&lt;.001 NA 5.9</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>Rpl</td>
<td></td>
</tr>
<tr>
<td>Protein S Haplotype</td>
<td>None</td>
<td></td>
<td>27.0</td>
<td>0.74 (0.60-0.90)</td>
<td>.003 .14</td>
<td>8.7</td>
<td>0.75 (0.58-0.97)</td>
<td>0.52 (0.30-0.91)</td>
<td>New</td>
<td></td>
</tr>
<tr>
<td>Protein C Haplotype</td>
<td>None</td>
<td></td>
<td>1.3</td>
<td>2.01 (1.12-3.63)</td>
<td>.02 .24</td>
<td>1.3</td>
<td>NP</td>
<td>NP</td>
<td>New</td>
<td></td>
</tr>
<tr>
<td>Factor IX</td>
<td>rs4197555 T→A</td>
<td></td>
<td>5.6</td>
<td>1.49 (1.08-2.05)</td>
<td>.02 .22</td>
<td>2.7</td>
<td>NP</td>
<td>NP</td>
<td>New</td>
<td></td>
</tr>
<tr>
<td>Factor X</td>
<td>rs693333 C→G</td>
<td></td>
<td>40.4</td>
<td>0.80 (0.68-0.96)</td>
<td>.01 .22</td>
<td>9.2</td>
<td>0.85 (0.66-1.10)</td>
<td>0.62 (0.44-0.91)</td>
<td>New</td>
<td></td>
</tr>
<tr>
<td>Factor XI</td>
<td>rs3822057 C→A</td>
<td></td>
<td>48.6</td>
<td>0.83 (0.70-0.99)</td>
<td>.04 .32</td>
<td>9.1</td>
<td>0.76 (0.58-0.99)</td>
<td>0.71 (0.51-0.99)</td>
<td>New</td>
<td></td>
</tr>
<tr>
<td>Fibrinogen α</td>
<td>rs2070006 G→A</td>
<td></td>
<td>38.3</td>
<td>1.25 (1.05-1.42)</td>
<td>.01 .22</td>
<td>8.7</td>
<td>1.18 (0.90-1.54)</td>
<td>1.60 (1.12-2.28)</td>
<td>Rpl</td>
<td></td>
</tr>
<tr>
<td>Fibrinogen β</td>
<td>rs2289252 C→T</td>
<td></td>
<td>39.5</td>
<td>1.31 (1.10-1.55)</td>
<td>.002 .13</td>
<td>10.9</td>
<td>1.09 (0.83-1.44)</td>
<td>1.79 (1.28-2.51)</td>
<td>New</td>
<td></td>
</tr>
<tr>
<td>Fibrinogen γ</td>
<td>rs2070006 G→A</td>
<td></td>
<td>42.5</td>
<td>1.20 (1.03-1.40)</td>
<td>.02 .25</td>
<td>5.9</td>
<td>1.22 (0.95-1.57)</td>
<td>1.55 (0.91-2.49)</td>
<td>Rpl</td>
<td></td>
</tr>
<tr>
<td>Tissue factor Haplotype</td>
<td>None</td>
<td></td>
<td>24.8</td>
<td>1.23 (1.02-1.49)</td>
<td>.03 .30</td>
<td>5.4</td>
<td>1.17 (0.91-1.50)</td>
<td>1.64 (1.04-2.58)</td>
<td>Rpl</td>
<td></td>
</tr>
<tr>
<td>Anticoagulation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein A</td>
<td>rs1999810 A→T</td>
<td></td>
<td>2.0</td>
<td>1.85 (1.12-3.03)</td>
<td>.08 .27</td>
<td>1.7</td>
<td>NP</td>
<td>NP</td>
<td>New</td>
<td></td>
</tr>
<tr>
<td>Factor V</td>
<td>rs1800788 C→T</td>
<td></td>
<td>21.3</td>
<td>1.31 (1.08-1.60)</td>
<td>.008 .21</td>
<td>6.2</td>
<td>1.27 (0.99-1.64)</td>
<td>1.83 (1.09-3.07)</td>
<td>New</td>
<td></td>
</tr>
<tr>
<td>Fibrinogen γ</td>
<td>rs2066854 A→T</td>
<td></td>
<td>24.8</td>
<td>1.23 (1.02-1.49)</td>
<td>.03 .30</td>
<td>5.4</td>
<td>1.17 (0.91-1.50)</td>
<td>1.64 (1.04-2.58)</td>
<td>Rpl</td>
<td></td>
</tr>
<tr>
<td>Antifibrinolysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAI-1</td>
<td>rs227631 A→G</td>
<td></td>
<td>42.7</td>
<td>1.19 (1.01-1.41)</td>
<td>.04 .32</td>
<td>7.5</td>
<td>1.17 (0.89-1.54)</td>
<td>1.43 (1.02-2.02)</td>
<td>Inc</td>
<td></td>
</tr>
<tr>
<td>PAI-1</td>
<td>rs227667 A→G</td>
<td></td>
<td>20.8</td>
<td>1.26 (1.03-1.55)</td>
<td>.02 .25</td>
<td>5.1</td>
<td>1.29 (0.99-1.65)</td>
<td>1.53 (0.87-2.70)</td>
<td>New</td>
<td></td>
</tr>
<tr>
<td>TAFI</td>
<td>rs17844078 T→C</td>
<td></td>
<td>11.5</td>
<td>1.38 (1.08-1.77)</td>
<td>.007 .27</td>
<td>4.1</td>
<td>1.31 (0.99-1.73)</td>
<td>2.51 (1.06-5.90)</td>
<td>New</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; E, exon; FDR, false discovery rate; I, intron; Inc, inconsistent associations in literature; MAF, minor allele frequency; NA, not applicable (false discovery rate); OR, odds ratio; P, promoter; PAI-1, plasminogen activator inhibitor 1; PAR %, population-attributable risk percentage; Rpl, replication of other reports in the literature; RS No., reference single nucleotide polymorphism accession identifier; TAFI, thrombin activatable fibrinolysis inhibitor; TFPI, tissue factor pathway inhibitor; U, untranslated region.

*For single nucleotide polymorphisms, the Seattle Program for Genomic Applications identifier is included when available in addition to an RS number.
†Estimates risk associated with each additional copy of minor allele compared with additional copy of major allele.
‡Estimates risk compared to those with no copies of the minor allele.

©2007 American Medical Association. All rights reserved.
between variants in the factor VII gene and VT risk were not confirmed in our data. Reports on a factor VIII non-synonymous substitution (reported as D1260E) have been inconsistent in the literature; we found a nonsignificant (P = .07) 21% increase in VT risk for this variant. A factor XII variant and factor XIIIa Val34Leu and factor XIIIb His93Arg have been associated with increased VT risk, but none of these associations were replicated in our study. Other variants in TFPI that have been reported in the literature were either not associated with risk in our data or were rare and not identified in the sequence panel. Variants in the TAFI promoter region associated with modest protection were not replicated.

**Table 5. Association Between Variants and Risk of Incident Venous Thrombosis Stratified by Known Risk Factors**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Variant Identifier</th>
<th>Without Known Risk Factor</th>
<th>With Known Risk Factor</th>
<th>P Value for Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor II</td>
<td>3696(I)/rs3136520</td>
<td>Coagulation</td>
<td>2.24 (1.28-3.94)</td>
<td>.42</td>
</tr>
<tr>
<td>Factor V</td>
<td>3578(I)/rs3753305</td>
<td></td>
<td>1.31 (1.06-1.61)</td>
<td>.14</td>
</tr>
<tr>
<td>PAI-1</td>
<td>664(P)/rs2227631</td>
<td></td>
<td>1.23 (1.00-1.51)</td>
<td>.81</td>
</tr>
<tr>
<td>TFPI</td>
<td>1502(P)/rs2192824</td>
<td></td>
<td>1.23 (1.00-1.51)</td>
<td>.81</td>
</tr>
</tbody>
</table>

**Interpretation of the FDR Statistic in Candidate Gene Studies**

We used a .20 threshold for the q value to guide the interpretation of gene-, haplotype-, and SNP-level findings in this candidate gene study. For moderate-sized association studies like this one, the global gene-level approach lacked power. This approach failed to detect potentially important variation within the gene. Our approach excluded from discovery analyses variants with well-established risk for VT (factor II 20210A and factor V Leiden) and included less–well-established potential genetic risks. Among the latter, 7 of the 9 were replicated in our data, yet only 2 of the 7 had a q value of less than .20. This finding suggests that a larger sample size would be appropriate for reliable replication of the most likely associations associated with variants in candidate genes. For example, in our data, factor II 20210A would have had a q value of approximately .30 if it had been included in our discovery work. Therefore, while the use of FDR is helpful in interpreting large numbers of tests, evidence for genetic effects can only be realized through multiple replication studies, with functional characterization of the variants.

**Limitations**

This comprehensive approach to risk evaluation among candidate genes at whole-gene, haplotype, and SNP levels inherently involved multiple tests, which can increase the risk of reporting false-positive associations. To better characterize this risk, we provided q values for FDR and a suggested q value threshold. This population association study is also prone to the potential confounding effects of population structure. We have adjusted for self-reported race in all analyses and restricted analyses to white individuals. The population of black individuals in our study was too small to address risk differences by race. The statistical power to detect significant associations with uncommon genetic variation was limited by sample size, and there was less than 81% power to detect changes in...
GENETIC VARIATIONS AND NONFATAL VENOUS THROMBOSIS IN POSTMENOPAusal WOMEN

risk of less than 50% among variants...drug therapies. JAMA. 1995;274:620-625.


