

## OBSTETRICS

# The risk of venous thromboembolism in oral contraceptive users: the role of genetic factors—a prospective cohort study of 240,000 women in the UK Biobank



Valeria Lo Faro, PhD; Therese Johansson, MSc; Åsa Johansson, PhD

**BACKGROUND:** More than 150 million women worldwide use oral contraceptives. Women with inherited thrombophilia and carriers of certain thrombophilia gene variants, such as factor V Leiden and the prothrombin, are at an increased risk for venous thromboembolism, especially when combined with oral contraceptive use. Venous thromboembolism is a complex disorder involving many genetic risk factors, and recently, polygenic risk scores have been proposed to capture a significant proportion of the genetic risk of venous thromboembolism.

**OBJECTIVE:** The aim of this study was to estimate the risk for developing venous thromboembolism when initiating oral contraceptive use (first 2 years) and during continued use among women with a high genetic liability.

**STUDY DESIGN:** We used a prospective study design in which 244,420 participants from the UK Biobank were followed from birth. The effect of oral contraceptive use during the first 2 years and in the remaining years of oral contraceptive use on the risk of developing venous thromboembolism was estimated using a Cox regression with a time-dependent exposure variable. Women were stratified according to their polygenic risk scores and whether they were carriers of factor V Leiden and/or prothrombin variants.

**RESULTS:** When genetic risk was not considered, an increased risk for venous thromboembolism was observed during the first 2 years of oral contraceptive use (hazard ratio, 3.09; 95% confidence interval, 3.00–3.20) but not during continued use (hazard ratio, 0.92; 95% confidence interval, 0.80–1.05). However, when genetic risk was considered, women in the highest polygenic risk score category had a more pro-

nounced risk of developing a venous thromboembolism during the first 2 years of oral contraceptive use (hazard ratio, 6.35; 95% confidence interval, 4.98–8.09), and a high risk was also observed among factor V Leiden (hazard ratio, 5.73; 95% confidence interval, 5.31–6.17) and prothrombin variant carriers (hazard ratio, 5.23; 95% confidence interval, 4.67–5.87). A high polygenic risk score in combination with being a factor V Leiden and prothrombin variant carrier conferred the highest risk for developing a venous thromboembolism during the first 2 years of oral contraceptive use (hazard ratio, 14.8; 95% confidence interval, 9.28–23.6). Women with a high genetic liability also had an increased risk during continued use but it was less pronounced, and the highest risk was conferred to carriers of both factor V Leiden and the prothrombin variant (hazard ratio, 4.93; 95% confidence interval, 3.16–7.7).

**CONCLUSION:** Evaluating polygenic risk can identify additional venous thromboembolism risk that is not captured in the commonly investigated genes for inherited thrombophilia. Our results indicate that oral contraceptive use is associated with an increased risk for developing a venous thromboembolism, particularly among women with a high genetic predisposition, and that oral contraceptive use dramatically increases the risk thereof short after initiation of use, which decreases with continued use. This suggests that the polygenic risk score could be used to identify women who are at high risk for developing a venous thromboembolism and advise them on alternative methods of contraception.

**Key words:** factor V Leiden, oral contraceptives, polygenic score, prothrombin G20210A, risk assessment, venous thromboembolism

## Introduction

Oral contraceptives (OCs) enable women to control their fertility.<sup>1</sup> However, studies have reported an increased

risk for thrombotic events among OC users.<sup>2</sup> The excess estrogenicity of OCs (the sum of estrogen and progestin contributions) increases the risk for a venous thromboembolism (VTE).<sup>3,4</sup> VTEs are a leading cause of cardiovascular death worldwide.<sup>5</sup> Each year in Europe, it is estimated that approximately 22,000 VTE events are related to OC use.<sup>6</sup> A VTE is a complex disorder that is influenced by both acquired and inherited factors. The acquired factors include, among others, the use of OCs.<sup>7</sup> The inherited factors include factor V Leiden (FVL) and prothrombin factor II (PTM) variants. From twin studies, VTE heritability has been estimated to be

50%.<sup>8</sup> However, today, known genetic variants can explain only 6% of the heritability.<sup>9</sup> Among women who use OCs, the risk for VTE is 3 to 5 times higher than for women who have never used OCs, and the highest risk is during the first 2 years of use.<sup>10</sup> Furthermore, the alteration in hemostatic imbalance and consequently the increased risk for a VTE when compared with the general population is more pronounced among women with a monogenic hereditary thrombophilia condition.<sup>11</sup> The World Health Organization (WHO) states that the use of OCs by these women is associated with an unacceptable health risk. However, VTE is a polygenic disorder

**Cite this article as:** Lo Faro V, Johansson T, Johansson A. The risk of venous thromboembolism in oral contraceptive users: the role of genetic factors—a prospective cohort study of 240,000 women in the UK Biobank. *Am J Obstet Gynecol* 2024;230:360.e1-13.

0002-9378

© 2023 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).  
<https://doi.org/10.1016/j.ajog.2023.09.012>



Click [Supplemental Materials](#) and [Video](#) under article title in Contents at [ajog.org](#)

## AJOG at a Glance

**Why was this study conducted?**

Oral contraceptive (OC) use increases the risk for venous thromboembolism (VTE) by a factor of 3 to 5. Factor V Leiden (FVL) and prothrombin G20210A (PTM) variants are known genetic risk factors for VTE. It is known that VTE is a polygenic disease and studies assessing the polygenic risk among OC users are lacking.

**Key findings**

Women with the highest polygenic risk have more than a 6-fold increased VTE risk during the first 2 years of OC use, which is higher than the risk among FVL or PTM carriers. With continued OC use, the increased risk is less pronounced.

**What does this add to what is known?**

Our study highlights the need to consider the polygenic effects of VTE in addition to the well-known hereditary thrombophilia variants when women initiate OC use.

and genetic liability for VTE can also be assessed using polygenic risk scores (PRSs).<sup>12</sup>

Currently, risk assessment in contraceptive counselling is based on clinical characteristics and a family history of VTE; the latter has shown poor sensitivity and predictive performances.<sup>13–15</sup> The main aim of this study was to estimate the risk for VTE associated with initiating OC use and with continued use among women with a high genetic liability, using both the PRS and the well-known genetic risk factors FVL and PTM. We also evaluated the performance of the PRS to accurately identify women with a high risk of developing VTE.

**Materials and Methods****Study cohort**

The UK Biobank (UKB) is a population-based cohort study in which more than 500,000 people, aged 37 to 72 years, were recruited from the general population through 22 assessment centers in the United Kingdom between 2006 and 2010. Participants were followed prospectively in different national registers.<sup>16</sup> Baseline information was collected at recruitment using touch-screen and nurse-administered questionnaires and through physical examinations. Biologic samples were also collected and almost all participants have been genotyped ([Supplementary Material](#)).

**Study design**

We investigated the OC-associated VTE risk among female participants in the UKB cohort. Our study was designed as a prospective study in which VTE was a binary outcome and the rate of VTE was assessed in all women in relation to their exposure to OCs. Women with missing information on their OC use, those who had no data of the covariates used, those who were not genotyped, or those who were not White European were excluded from the analyses ([Supplemental Figure](#)), giving a sample of 244,420 women in the analyses. Our study was designed to follow the women from birth (age, 0 years) until the first of the following events occurred: VTE diagnosis, end-of-study follow-up (ie, age at recruitment), underwent a bilateral oophorectomy or hysterectomy, or menopause started. To examine whether the use of OCs in combination with a high genetic liability for VTE confers an increased risk, we stratified the cohort into 10 deciles of risk according to the PRS scores (using the first decile as the reference) and/or according to their FVL and PTM carrier status (using the noncarriers as the reference). This study was approved by the UKB (application #41143) and the Swedish Ethical Review Authority (dnr: 2020-04415).

**Assessment of exposure, outcome, and covariates**

Information on OC use, including the age at initiation and discontinuation, was assessed during the initial assessment visit. The relevant UKB data fields referenced included 2784 (ever taken OC pill), 2794 (age started OC pill), and 2804 (age when an OC pill was last used). The first occurrence of VTE in the UKB was based on medical history and linkage to hospital admission data and a cause of death register. The UKB data fields included the following International Classification of Diseases, Ninth Revision (ICD9) and ICD10 codes: 4151, 4511, 4532, I80, I81, I82, and I26, which were extracted from health records, and self-reported VTE occurrence extracted from field codes 20002 (1068, 1093, and 1094). The [Supplemental Material](#) contains information on the assessment of covariates.

**Genotyping and polygenic risk scores**

The UKB participants had been genotyped using the UKB Axiom array (ThermoFisher Scientific, Santa Clara, CA) and the UK BiLEVE array, and untyped variants were imputed using SHAPEIT.<sup>16</sup> From the genetic data, we extracted information for rs6025 (FVL) (effect allele T; allele frequency, 0.02) in the *F5* gene and rs1799963 (PTM) (effect allele A; allele frequency, 0.01) in the *F2* gene ([Supplemental Material](#)). The PRS for VTE used in this study had been calculated already by Genomics PLC under the UK Biobank project 9659 and was provided by the UKB (UKB data field 26289 - Standard PRS for VTE).<sup>17</sup> This particular PRS had been trained on the Electronic Medical Records and Genomics (eMERGE) cohort (releases 2, 3, 5, and 6), which is a consortium of 10 participating sites that jointly performed genome-wide association studies (GWAS) and made the respective summary statistics freely available.<sup>18</sup> These cohorts did not overlap with that of the UKB in terms of participants. The PRS was constructed from the VTE GWAS summary statistics, including 29,799 VTE cases and 475,303 controls from the

eMERGE cohorts. All genetic variants with an imputation quality score  $>0.8$  were used to generate the PRS weights. In addition, any genetic variants that showed large differences in allele frequency between the UKB genetically inferred ancestry groups and either the GnomAD or the 1000 Genomes Project and those with evidence of large departures from the Hardy-Weinberg equilibrium ( $P$  value  $>1e-10$ ) were excluded. The PRS algorithm was constructed using a Bayesian approach, which has been described previously.<sup>17</sup> The PRS has already been validated in the UKB in a recent study in which it was applied to determine the risk for deep vein thrombosis (a manifestation of a venous thromboembolism).<sup>19</sup> Its performance was then compared with 2 other PRSs, 1 that was trained using half of the UKB and 1 that used the Global Biobank Meta-analysis Initiative consortium effort. It was observed that the area under the curve (AUC) estimates improved from 0.60 (95% confidence interval [CI], 0.59–0.61) when conventional risk factors were considered (sex, age, and principal components) to 0.66 (95% CI, 0.65–0.67) when the eMERGE PRS was also included. In addition, we also validated the PRS in terms of VTE as part of the current study. For calculating the AUC and its 95% CIs, we used the R package pROC (R Core Team, Vienna, Austria).<sup>20</sup>

### Cox regression

Cox regression analyses were performed to calculate the instantaneous VTE risk during the period of OC use. We only considered first events of VTE in our study because women were censored after the first VTE diagnosis. The follow-up started at birth and age was used as the primary timescale. Women were followed until the first of the following events occurred: VTE diagnosis, end-of-study follow-up (ie, age at assessment center visit), when women underwent a bilateral oophorectomy or a hysterectomy, or when they entered menopause, whichever came first. To adjust for potential confounding, we included the following covariates: year of birth, body

mass index (BMI) at recruitment, pregnancy period, Townsend deprivation index (TDI) as a proxy for socioeconomic status, smoking status, and the first 4 genetic principal components (Supplemental Material). The genetic principal components were included as covariates to adjust for confounding owing to population stratification and were computed based on the genetic kinship between the individuals of the cohort.<sup>16,21</sup> The use of OCs was modelled as a time-varying variable to which all women were unexposed at age = 0 but the exposure status changed to exposed = 1 when women initiated OC use (Supplemental Material). The value of the exposure could also change from “first two years of use” to “remaining years of use” for women who continued their use for more than 2 years. This means that in the analyses, the incidence rate of VTE during the first 2 years after initiating OC use was compared with the incidence rate among women of the same age who have not yet used OCs. Similarly, the rate during the remaining years of use and up to 2 years after cessation was compared with that of women of the same age who have not yet used OCs. When estimating the effect during use, women were censored 2 years after stopping OC use. The reason for considering the period up to 2 years after discontinuation as continued use is because there is a risk for some women who developed a VTE just before they stopped OC use, but it looks like the events occurred after they stopped OC use, which will introduce protopathic bias, also referred to as reverse causality. Therefore, we included a 2-year lag time that has been discussed previously.<sup>22</sup>

A second exposure variable was used to stratify women into high or low genetic VTE risk groups (Supplemental Material). For the analyses that include the PRS, we defined the reference group as women with the lowest genetic liability for VTE, that is, being in the first PRS decile and not being carriers of FVL and PTM, as comparison for the higher genetic VTE risk groups. In the analyses of the risk conferred by the presence of FVL and PTM, we used all noncarriers as

the reference group, irrespective of their PRS status, for comparison with the FVL and PTM carriers. Cox regression modelling was performed using the “survival” R package (R Core Team) to calculate the hazard ratio (HR) and its CIs.<sup>23,24</sup>

### Results

A total of 244,420 women were included in our analyses (Table) of which 10,856 experienced a first ever VTE event during the follow-up period. A total of 193,371 participants initiated OC use at some time point during the follow-up. The never-user group had a larger number of women who reported a VTE episode. This was most probably because women who never used OCs were older at the time of recruitment and therefore were more likely to have been diagnosed with a VTE than women who were younger at the time of recruitment. Among OC users, 8682 were carriers of FVL and 4119 were carriers of PTM. The frequency of FVL was 4.48% among OC users and 4.51% among never users, and for PTM, the frequency was 2.13% among users and 2.20% among never users. There was no significant difference in the PRS between ever and never users of OCs, indicating that bias caused by confounding by indication was unlikely to affect our results. The descriptive statistics for the different genetic risk groups and time of OC use analyzed in this study are summarized in Supplemental Table 1.

### Genetic predisposition and risk for venous thromboembolism

We categorized the women into different risk groups, namely those with the highest PRS for VTE (referred to as the 10<sup>th</sup> decile), those who were carriers of FVL, and/or those who were carriers of PTM. A total of 24,291 women were in the highest PRS category, whereas (independently of the PRS category) 10,985 women were carriers of FVL (either homozygous or heterozygous), and 5244 were carriers of PTM (Supplemental Table 2). We estimated the VTE risk in the entire cohort and in the subgroup of never users (Supplemental

**TABLE**  
**Characteristics of the entire study cohort**

Characteristics	Users of oral contraceptive	Never users of oral contraceptive	P value <sup>a</sup>
Number (%)	193,371	51,049	—
Venous thromboembolism events, n (%)	7687	3169	<2.2e-16
Year of birth, median (Q1–Q3)	1952 (1946–1959)	1945 (1942–1951)	<2.2e-16
Body mass index, median (Q1–Q3)	25.9 (23.3–29.4)	26.5 (23.6–30.05)	<2.2e-16
Age at presentation, median (Q1–Q3)	56 (49–62)	63 (57–66)	<2.2e-16
Age at first venous thromboembolism episode, median (Q1–Q3)	51 (33–64)	57 (35–69)	<2.2e-16
Age at oral contraceptive initiation, median (Q1–Q3)	21 (18–24)	—	—
Age at oral contraceptive discontinuation, median (Q1–Q3)	30 (26–37)	—	—
Duration of oral contraceptive use, median (Q1–Q3)	9 (4–15)	—	—
First 2 years of use, mean (full range)	1.8 (1–2)	—	—
Remaining years of use, median (Q1–Q3)	3 (2–7)	—	—
Age at menopause, median (Q1–Q3)	50 (45–52)	50 (45–53)	1.968e-13
Age at first delivery	25 (22–29)	24 (22–27)	<2.2e-16
Had hysterectomy, n (%)	34,273	11,373	<2.2e-16
Had bilateral oophorectomy, n (%)	14,856	4893	<2.2e-16
Townsend deprivation index, median (Q1–Q3)	-2.30 (-3.7 to 0.12)	-2.17 (-3.62 to 0.46)	<2.2e-16
FVL carriers, n (%)	8682	2303	8.345e-1
PTM carriers, n (%)	4119	1125	3.059e-1
Delivery, n (%)	133,150	32,931	<2.2e-16
Smoking, n (%)	55,479	11,773	<2.2e-16
Polygenic Risk score	-0.03 (-0.66 to 0.63)	-0.02 (-0.65 to 0.64)	3.2e-1

The data are presented as median (Q1, first quartile; Q3, third quartile) for continuous data and as total number and percentage for binary data.

FVL, factor V Leiden; PTM, prothrombin factor II.

<sup>a</sup> The Mann-Whitney U test was used to evaluate quantitative traits and the Pearson chi-square test was used for binary traits without considering any potential confounding.

Lo Faro. Genetic factors and risk of venous thromboembolism in women using oral contraceptives. *Am J Obstet Gynecol* 2024.

Table 3). All high genetic risk groups were associated with a significantly higher incidence rate of VTE when compared with the respective reference groups.

We also validated the PRS in the UKB. We estimated the AUC for a base model (including age and genetic principal components), a base plus the FVL and/or PTM model, and a base plus the PRS model. We observed that the base plus the PRS model improved classification when compared with the base and the base plus FVL and PTM models. The carrier status for the 2 variants increased the prediction by around 1.0% in the AUC, whereas the PRS increased the prediction in the AUC by 3.5% (Figure 1,

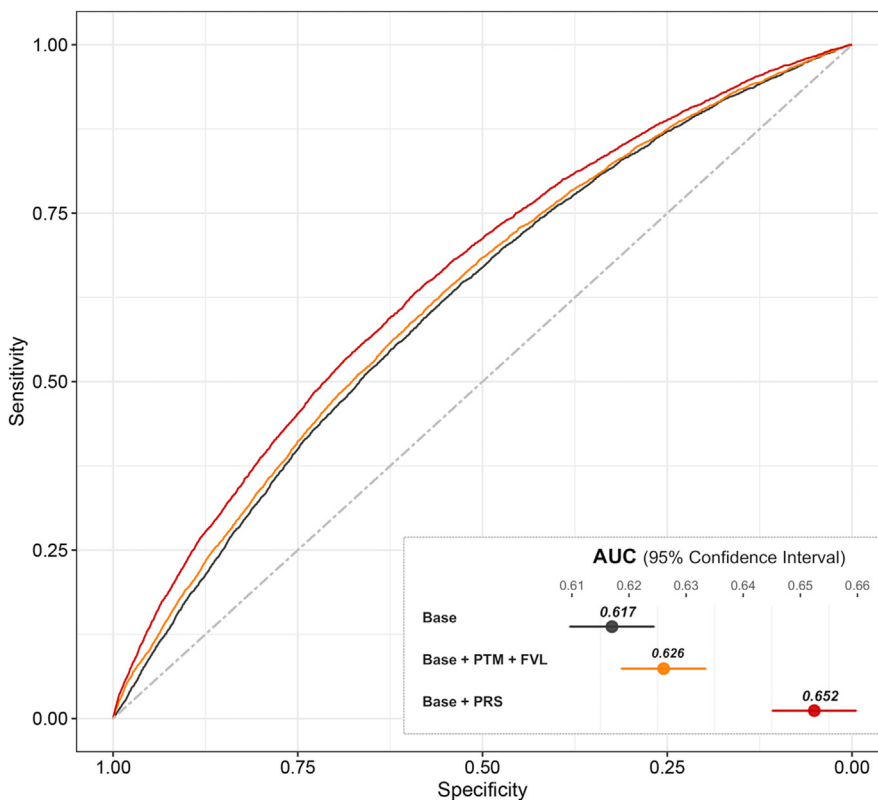
Supplemental Table 4). Subsequently, we estimated the odds ratio and 95% CI using a logistic regression for the PRS decile in our cohort (Figure 2). Here, we observed a trend of higher odds of VTE among those in higher PRS deciles.

#### Effect of duration of oral contraceptive use on venous thromboembolism risk

The association between the duration of OC use and VTE risk was first estimated for all women without considering their genetic predisposition for VTE. The risk during the first 2 years of OC use was associated with an increased hazard of developing a VTE (HR, 3.09; 95% CI,

3.00–3.20). During the remaining years of use, we found no association between OC use and VTE (HR, 0.92; 95% CI, 0.80–1.05). We also stratified the women according to their PRS (independent of being FVL or PTM carriers) into 1<sup>st</sup> and 10<sup>th</sup> PRS deciles and estimated the interaction effect for the first 2 years of OC use in comparison with never OC users. The effect during the first 2 years of use was significantly higher (interaction  $P < .001$ ) among women in the 1<sup>st</sup> decile than among those in the 10<sup>th</sup> PRS decile. In contrast, in the 2 strata containing FVL and PTM carriers, the first 2 years of OC use were associated with a similar risk as that among women in the 1<sup>st</sup> PRS decile.

**FIGURE 1**  
Discriminatory ability of VTE polygenic risk scores among female participants



Receiver operating characteristic curves assess the discriminative power of different significant models. The grey dotted line with an AUC of 50% is used as reference.

AUC, area under the curve; FVL, factor V Leiden; PRS, polygenic risk score; PTM, prothrombin factor II; UKB, UK Biobank; VTE, venous thromboembolism.

Lo Faro. Genetic factors and risk of venous thromboembolism in women using oral contraceptives. *Am J Obstet Gynecol* 2024.

### The combined effect of oral contraceptive use and genetic risk on venous thromboembolism

Among the women in the 10<sup>th</sup> PRS decile with a genetic risk factor, the HR during the first 2 years of OC use was associated with an increase in the VTE risk (HR, 6.35; 95% CI, 4.98–8.09) when compared with the reference category (Figure 3). The effect remained significant but with a less pronounced risk during the remaining years of OC use (HR, 2.12; 95% CI, 1.81–2.49). Among women in the 10<sup>th</sup> PRS decile as a genetic risk factor (neither FVL nor PTM carriers), the HR during the first 2 years of OC use was associated with an increased VTE risk (HR, 5.82; 95% CI, 4.48–7.91) (Figure 3). The effect remained significant but with a less pronounced risk

during the remaining years of OC use (HR, 1.94; 95% CI, 1.06–2.35). We also estimated the HR for those women in the 10<sup>th</sup> decile PRS who were carriers of FVL and/or PTM in comparison with the reference category. The first 2 years of OC use showed an increased HR (HR, 8.78; 95% CI, 6.12–12.6) for those in the 10<sup>th</sup> decile PRS and who were carriers of FVL, whereas the HR was 10.58 (95% CI, 7.48–14.97) for those in the 10<sup>th</sup> decile PRS and who were carriers of PTM. During the remaining years of use, the HR was 3.12 (95% CI, 2.5–3.88) and 3.64 (95% CI, 2.58–5.13) for those in the 10<sup>th</sup> decile of PRS and who were carriers of FVL and PTM, respectively (Figure 3). Among women with FVL as a genetic risk factor, the HR was increased during the first 2 years of OC use (HR,

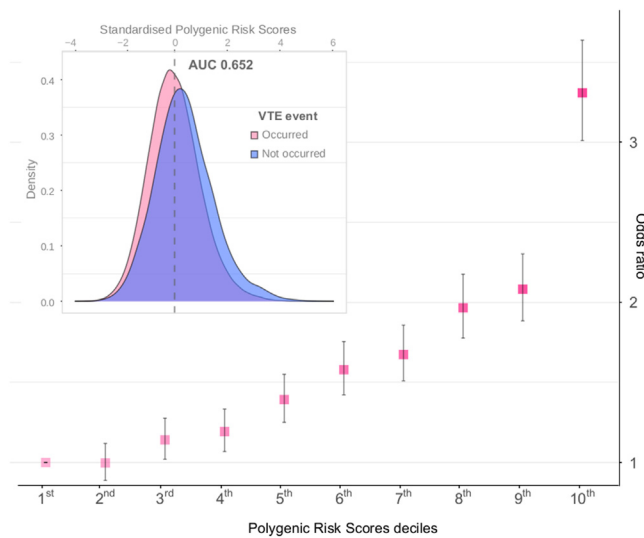
5.73; 95% CI, 5.31–6.17) regardless of which PRS category they belonged to. The HR remained significantly increased during the remaining years of OC use, but the risk was less pronounced (HR, 2; 95% CI, 1.86–2.16). Similarly, women with PTM as a genetic risk factor had an increased HR during the first 2 years of OC use (HR, 5.23; 95% CI, 4.67–5.87) and also during the remaining years of use (HR, 1.76; 95% CI, 1.57–1.97). The HR for women who carried both FVL and PTM was 9 (95% CI, 6.07–13.34) in the first 2 years, and 3.39 (95% CI, 2.38–4.83) in the remaining years.

We estimated the risk among women in the 10<sup>th</sup> decile PRS and who were both FVL and PTM carriers; in the first 2 years of OC use, the HR was 14.8 (95% CI, 9.28–23.6). A less pronounced increase in the hazard for VTE was observed during the remaining years of use (HR, 4.93; 95% CI, 3.16–7.7). However, it should be highlighted that the total number of VTE events was small (n=13 in the first 2 years of OC use; n=14 in the remaining years of use), and CIs for these estimates were wide.

### Comment Principal findings

We estimated the risk for VTE among women who used OCs in relation to their genetic predisposition. We showed that women in the highest PRS decile had a more than 6-fold increased risk for VTE during the first 2 years of OC use when compared with never users in the lowest genetic risk group. This increased risk was higher than the risk associated with being an FVL or a PTM carrier. Our results highlight that polygenic risk has a higher impact on the occurrence of VTE in the context of OC use.<sup>11,25–27</sup> Therefore, combining genetic liability (including several common low-effect variants) with clinical risk factors may allow better VTE risk stratification associated with OC use than when only considering FVL and PTM carrier status. Our study also showed that there is a discernible difference in the magnitude of the effect of OC use between the first 2 years of OC use and the remaining years of OC use (considered until 2 years after discontinuation), with a multiple-fold

**FIGURE 2**  
Odds ratio estimates for each VTE polygenic risk score decile



The first decile was used as reference for the others. The odds ratio and 95% confidence intervals were estimated using logistic regressions. Each point indicates the odds ratio and the error bar indicates the lower and upper 95% confidence interval values for each odds ratio. In the upper part of the figure, the density distribution plot of VTE cases vs controls is shown.

AUC, area under the curve; VTE, venous thromboembolism.

Lo Faro. Genetic factors and risk of venous thromboembolism in women using oral contraceptives. *Am J Obstet Gynecol* 2024.

increased risk associated with the first 2 years of use. This highlights the importance of treating OC use as a time-varying exposure variable instead of estimating an average HR for the duration of the follow-up for users vs never users.

### Results in the context of what is known

OC use has been shown previously to be associated with a 3 to 5 times higher risk for VTE with the highest risk observed during the first 2 years of use.<sup>28,29</sup> Consistent with this, in our study, the HR for VTE among women during the first 2 years of OC use was 3.09 when compared with never users and when genetic information was considered. The incidence of VTE among premenopausal women is about 3 per 10,000 women per year.<sup>30</sup> However, given that more than 150 million women worldwide use OCs, even a small increase in the risk for VTE associated with OC use leads to a substantial increase in the number of VTE cases. From the reported estimates of the different AUC

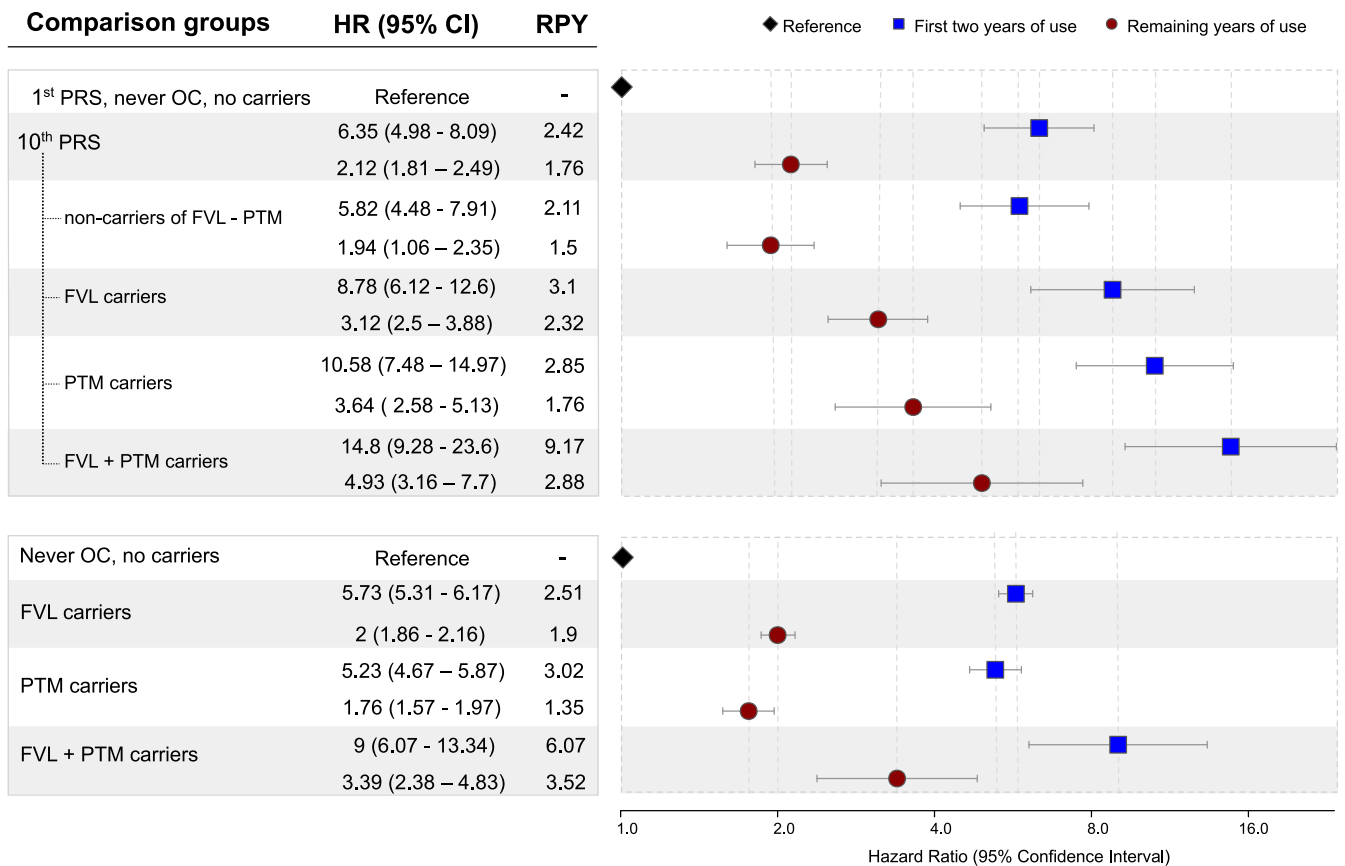
models, we observed that the PRS improved risk classification with an increased prediction rate of 3.5%. In the entire cohort of women in this study, a high PRS was associated with a higher rate of VTE (HR, 2.2), which was slightly higher than the risk of being a carrier of FVL or PTM (1.83 and 1.45, respectively), whereas the VTE risk reported in the current literature ranges from 3 to 5. The contribution of the PRS to the risk seems to be modest, however, the AUC calculated for different models, including the base model, base plus FVL and/or PTM model, and base plus the PRS model, showed that the PRS improved classification when compared with the base model and the model that included FVL and PTM status. The carrier status for the 2 variants increased prediction by around 1.0% in the AUC model in comparison with an increase of 3.5% in the AUC for the model that included the PRS. A study conducted on the VTE risk in a general population (independent of OC use and also including males) that used a different PRS than this study reported an

improvement of 4% in risk prediction, which is in line with our AUC estimate.<sup>9</sup> However, the PRS does not capture all genetic risk. In fact, the PRS additively incorporates known risk-associated loci but does not consider interactions (ie, gene-gene or gene-environment) or variants that act only in specific genetic backgrounds (ie, epistasis). In addition, other sources of variability that make the PRS unable to capture all genetic risk may be derived from differences in the allele frequencies of the common causative alleles and changes in environmental exposures. In addition, the performance of the PRS is highly dependent on the GWAS summary statistics used for constructing the PRS. The cohort used for our PRS was the largest available data set. However, it is anticipated that increasing the sample size of GWASs will lead to the identification of more VTE-associated genetic variants, boosting the statistical power, robustness, and clinical use of PRSs.

### Clinical implications

The WHO classified OCs as being an unacceptable health risk for women with known thrombogenic variants, but it discourages global screening for thrombophilia before prescribing OCs because of the low prevalence of thrombophilia (7%–8% among Europeans) and the high cost of screening.<sup>31</sup> VTE is a polygenic disease involving thousands of genetic variants that collectively contribute to the risk of a thrombotic event. Because the PRS takes into account a greater proportion of the genetic contribution to the disease, using the full spectrum of genetic risks for thrombotic events may improve risk stratification and numerically identify more individuals at higher risk among OC users than when using only FVL and PTM carrier status. A recent study has also shown that considering the polygenic background improved the risk accuracy estimates among individuals who carry a monogenic risk variant, which may better inform decision-making and refine risk estimates during counselling.<sup>32</sup> For women with a high genetic risk for VTE, this means that using OC dramatically increases their already high risk for VTE,

**FIGURE 3**  
**Venous thromboembolism risk according to genetics and oral contraceptive use**



The HRs for a venous thromboembolism were calculated for both the first 2 years of OC use (shown as blue squares) in the upper part of each group and for the remaining years of OC use (shown as red circles) in the lower part of each comparison group. The HRs for a venous thromboembolism and their error bars, which indicate 95% CIs, are shown in the HR (95% CI) column. The rate per 1000 person-years (RPY) column shows the rate of venous thromboembolism per 1000 person-years that occurred in each group and by the time of occurrence. For the analyses with the PRS, the reference category included individuals in the lowest decile of the PRS (1<sup>st</sup> PRS) who were not carriers of either FVL or the PTM and who had never used OCs. In the analysis of the risk of FVL and/or PTM, we used nonuser and noncarriers as the control category irrespective of any PRS status. All models were adjusted for body mass index, year of birth, pregnancy and postpartum periods, smoking status, Townsend deprivation index, and the first 4 principal components.

CI, confidence interval; FVL, factor V Leiden; HR, hazard ratios; OC, oral contraceptive; PRS, polygenic risk score; PTM, prothrombin factor II variant; RPY, rate per 1,000 person-years.

Lo Faro. Genetic factors and risk of venous thromboembolism in women using oral contraceptives. *Am J Obstet Gynecol* 2024.

especially in the first 2 years of use. Therefore, women in this situation need careful counselling about their contraceptive methods, including evaluation of other more appropriate contraceptive choices. These implications of using the PRS could better inform and improve the decision-making process and therefore refine the risk estimates during counselling. The presence of other risk factors for thrombosis, such as dyslipidemia, smoking, and obesity, should be considered when advising these women about oral contraceptive therapy.

Because genotyping can nowadays be done at a low cost, efforts to integrate genotyping data in healthcare systems may facilitate the use of the PRS as a tool to improve the identification of high-risk women among OC users. A preliminary health economic analysis study has shown the potential cost benefits of using PRS in cardiovascular disease prevention in the Finnish health system.<sup>33</sup> Indeed, it was found that the cardiovascular disease PRS, together with traditional risk factors, would have cost benefits if employed using a targeted

approach. However, evaluation studies in clinical settings among women who use OCs in relation to their genetic risk of VTE should be carried out to ascertain if there is a cost benefit. We also believe that for a proper individual risk evaluation, the future risk assessments should be performed in a clinical setting where the incorporation and use of PRS will be assessed in high-risk groups, such as smokers, individuals with obesity, and those with cardiovascular disease. Evaluations in different populations will also be required. It is important to note that

the characteristics of PRSs enable opportunities for earlier prevention. In general, for cardiovascular disease, risk factors are not measured early in life. In contrast, individuals can be genotyped early in life and have their PRS determined for a wide range of diseases. For those with a significantly increased lifetime risk for disease, targeted interventions could be used to reduce their risk, for example, through guidance on the use of drug therapy. An important consideration for the applicability of the PRS is its associated cost. In 2014, a review suggested that the cost of hospitalization for VTE ranged from about \$3000 to about \$8700. Nowadays, the cost of genotyping SNPs across the genome, including for both FVL and PTM, to determine a PRS for any given disease has decreased substantially. Based on current prices for genotyping arrays and the required bioinformatic analysis, a recent study estimated that the 1-time PRS cost is between \$80 and \$120.<sup>34</sup> Therefore, it is possible that there would be a significant health and economic benefit by performing genotyping for genetic risk predictions over time. However, further studies are needed to address issues related to the effectiveness and ethics of PRS-based screenings before they can be implemented in clinical settings.<sup>35</sup>

So far, most studies have evaluated the combined effect of OC use and FVL and PTM carrier status. Therefore, information on the combined effect of OC use and common genetic variants is still scarce.<sup>25,26</sup> In this study, we examined the risk in 244,420 participants, which allowed us to obtain more precise estimates, and analyzed OC use as a time-varying exposure. We found that the risk for VTE was not constant over time among OC users. This study contributes to a more accurate estimation of the risk associated with OC use among women with defects in 2 thrombophilia genes. From our study, the presence of the highest polygenic risk, the FVL variant, and the PTM variant among women who use OC seems to have an additive effect. Because we found only 27 OC users with a VTE event in the 10th PRS decile and who carried both mutations,

the confidence interval is large and the estimated risk should be considered an approximation (Figure 3, “10<sup>th</sup>+PTM+FVL”). In this study, there were very few or no homozygotes for FVL or PTM. A limited number of studies in the literature have described patients who are homozygous carriers of FVL and, usually, these patients develop their first thrombosis at a younger age with a 10- to 80-fold increased VTE risk when compared with controls.<sup>36–38</sup> Individuals who are homozygous for PTM are the rarest with only 141 homozygous PTM cases, mostly Southern European, found until 2022.<sup>39,40</sup>

### Research implications

We emphasize that not only individuals with FVL and PTM but also those with a high polygenic liability for VTE are at high risk and should therefore consider alternatives to OC use. Evaluations of the benefit of introducing global screening

for thrombophilia among those who wish to start OC therapy are based only on economic analyses. These analyses do not examine the duration of therapy and the benefits of knowing one's genetic risk. For women who are FVL carriers, the pharmacoeconomic evaluation by Smith et al<sup>41</sup> (that included 15 years of OC use) found that screening and counselling was an economically favorable strategy. The next step in the pharmacoeconomic evaluation will be to assess the impact of introducing global screening before starting OC.<sup>42</sup>

### Strengths

We used a combination of genetic and clinical data from a large group of women with a long duration of OC use and follow-up. Our estimates may explain why some women, even non-carriers of FVL and PTM variants, are at higher risk for developing VTE when using OCs. We showed that OC use in

### GLOSSARY

*Area under the curve (AUC):* scalar value measuring the overall performance of a binary classifier with values ranging from 0.5 to 1.0. The minimum value represents the performance of a random classifier, and the maximum value corresponds to a perfect classifier.

*Body mass index (BMI):* a person's weight in kilograms divided by the square of their height in meters.

*Cox regression:* model when the outcome is the length of time until an event occurs. It calculates the hazard of an event, which is defined as the conditional probability of a single non-repeatable event occurring in a given time interval, assuming that the person has not experienced the event before that time.

*Factor V Leiden (FVL):* abnormal factor V protein resulting from a point mutation in the factor V gene. The result of this mutation is a protein that is relatively resistant to degradation and, in turn, an increase in thrombin generation.

*Genetic principal components:* covariance pattern among individuals used as a covariate to reduce the effect of confounding in exposure and outcome.

*Genome-wide association study (GWAS):* a method used to analyze common genetic variants that are studied for association with a trait of interest by comparing the frequency of variants between individuals with a trait or disease and those without.

*Hazard ratio (HR):* an estimate of the relative hazard rate, which is the incidence rate of an event among exposed in relation to unexposed individuals.

*Polygenic risk score (PRS):* estimate that represents the individual genetic liability for a trait of interest.

*Prothrombin G20210A (PTM):* abnormality in the promoter region of the prothrombin gene caused by a mutation that leads to excessive accumulation of prothrombin.

*Townsend deprivation index (TDI):* estimation of area-based social deprivation scores (considering unemployment, overcrowding, non-car ownership, and non-home ownership) based on national census data.

*Time-dependent covariate:* covariate changing state over time during an observed period.

the presence of a high genetic liability is a circumstantial VTE risk factor among women of reproductive age.

### Limitations

First, the UKB cohort consisted of healthier individuals when compared with the general United Kingdom population (ie, the individuals who volunteered to participate have healthier lifestyles, higher levels of education, and better health than the general United Kingdom population), and only White European women were analyzed, thus, the results of the combined effect of OC and genetic risk factors should be replicated in other larger populations and among people with different ancestries. Consequently, additional studies are warranted. Second, based on the birth years of the UKB participants and the year in which they initiated OC use, our results are mainly based on the second generation of combined OCs and on the oral route of administration. Third, exposure to OCs was assessed by self-reported questionnaires, which are likely to introduce recall bias. Fourth, further studies are needed to show if the predictive accuracy of already established clinical risk factors will be improved by the addition of PRSs.

### Conclusion

Common genetic variants may capture additional risks not accounted for by traditional clinical and genetic factors. OC use is associated with a dramatically increased risk for VTE among highly genetically predisposed women and not only among carriers of the known FVL and PTM variants, especially at the beginning of use. Further studies in other populations and ancestries are needed to confirm our findings. ■

### Acknowledgments

We acknowledge all of the participants and the staff involved in the United Kingdom Biobank for their valuable contribution.

### References

1. Morimont L, Haguët H, Dogné JM, Gaspard U, Douxfils J. Combined oral contraceptives and venous thromboembolism: review

and perspective to mitigate the risk. *Front Endocrinol* 2021;12:769187.

2. Treger S, Ackerman S, Kaplan V, Ghanem S, Nadir Y. Progestin type affects the increase of heparanase level and procoagulant activity mediated by the estrogen receptor. *Hum Reprod* 2021;36:61–9.

3. Tepper NK, Whiteman MK, Marchbanks PA, James AH, Curtis KM. Progestin-only contraception and thromboembolism: a systematic review. *Contraception* 2016;94:678–700.

4. Mantha S, Karp R, Raghavan V, Terrin N, Bauer KA, Zwicker JI. Assessing the risk of venous thromboembolic events in women taking progestin-only contraception: a meta-analysis. *BMJ* 2012;345:e4944.

5. Blann AD, Lip GYH. Venous thromboembolism. *BMJ* 2006;332:215–9.

6. Odland V, Milsom I, Persson I, Victor A. Can changes in sex hormone binding globulin predict the risk of venous thromboembolism with combined oral contraceptive pills? *Acta Obstet Gynecol Scand* 2002;81:482–90.

7. Gutmann CH. Venous thrombosis: risk factors, management and complications. Hauppauge, NY: Nova Science Publishers Incorporated; 2014.

8. 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;14:328–32.

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

10. Vandenbroucke JP, Rosing J, Bloemenkamp KW, et al. Oral contraceptives and the risk of venous thrombosis. *N Engl J Med* 2001;344:1527–35.

11. van Vlijmen EFW, Wiewel-Verschueren S, Monster TBM, Meijer K. Combined oral contraceptives, thrombophilia and the risk of venous thromboembolism: a systematic review and meta-analysis. *J Thromb Haemost* 2016;14:1393–403.

12. Khera AV, Chaffin M, Aragam KG, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet* 2018;50:1219–24.

13. Cosmi B, Legnani C, Bernardi F, Coccheri S, Palareti G. Role of family history in identifying women with thrombophilia and higher risk of venous thromboembolism during oral contraception. *Arch Intern Med* 2003;163:1105–9.

14. Aznar J, Mira Y, Vayá A, Ferrando F, Villa P. Is family history sufficient to identify women with risk of venous thromboembolism before commencing the contraceptive pill? *Clin Appl Thromb Hemost* 2002;8:139–41.

15. Bezemer ID, van der Meer FJM, Eikenboom JCJ, Rosendaal FR, Doggen CJM. The value of family history as a risk indicator for

venous thrombosis. *Arch Intern Med* 2009;169:610–5.

16. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018;562:203–9.

17. Thompson DJ, Wells D, Selzam S, et al. UK Biobank release and systematic evaluation of optimised polygenic risk scores for 53 diseases and quantitative traits. *medRxiv*. 2022. Available at: <https://www.medrxiv.org/content/10.1101/2022.06.16.22276246v2>. Accessed October 11, 2023.

18. McCarty CA, Chisholm RL, Chute CG, et al. The eMERGE Network: a consortium of biorepositories linked to electronic medical records data for conducting genomic studies. *BMC Med Genomics* 2011;4:13.

19. Lo Faro V, Johansson T, Höglund J, Hadizadeh F, Johansson Å. Polygenic risk scores and risk stratification in deep vein thrombosis. *Thromb Res* 2023;228:151–62.

20. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011;12:1–8.

21. Haworth S, Mitchell R, Corbin L, et al. Apparent latent structure within the UK Biobank sample has implications for epidemiological analysis. *Nat Commun* 2019;10:333.

22. Tamim H, Monfared AAT, LeLorier J. Application of lag-time into exposure definitions to control for protopathic bias. *Pharmacoepidemiol Drug Saf* 2007;16:250–8.

23. Thiébaud ACM, Bénichou J. Choice of time-scale in Cox's model analysis of epidemiologic cohort data: a simulation study. *Stat Med* 2004;23:3803–20.

24. Cologne J, Hsu WL, Abbott RD, et al. Proportional hazards regression in epidemiologic follow-up studies: an intuitive consideration of primary time scale. *Epidemiology* 2012;23:565–73.

25. Vandenbroucke JP, Koster T, Briët E, Reitsma PH, Bertina RM, Rosendaal FR. Increased risk of venous thrombosis in oral-contraceptive users who are carriers of factor V Leiden mutation. *Lancet* 1994;344:1453–7.

26. Martinelli I, Taioli E, Bucciarelli P, Akhavan S, Mannucci PM. Interaction between the G20210A mutation of the prothrombin gene and oral contraceptive use in deep vein thrombosis. *Arterioscler Thromb Vasc Biol* 1999;19:700–3.

27. Legnani C, Palareti G, Guazzaloca G, et al. Venous thromboembolism in young women; role of thrombophilic mutations and oral contraceptive use. *Eur Heart J* 2002;23:984–90.

28. Sidney S, Petitti DB, Soff GA, Cundiff DL, Tolan KK, Quesenberry CP. Venous thromboembolic disease in users of low-estrogen combined estrogen-progestin oral contraceptives. *Contraception* 2004;70:3–10.

29. van Hylckama Vlieg A, Helmerhorst FM, Vandenbroucke JP, Doggen CJM, Rosendaal FR. The venous thrombotic risk of oral contraceptives, effects of oestrogen dose

and progestogen type: results of the MEGA case-control study. *BMJ* 2009;339:b2921.

30. Gialeraki A, Valsami S, Pittaras T, Panayiotakopoulos G, Politou M. Oral contraceptives and HRT risk of thrombosis. *Clin Appl Thromb Hemost* 2018;24:217–25.

31. World Health Organization. Medical eligibility criteria for contraceptive use, 5th ed. Geneva, Switzerland: World Health Organization; 2015.

32. Fahed AC, Wang M, Homburger JR, et al. Polygenic background modifies penetrance of monogenic variants for tier 1 genomic conditions. *Nat Commun* 2020;11:3635.

33. Hynninen Y, Linna M, Vilkkumaa E. Value of genetic testing in the prevention of coronary heart disease events. *PLoS One* 2019;14:e0210010.

34. Mujwara D, Henno G, Vernon ST, et al. Integrating a polygenic risk score for coronary artery disease as a risk-enhancing factor in the pooled cohort equation: A cost-effectiveness analysis study. *J Am Heart Assoc* 2022;11:e025236.

35. Lambert SA, Abraham G, Inouye M. Towards clinical utility of polygenic risk scores. *Hum Mol Genet* 2019;28:R133–42.

36. Emmerich J, Rosendaal FR, Cattaneo M, et al. Combined effect of factor V Leiden and prothrombin 20210A on the risk of venous thromboembolism—pooled analysis of 8 case-control studies including 2310 cases and 3204 controls. Study group for pooled-analysis in

venous thromboembolism. *Thromb Haemost* 2001;86:809–16.

37. Rosendaal FR, Doggen CJ, Zivelin A, et al. Geographic distribution of the 20210 G to A prothrombin variant. *Thromb Haemost* 1998;79:706–8.

38. Folsom AR, Cushman M, Tsai MY, et al. A prospective study of venous thromboembolism in relation to factor V Leiden and related factors. *Blood* 2002;99:2720–5.

39. Bosler D, Mattson J, Crisan D. Phenotypic heterogeneity in patients with homozygous prothrombin 20210AA genotype. A paper from the 2005 William Beaumont Hospital Symposium on Molecular Pathology. *J Mol Diagn* 2006;8:420–5.

40. Elkattawy S, Alyacoub R, Singh KS, Fichadiya H, Kessler W. Prothrombin G20210A gene mutation-induced recurrent deep vein thrombosis and pulmonary embolism: case report and literature review. *J Investig Med High Impact Case Rep* 2022;10:232470 96211058486.

41. Smith KJ, Monsef BS, Ragni MV. Should female relatives of factor V Leiden carriers be screened prior to oral contraceptive use? A cost-effectiveness analysis. *Thromb Haemost* 2008;100:447–52.

42. McDaid A, Logette E, Buchillier V, et al. Risk prediction of developing venous thrombosis in combined oral contraceptive users. *PLoS One* 2017;12:e0182041.

## Author and article information

From the Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala University, Uppsala, Sweden (Dr Faro, Mr T Johansson, and Dr A Johansson); and Centre for Women's Mental Health during the Reproductive Lifespan — Womher, Uppsala University, Sweden (Mr T Johansson).

Received April 26, 2023; revised Aug. 31, 2023; accepted Sept. 13, 2023.

The authors report no conflict of interest.

This study was funded by The Swedish Heart Lung Foundation (20200687), the Swedish Research Council (2019-01497), the Borgström Marcus and Johansson Gustaf Adolf foundations, and the Uppsala University Center for Women's Mental Health during the Reproductive Lifespan. The computations were enabled by resources in project SNIC 2018/8-372 and project sens2017538 provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS) at the Uppsala Multidisciplinary Center for Advanced Computational Science (JPPMAX) that was funded by the Swedish Research Council through grant agreement number 2022-06725. The funders had no role in the study design, data collection, data analysis, data interpretation, writing of the manuscript, or the decision to submit for publication.

The data used for this study are available for bona fide researchers and can be accessed by an application to the UK Biobank.

Corresponding author: Valeria Lo Faro, PhD. [valeria.lo.faro@igp.uu.se](mailto:valeria.lo.faro@igp.uu.se)

## Supplemental Material

### Study population

The United Kingdom Biobank (UKB) is a population-based cohort of more than 500,000 participants from the general population residing in the United Kingdom between 2006 and 2010. Participants were between the ages of 40 and 79 years of age at recruitment. Participants attended 1 of 22 assessment centers across England, Scotland, and Wales. Health and lifestyle data have been collected from questionnaires, interviews, physical measures, and from various health registers. Participants included in our study were women who reported that they were White Irish, White British, or other White, followed by the exclusion of ethnic outliers identified by principal component analysis of the genotype data and with complete data on lifestyle.

### Venous thromboembolism outcome

Identification of the first venous thromboembolism (VTE) episode was obtained from the UKB “First occurrence” data field 41202, including diagnostics recorded before and after the first evaluation visit. These data were regenerated by mapping the clinical codes of primary care, hospital admissions, death records, and self-reported diseases to the 3-character International Classification of Diseases (ICD-10 and ICD-9) codes. For each participant, the data field providing the earliest VTE diagnosis was identified from 1 source. Self-reported medical conditions are derived from data fields 20002 and obtained during initial assessment visits during which participants were asked by trained nurses if they have ever been told by doctors that they have a disease. The interviewer recorded the first diagnosis for each of the illnesses reported by the participant. They could specify either their age or the year in which the diagnosis occurred. If a participant identified both an ICD10 code in the registration and a self-declared medical condition, the earliest date was used. The age at the first diagnosis of VTE was calculated based on the age difference between the date of diagnosis of VTE and the date of

birth and was rounded downward to allow comparison, for example, with the age at the beginning of oral contraceptive use.

### Assessment of exposure

The UKB data fields 2784 (ever taken oral contraceptive pill), 2794 (age started oral contraceptive pill), and 2804 (age when last used oral contraceptive pill) were used to assess oral contraceptive use. Women were followed from birth until the first occurrence of a VTE or to the end of follow-up (ie, age at assessment center visit) or when women had a bilateral oophorectomy, hysterectomy, or entered menopause, whichever came first. For participants who were still using oral contraceptives, the age of last use was set to the age at assessment and those with no information about oral contraceptive use were excluded. Information on menopausal status was obtained during the initial visit. Age at menopause (last menstrual cycle), age at first live birth, age at bilateral oophorectomy, and the age at hysterectomy were assessed from the touchscreen questionnaire completed at the initial visit to the assessment center. Genetic principal components were included as covariates in all models to adjust for confounding owing to population stratification and to reduce the effect of confounding owing to co-incident geographic exposure and outcome.<sup>1</sup> The genetic principal components in the UKB are computed based on the genetic kinship between the individuals of the cohort.<sup>2</sup> Information on the year of birth and Townsend Deprivation Index (TDI)—used as a proxy for socioeconomic status—were obtained from the local National Health Service Primary Care Trust register before participation in the UKB. The TDI score was calculated using a combination of 4 census variables for each geographical area. For the UK Biobank, the TDI scores were calculated using data immediately before the participant joined the UKB and were based on the previous national census output areas. Each participant was assigned a score corresponding to the output area in which their postcode is located.

The sources of exposures, VTE, and covariates information used in this study are listed in below.

Covariates	Data-fields	Source
Diagnosis of venous thromboembolism	41202, 20002	Registers and Self-reported
Date of first venous thromboembolism event	131396, 131398, 131400	Registry
Year of birth	34	Registry
Age	21003	Registry
Townsend deprivation index	189	Derived variable
Age at first live birth	2754	Self-reported
Body mass index	21001	Derived variable
Age started smoking	3436, 2867	Derived variable
Age stopped smoking	2897	Derived variable
Age at bilateral oophorectomy (both ovaries removed)	3882	Self-reported
Age at hysterectomy	2824	Self-reported
Genetic principal components	22009	Derived variable
Age at menopause (last menstrual period)	3581	Self-reported
Age when initiated OC	2794	Self-reported
Age when discontinued OC	2804	Self-reported

### Time-varying covariates and exposure

The Cox proportional hazards model with time-dependent covariates provides a more accurate estimate than when these covariates are used as fixed effects that occurred during the follow-up period. Therefore, confounders with information on the changes that occurred during the follow-up time were incorporated into the model as time-

varying covariates. The inclusion of time-varying covariates and exposure was achieved by applying the counting process approach.<sup>3</sup> In this approach, the data for each individual were divided into multiple episodes. All covariates can be treated as time-fixed within each episode. To arrange and structure data with multiple records for each individual, the `tmerge` function in the R package `survival` was used. Age was used as the primary time scale to allow for a natural nonparametric adjustment of age in the Cox model and follow-up started at age 0. Time-dependent covariates used in the models were oral contraceptive use, pregnancy and postpartum period, and smoking habit. Here, to investigate whether there is a difference in risk associated with length of use, we allowed the exposure variable, in this case oral contraceptive use, to change the value at the first 2 years of use, after 2 years of use, at cessation, and 2 years after the cessation. For example, we defined a variable, `tgroup`, with the following states: `tgroup = 1` describes the first time interval in which an individual is classified as a nonuser (used as the reference group); `tgroup = 2` describes the second time interval in which the individual has been a user for <2 years; and `tgroup = 3` describes the time interval of >2 years of use until the age at cessation (remaining years of use). This approach correctly characterizes the exposure status and classifies the person-time of the user before oral contraceptive initiation as unexposed follow-up time, aiming to avoid immortal time bias. Using the same approach, a period including pregnancy and postpartum was introduced into the model as a time-dependent covariate. This period was calculated by subtracting and adding one unit from the time of delivery, which was obtained from data field 2754 (age at first live birth). A woman was considered unexposed in the year before the time of delivery and exposed 1 year after delivery.

### Genotyping

The third release of imputed genotypes from the UKB data that had already been imputed using a combined 1000 Genomes/UK10K reference panel, was used for the current study. In this study we extracted information for two individual SNPs; rs6025 (effect allele T; allele frequency= 0.02; p-value Hardy-Weinberg= 0.84) in the *FVL* gene and rs179963 (effect allele A; allele frequency= 0.01, p-value Hardy-Weinberg= 1) in the *PTM* gene. Both variants had been directly genotyped or imputed, respectively.

### Polygenic risk score

Polygenic risk scores (PRSs) were obtained from data fields 26289 (standard PRS for venous thromboembolic disease [VTE]) for all the UKB participants. Data supporting these scores were calculated entirely from external genome-wide association study data (the standard PRS set). This work was conducted by the Genomics PLC under UK Biobank project 9659. First, we extracted all White females with PRS scores ( $n=244,420$ ), and then the PRS scores were standardized to have mean=0 and standard deviation=1. Women were categorized into decile groups with the 10<sup>th</sup> decile being considered the highest PRS risk category (the 10% of the women with the highest polygenic risk) and the 1<sup>st</sup> decile being the lowest PRS risk category (the 10% of the women with the lowest polygenic risk).

### Cox regression

All Cox regression models were fitted using the `Coxph` function in the `survival` package. The estimates and *P* values extracted from the model were for the first 2 years of use and the remaining years of use using women who had not initiated oral contraceptive use as the reference group. A second exposure variable was used to stratify women into high and low genetic risk

for VTE. Here, we defined the reference (low-risk) group as women being in the first PRS risk decile and not being carriers of *FVL* and *PTM*. We performed analyses with different definitions of high-risk genetic groups based on combinations of being in the 10<sup>th</sup> PRS decile and/or *FVL* mutation and/or *PTM* carriers. In the combined analyses of genetic risk related to the PRS and oral contraceptive use, the reference group consisted of never users with low genetic risk, that is, those women who had not yet initiated oral contraceptive use, who were in the first PRS decile, and who were non-carriers of *FVL* and *PTM*. In the analysis of the risk of *FVL* and *PTM* variants, we instead used all noncarriers as the control category, irrespective of any PRS status. To ensure effective control of age, we used age as a timescale in the Cox models, which is an optimal approach to adjust for age in a nonlinear fashion and useful when age dramatically influences the hazard rates of the outcomes.

### Interaction analysis

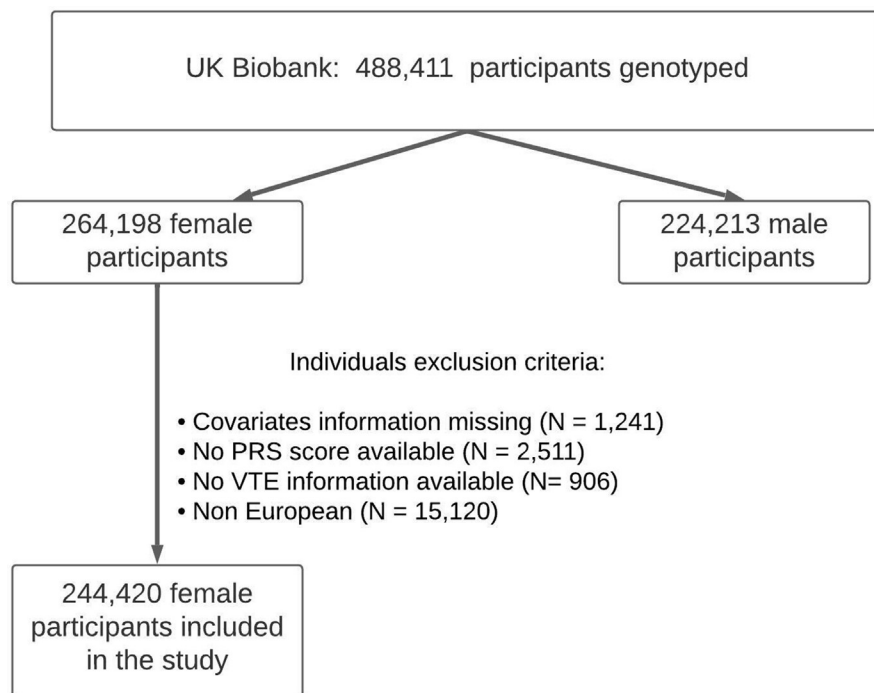
We performed interaction analyses using the Cox proportional hazard regression to formally test for an interaction effect between the genetic risk and oral contraceptive use on VTE risk. Interaction terms for the genetic risk (PRS, *FVL*, or *PTM*) and the time-dependent oral contraceptive exposure were then included in the Cox regression models.

### Supplemental References

1. Haworth S, Mitchell R, Corbin L, et al. Apparent latent structure within the UK Biobank sample has implications for epidemiological analysis. *Nat Commun* 2019;10:333.
2. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018;562:203–9.
3. Therneau TM, Grambsch PM. *Modeling survival data: extending the Cox model*. Berlin, Germany: Springer Science & Business Media; 2013.

## SUPPLEMENTAL FIGURE

## Workflow for inclusion and exclusion of UKB participants



PRS, polygenic risk score; UKB, UK Biobank; VTE, venous thromboembolism.

Lo Faro. Genetic factors and risk of venous thromboembolism in women using oral contraceptives. *Am J Obstet Gynecol* 2024.