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, Åsa Johansson, PhD

PII: S0002-9378(23)00618-X

DOI: https://doi.org/10.1016/j.ajog.2023.09.012

Reference: YMOB 15263

To appear in: *American Journal of Obstetrics and Gynecology*

Received Date: 26 April 2023

Revised Date: 31 August 2023

Accepted Date: 13 September 2023

Please cite this article as: Lo Faro V, Johansson T, Johansson Å, 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, *American Journal of Obstetrics and Gynecology* (2023), doi: https://doi.org/10.1016/j.ajog.2023.09.012.

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[1], Therese Johansson[1,2], Åsa Johansson, PhD[1]

[1] Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala University, Uppsala, Sweden

[2] Centre for Women's Mental Health during the Reproductive Lifespan – Womher, Uppsala University, Sweden

Conflict of Interest

The authors report that they have no competing interests.

Sources of Funding

This work 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 funders had no role in study design, data collection, data analysis, data interpretation, writing of the manuscript, or the decision to submit for publication.

Correspondence

Valeria Lo Faro, Ph.D. (valeria.lo.faro@igp.uu.se)

Word count Manuscript: 4550

Word count abstract: 464

Condensation page:

Tweetable statement: Polygenic predisposition, and not only the well-known inherited thrombophilia variants, dramatically increases the risk of venous thromboembolism in women using oral contraceptives.

Short Title: Genetic factors and risk of venous thromboembolism in women using oral contraceptives

AJOG at a Glance:

A. Why was this study conducted?

- Oral contraceptive (OC) use increases the risk of venous thromboembolism (VTE) by a factor of three-to-five.
- 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 in OC users are lacking.

B. What are the key findings?

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

C. What does this study add to what is already known?

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

Abstract:

Background. Over 150 million women worldwide use oral contraceptives. Women with inherited thrombophilia and carriers of certain thrombophilia gene variants, such as factor V Leiden and prothrombin mutation, are at increased risk of venous thromboembolism, especially in combination 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 is to estimate the risk of venous thromboembolism when initiating oral contraceptive use (first two years) and during continued use in 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 two years and in the remaining years of OC use on the risk of venous thromboembolism risk was estimated using 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 of venous thromboembolism was observed during the first two years of oral contraceptive use (hazard ratio=3.09; 95% CI = 3.00 - 3.20), but not during continued use (hazard ratio =0.92;95% CI, 0.80 – 1.05). However,

when genetic risk was considered, women with the highest polygenic risk scores risk category had a more pronounced risk of venous thromboembolism during the first two years of oral contraceptive use (hazard ratio = 6.35; 95% CI, 4.98 - 8.09), and a high risk was also observed in factor V Leiden (hazard ratio, 5.73 [95% CI, 5.31- 6.17]) and prothrombin variant carriers (hazard ratio, 5.23 [95% CI, 4.67 – 5.87]). A high polygenic risk score in combination with being factor V Leiden and prothrombin variant carrier resulted in the highest risk of venous thromboembolism during the first two years of oral contraceptive use (hazard ratio, 14.8 [95% CI, 9.28 - 23.6]). Women with a high genetic liability also had an increased risk during continued use but less pronounced, with the highest risk in combination with being carriers of both factor V Leiden and prothrombin variant (hazard ratio, 4.93 [95% CI, 3.16 - 7.7]).

Conclusions. Polygenic risk can capture 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 of venous thromboembolism, particularly in women with a high genetic predisposition, and that oral contraceptive use dramatically increases the risk short after initiation of use, which decreases with continued use. This suggests that polygenic risk score could be used to identify women at high risk of developing venous thromboembolism and advise them on alternative methods of contraception.

Word count Abstract: 464

Keywords: venous thromboembolism; polygenic score; oral contraceptives; risk assessment; factor V Leiden; prothrombin G20210A.

Introduction

Oral contraceptives (OC) enable women to control their fertility.[1] However, studies have reported an increased risk of thrombotic events in OC users.[2] Excess estrogenicity of OC (the sum of estrogen and progestin contributions) increases the risk of venous thromboembolism (VTE).[3,4] VTE is a leading cause of cardiovascular death worldwide.[5] Each year in Europe, it is estimated that approximately 22,000 VTE events are related to OC use.[6] VTE is a complex disorder that is influenced by both acquired and inherited factors. The acquired factors include, among others, the use of OCs.[7] The inherited factors are represented by the Factor V Leiden (FVL) and prothrombin Factor II variants (PTM). From twin studies, VTE heritability has been estimated to be 50%.[8] However, today known genetic variants can explain only 6% of the heritability.[9] In women who use OC, the risk of VTE is three-to-five times higher compared to women who have never used OC, with the highest risk during the first two years of use.[10] Furthermore, the alteration in hemostatic imbalance and consequently the increased risk of VTE compared to the general population is more pronounced in women with a monogenic hereditary thrombophilia condition.[11] The World Health Organization (WHO) states that the use of OC in these women is associated with an unacceptable health risk. However, VTE is a polygenic disorder and genetic liability to VTE can also be assessed as polygenic risk scores (PRS).[12]

Currently, risk assessment in contraceptive counselling is based on clinical characteristics and family history of VTE; the latter has shown poor sensitivity and predictive performances.[13–15] The main aim of this study was to estimate the risk of VTE associated with initiating OC use and with continued use in women with a high genetic liability, using both 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.

Methods

Study cohort

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

Study design

We investigated the OC-associated VTE risk in female participants in the UKB cohort. Our study was designed as a prospective study, where VTE is a binary outcome and the rate of VTE was assessed in all women in relation to their exposure to OC. Women with missing information about their OC use, any of the covariates used, not genotyped, or were not white European, were excluded from the analyses (Supplementary Figure 1), resulting in 244,420 women in the analyses. Our study was designed to follow the women from birth (age = 0)

until the first of the following events occurred: VTE diagnosis, end-of-study follow-up (i.e., age at recruitment), having bilateral oophorectomy or hysterectomy, or entered menopause. To examine whether the use of OC in combination with a high genetic liability for VTE confers an increased risk, we stratified the cohort in ten deciles of risk according to the PRS scores (using the 1st decile as the reference) and/or according to their carrier status of FVL and PTM (using the non-carriers as the reference). This study was approved by UKB (application #41143) and the Swedish Ethical Review Authority (dnr: 2020-04415).

Assessment of exposure, outcome, and covariates

Information on OC use**,** including age when initiating and discontinuing, was assessed during the initial assessment visit. The relevant UKB data fields include 2784 (ever taken OC pill), 2794 (age started OC pill), and 2804 (age when last used OC pill). The first occurrence of VTE in the UKB was based on medical history and linkage to data on hospital admissions and cause of death register. The UKB data fields include the following ICD9 and IC10 codes 4151, 4511, 4532, I80, I81, I82, and I26 extracted from health records, and self-reported VTE extracted from field codes 20002 (1068, 1093 and 1094). See Supplementary Material for information on the assessment of covariates.

Genotyping and polygenic risk scores

The UKB participants had been genotyped using the UKB Axiom array and the UK BiLEVE array and untyped variants have been imputed using SHAPEIT.[16] 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 (Supplementary

Material). The PRS for VTE used in this study had already been calculated by Genomics PLC under UK Biobank project 9659, and was provided by UKB (UKB data field 26289 - Standard PRS for VTE).[17] This particular PRS had been trained on the eMERGE (The Electronic Medical Records and Genomics) cohort (releases 2, 3, 5, and 6), which is a consortium of ten participating sites that jointly perform genome-wide association studies (GWAS) and makes the respective summary statistics freely available.[18] These cohorts do not overlap with the UK Biobank with regards to 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 UKB genetically inferred ancestry groups and either GnomAD or the 1000 Genomes Project, and those with evidence of large departures from Hardy-Weinberg equilibrium (p-value > 1e-10) were excluded. The PRS algorithm was constructed using a Bayesian approach, which has been described previously.[17] The PRS has already been validated in the UKB in a recent article where it was applied to the risk of deep vein thrombosis (a manifestation of venous thromboembolism).[19] Its performance was then compared to two other PRS, one trained in half of the UKB and one from the Global Biobank Meta-analysis Initiative consortium effort. It was observed that the area under the curve (AUC) estimates improved from 0.60 (95 % CI:0.59–0.61) of the conventional risk factors (sex, age, and principal components) to 0.66 (95 % CI: 0.65-0.67) when also the eMERGE PRS was included. In addition, we also validated the PRS in relation to VTE as part of the current study. For calculating the AUC and its 95 % confidence intervals, we used the R package pROC.[20]

Cox regression

Cox regression analyses were performed to calculate the instantaneous VTE risk during the use of OC. We only considered first events of VTE in our study since women are censored after the first VTE diagnosis. The follow-up started at birth and age was used as the primary time-scale. Women were followed until one of the first of the following events occurred: VTE diagnosis, end-of-study follow-up (i.e., age at assessment center visit), when women had a bilateral oophorectomy, a hysterectomy, or 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 four genetic principal components (see Supplementary Material for more details). The genetic principal components were included as covariates to adjust for confounding due to population stratification and computed based on the genetic kinship between the individuals of the cohort.[16,21] The use of OC was modelled as a time-varying variable where all women were unexposed at age = 0 but the exposure status changed to exposed = 1 when women initiated the OC use (Supplementary 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 two years. This means that in the analyses, the incidence rate of VTE during the first two years after initiating the use is compared to the incidence rate of women of the same age who have not so far used OC. Similarly, the rate during remaining years of use and up until two years after cessation is compared to women of the same age who have not so far used OC. When estimating the effect during use, women were censored two years after stopped using OC. The reason for considering up to two years after discontinuation as continued use is

because there is a risk that some women developed a VTE just before they stopped, but it looks like the events occurred after they stopped using them which will introduce protopathic bias, also referred to as “reverse causality”. Therefore, we included a two-year lag time, as has been discussed previously.[22]

A second exposure variable was used to stratify women into high/low genetic VTE risk (Supplementary Material for more details). For the analyses that include the PRS, we defined the reference group as women with the lowest genetic liability to VTE i.e., being in the 1st PRS decile and not being carriers of FVL and PTM in order to compare to the high genetic VTE risk. In the analyses of the risk of FVL and PTM, we used all non-carriers as the reference group, irrespective of their PRS status, in order to compare to the FVL and PTM carriers. The Cox regression modelling was performed using the 'survival' R package and hazard ratio (HR) and its confidence intervals were calculated.[23,24]

Results

A total of 244,420 women were included in our analyses (Table 1), of which 10,856 experienced a first ever VTE event during the follow-up. A total of 193,371 had initiated OC at some time points during the follow-up. The never user group included a larger number of women who reported a VTE episode. This is most probably because never user women 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, 8,682 were carriers of FVL and 4,119 of PTM. The frequency of FVL was 4.48% in OC users and 4.51% in never users, and for PTM was 2.13% in users and 2.20% in never users. There was no

significant difference in PRS between ever and never users of OC, indicating that bias due to confounding by indication is unlikely to affect our results. Descriptive statistics for the different genetic risk groups and time of OC use analyzed in this study are summarized in Supplementary Table 1.

Genetic predisposition and risk of venous thromboembolism

We categorized the women as those having the highest PRS VTE (referred to as the 10th decile), being carriers of the FVL, and/or being carriers of PTM. A total of 24,291 were in the highest PRS category, while, (independently from the PRS category) 10,985 women were carriers of FVL (either homozygous or heterozygous), and 5,244 of PTM (Supplementary Table 2). We estimated the VTE risk in the entire cohort and in the subgroup of never users (Supplementary Table 3). All high genetic-risk groups were associated with a significantly higher incidence rate of VTE compared with the respective reference group.

We also validated the PRS in the UKB. We estimated the AUC for a Base model (including age and genetic principal components), the Base plus the FVL and/or PTM, and the Base plus the PRS. We observed that the Base plus the PRS model improved classification over both the Base and the Base plus FVL and PTM models. The carrier status for the two variants increased prediction by around 1.0 % in the AUC, compared with 3.5 % in the AUC for the PRS (Figure 1, Supplementary Table 4). After, we estimated the odds ratio and 95 % confidence interval using logistic regression for the PRS decile in our cohort (Figure 2). Here, we observed a trend of higher odds of VTE in those being in the 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 in all women, without considering their genetic predisposition to VTE. The risk during the first two years of OC use was associated with an increased hazard of 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 based on their PRS (independently from being FVL or PTM carriers) into 1st and 10th PRS deciles and estimated the interaction effect in the first two years of OC use, compared to never OC users. The effect during the first two years of use was significantly higher (interaction $P < 0.001$) in the 1st compared to the 10th PRS decile. In contrast, in the two strata with FVL and PTM carriers, the first two years of OC use were associated with similar risk to the women in the 1st PRS decile.

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

Among the women in the 10th PRS decile as a genetic risk factor, the HR during the first two years of OC use was associated with an increase in VTE risk (HR, 6.35 [95% CI, 4.98 - 8.09]) as compared to 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 10th PRS decile as a genetic risk factor (neither FVL nor PTM carriers), the HR during the first two 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 of those women being in the 10th decile PRS and carriers of FVL and/or PTM as compared to the reference category. The first two years of OC use showed an increased HR (HR, 8.78 [95% CI, 6.12 - 12.6] in the 10th decile PRS and carriers of FVL, while the HR was 10.58 (95%

CI, 7.48 - 14.97) in the 10th decile PRS and 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) in the 10th decile of PRS and carriers of FVL and PTM, respectively (Figure 3). Among women with FVL as a genetic risk factor, the HR was increased during the first two years of OC use (HR, 5.73 [95% CI, 5.31 - 6.17]), regardless of which PRS category they belonged to. The HR remained significant 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 hazard rate during the first two 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 carrying both FVL and PTM was 9 (95% CI, 6.07 - 13.34) in the first two years, and 3.39 (95% CI, 2.38 - 4.83) in the remaining years.

We estimated the risk in women being in the 10th decile PRS and who were both FVL and PTM carriers; in the first two 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 two years of OC use; N=14 in the remaining years of use), and confidence intervals for these estimates are wide.

COMMENT

Principal Findings

We estimated the risk of VTE in women using OC in relation to their genetic predisposition. We showed that women in the highest PRS decile had a more than 6-fold increased risk of VTE during the first two years of OC use compared to never users in the lowest genetic risk. This increased risk is higher than the risk of being an FVL or PTM carrier. Our results highlight that polygenic risk has an increased impact on the occurrence of VTE in relation to OC.[11,25–27] Therefore, combining genetic liability (including several common low-effect variants) with clinical risk factors may allow better VTE risk stratification associated with OC use compared to 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 two years of OC use and the remaining years of OC use (considered until two years after discontinuation), with a many-fold increased risk associated with the first two years. This highlights the importance of treating OC use as a time-varying exposure variable rather than estimating an average HR of the duration of the follow-up for users versus never users.

Results in the Context of What is Known

OC use has previously been shown to be associated with a three-to-five times higher risk of VTE, with the highest risk observed during the first two years of use.[28,29] Consistent with this, in our study, the hazard ratio for VTE in women using OC in the first two years was 3.09, compared to never users when genetic information was considered. The incidence of VTE in premenopausal women is about 3 per 10,000 women per year.[30] However, given that more than 150 million women worldwide use OC, even a small increase in the risk of VTE associated with OC use results in a substantial increase in the number of VTE cases. From the reported estimates of the different AUC models, we observed that the PRS improved classification with an increased prediction of 3.5%. In the cohort of all women from our study, a high PRS was

associated with a higher rate of VTE (HR = 2.2), which was slightly higher compared to 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 appears to be modest, however, the AUC calculated in different models, containing the Base, Base plus FVL and/or PTM, and Base plus the PRS, showed that the PRS improved classification over the Base model and the model including FVL and PTM. The carrier status for the two variants increased prediction by around 1.0 % in the AUC, compared with 3.5 % in the AUC for the PRS. A study conducted for VTE risk in a general population (independent of OC use and also including males) that used a different PRS from ours, reported an improvement of 4 %, which is in line with our AUC estimate.[9] However, the PRS does not capture all genetic risk. In fact, the PRS additively incorporates known risk-associated loci but does not consider interactions (i.e. gene-gene or gene-environment) or variants that act only in specific genetic backgrounds (i.e. 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. Also, 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 dataset. 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 utility of PRSs.

Clinical Implications

The WHO classifies OC as an unacceptable health risk for women with known thrombogenic variants, but it discourages global screening for thrombophilia before prescribing OC due to the low prevalence of thrombophilia (7-8% among Europeans) and high cost of screening.[31]

345 VTE is a polygenic disease involving thousands of genetic variants that collectively contribute
346 to the risk of a thrombotic event. As the PRS accounts for a greater proportion of the genetic
347 contribution to disease, using the full spectrum of genetic risk for thrombotic events may
348 improve risk stratification and identify numerically more individuals at higher risk among oral
349 contraceptive users, compared to only FVL and PTM carriers. A recent study has also shown
350 that considering polygenic background improves the risk accuracy estimates in individuals
351 who carry a monogenic risk variant, which may better inform decision-making and refine risk
352 estimates during counselling.[32] For women with a high genetic risk of VTE, this means that
353 using OC dramatically increases their already high risk of VTE, especially in the first two years.
354 Therefore, women in this situation need careful counselling about their contraceptive
355 methods, including evaluation of other more appropriate contraceptive choices. These
356 implications of PRS could better inform and improve the decision-making process, and
357 therefore refine risk estimates during counselling. The presence of other risk factors for
358 thrombosis, such as dyslipidemia, smoking, and obesity, should be considered when advising
359 these women about oral contraceptive therapy.
360
361 Because genotyping can nowadays be done at a low cost, efforts to integrate genotyping data
362 in healthcare systems may facilitate the use of the PRS as a tool to improve the identification
363 of high-risk women when using oral contraceptives. A preliminary health economic analysis
364 study has shown the potential cost benefits of using PRS in cardiovascular disease prevention
365 in the Finnish health system.[33] Indeed, it was found that cardiovascular disease PRS together
366 with traditional risk factors would be cost-beneficial if deployed in a targeted approach.
367 However, evaluation studies in clinical settings on women using oral contraceptives in relation
368 to the 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 assessment should be performed in a clinical setting, where the incorporation and utility of PRS will be assessed in high-risk groups, such as in smokers, in obese, and in patients with cardiovascular disease. Evaluations also in different populations will be required. It is important to note that the characteristics of PRSs open up 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 done for a wide range of diseases. For those with a significantly increased lifetime risk of disease, targeted interventions could be used to reduce their risk, for example through guidance on 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 both FVL and PTM, to construct 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 one-time PRS cost is between $80 and $120.[34] Therefore, it is possible that there would be a significant health and economic benefit by performing genotyping to be used 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.[35]

So far, most studies have evaluated the combined effect of OC use with FVL and PTM. Therefore, information on the combined effect of OC and common genetic variants is still scarce.[25,26] Here, 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 of VTE was not constant over time in OC users. This study contributes to a more accurate

393 estimation of the risk associated with OC use in women with defects in two thrombophilia
394 genes. From our study, the presence of the highest polygenic risk, FVL, and PTM, in women
395 using OC seems to have an additive effect. Because we found only 27 OC users with a VTE
396 event in the 10th PRS and carrying both mutations, the confidence interval is large and the
397 estimated risk should be considered an approximation (Figure 2, "10th+PTM+FVL"). Here,
398 there were very few or no homozygotes for FVL or PTM. A limited number of studies in the
399 literature have described patients who are homozygous carriers of FVL and, usually, these
400 patients develop their first thrombosis at a younger age with a 10-to-80-fold increased VTE
401 risk compared to controls.[36–38] Individuals who are homozygous for PTM are the rarest, with
402 only 141 homozygous PTM cases, mostly Southern European, found until 2022.[39,40]
403
404 **Research Implications**
405 We emphasize that not only individuals with FVL and PTM but also those with a high polygenic
406 liability to VTE are at high risk and should therefore consider alternatives to OC. Evaluations
407 of the benefit of introducing global screening for thrombophilia in those who wish to start OC
408 therapy are based only on economic analyses. These analyses do not examine the duration of
409 therapy and the benefits of knowing one's genetic risk. For women who are FVL carriers, the
410 pharmacoeconomic evaluation by Smith (which included 15 years of OC use) found that
411 screening and counselling was an economically favorable strategy.[41] The next step in the
412 pharmacoeconomic evaluation will be to assess the impact of introducing global screening
413 before starting OC.[42]
414
415 **Strengthens**

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, are at higher risk of developing VTE when using OC. We showed that OC use in the presence of a high genetic liability is a circumstantial VTE risk factor in women of reproductive age.

Limitations

First, the UKB cohort consists of healthier individuals compared to the general UK population (i.e., the individuals that volunteer to participate have healthier lifestyles, higher levels of education, and better health than the general UK population), and only analyzed white European women, so the results of the combined effect of OC and genetic risk factors should be replicated in other larger populations and 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 oral contraceptives and on the oral route of administration. Third, exposure to OC was assessed by self-report questionnaires, which is 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.

Conclusions

Common genetic variants may capture additional risk not accounted for by traditional clinical and genetic factors. OC use is associated with a dramatically increased risk of VTE in highly genetically predisposed women, not only in carriers of the known FVL and PTM variants,

especially at the beginning of use. Further studies, also in other populations and ancestries, are needed to confirm our findings.

Glossary of Terms

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 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, which 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 data from national census data.

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

Acknowledgements

We acknowledge all of the participants and staff involved in the UKB for their valuable contribution. 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 Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX), funded by the Swedish Research Council through grant agreement no. 2022-06725.

Authorship Contributions

VLF, TJ, and ÅJ conceptualized, and designed the experimental setup. VLF performed data analyses. VLF and ÅJ interpreted the results. VLF, TJ, and ÅJ wrote the main manuscript. ÅJ provided infrastructure and financial support. All authors have read and agreed to the published version of the manuscript.

Data and code availability

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

References:

1. Morimont, L., Haguet, H., Dogné, J.-M., Gaspard, U. & Douxfils, J. Combined Oral Contraceptives and Venous Thromboembolism: Review and Perspective to Mitigate the Risk. *Front. Endocrinol.* **12**, 769187 (2021).
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.* **36**, 61–69 (2021).
3. Tepper, N. K., Whiteman, M. K., Marchbanks, P. A., James, A. H. & Curtis, K. M. Progestin-only contraception and thromboembolism: A systematic review. *Contraception* **94**, 678–700 (2016).
4. Mantha, S. *et al.* Assessing the risk of venous thromboembolic events in women taking progestin-only contraception: a meta-analysis. *BMJ* **345**, e4944 (2012).
5. Blann, A. D. & Lip, G. Y. H. Venous thromboembolism. *BMJ* **332**, 215–219 (2006).
6. Odlind, 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.* **81**, 482–490 (2002).
7. Gutmann, C. H. *Venous Thrombosis: Risk Factors, Management and Complications*. (Nova Science Pub Incorporated, 2014).
8. Larsen, T. B. *et al.* Major genetic susceptibility for venous thromboembolism in men: a study of Danish twins. *Epidemiology* **14**, 328–332 (2003).
9. Ghouse, J. *et al.* Genome-wide meta-analysis identifies 93 risk loci and enables risk prediction equivalent to monogenic forms of venous thromboembolism. *Nat. Genet.* (2023) doi:10.1038/s41588-022-01286-7.
10. Vandenbroucke, J. P. *et al.* Oral contraceptives and the risk of venous thrombosis. *N. Engl. J. Med.* **344**, 1527–1535 (2001).
11. van Vlijmen, E. F. W., Wiewel-Verschueren, S., Monster, T. B. M. & Meijer, K. Combined oral

contraceptives, thrombophilia and the risk of venous thromboembolism: a systematic review and meta-analysis. *J. Thromb. Haemost.* **14**, 1393–1403 (2016).

12. Khera, A. V. *et al.* Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat. Genet.* **50**, 1219–1224 (2018).
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.* **163**, 1105–1109 (2003).
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.* **8**, 139–141 (2002).
15. Bezemer, I. D., van der Meer, F. J. M., Eikenboom, J. C. J., Rosendaal, F. R. & Doggen, C. J. M. The value of family history as a risk indicator for venous thrombosis. *Arch. Intern. Med.* **169**, 610–615 (2009).
16. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
17. Thompson, D. J. *et al.* UK Biobank release and systematic evaluation of optimised polygenic risk scores for 53 diseases and quantitative traits. *medRxiv* 2022.06.16.22276246 (2022) doi:10.1101/2022.06.16.22276246.
18. McCarty, C. A. *et al.* The eMERGE Network: a consortium of biorepositories linked to electronic medical records data for conducting genomic studies. *BMC Med. Genomics* **4**, 13 (2011).
19. Lo Faro, V., Johansson, T., Höglund, J., Hadizadeh, F. & Johansson, Å. Polygenic risk scores and risk stratification in deep vein thrombosis. *Thromb. Res.* **228**, 151–162 (2023).
20. Robin, X. *et al.* pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* **12**, 1–8 (2011).
21. Haworth, S. *et al.* Apparent latent structure within the UK Biobank sample has implications for epidemiological analysis. *Nat. Commun.* **10**, 333 (2019).

547 22. Tamim, H., Monfared, A. A. T. & LeLorier, J. Application of lag-time into exposure definitions to
548 control for protopathic bias. *Pharmacoepidemiol. Drug Saf.* **16**, 250–258 (2007).
549 23. Thiébaut, A. C. M. & Bénichou, J. Choice of time-scale in Cox's model analysis of epidemiologic
550 cohort data: a simulation study. *Stat. Med.* **23**, 3803–3820 (2004).
551 24. Cologne, J. *et al.* Proportional hazards regression in epidemiologic follow-up studies: an intuitive
552 consideration of primary time scale. *Epidemiology* **23**, 565–573 (2012).
553 25. Vandenbroucke, J. P. *et al.* Increased risk of venous thrombosis in oral-contraceptive users who
554 are carriers of factor V Leiden mutation. *Lancet* **344**, 1453–1457 (1994).
555 26. Martinelli, I., Taioli, E., Bucciarelli, P., Akhavan, S. & Mannucci, P. M. Interaction between the
556 G20210A mutation of the prothrombin gene and oral contraceptive use in deep vein
557 thrombosis. *Arterioscler. Thromb. Vasc. Biol.* **19**, 700–703 (1999).
558 27. Legnani, C. *et al.* Venous thromboembolism in young women; role of thrombophilic mutations
559 and oral contraceptive use. *Eur. Heart J.* **23**, 984–990 (2002).
560 28. Sidney, S. *et al.* Venous thromboembolic disease in users of low-estrogen combined estrogen-
561 progestin oral contraceptives. *Contraception* **70**, 3–10 (2004).
562 29. van Hylckama Vlieg, A., Helmerhorst, F. M., Vandenbroucke, J. P., Doggen, C. J. M. & Rosendaal,
563 F. R. The venous thrombotic risk of oral contraceptives, effects of oestrogen dose and
564 progestogen type: results of the MEGA case-control study. *BMJ* **339**, b2921 (2009).
565 30. Gialeraki, A., Valsami, S., Pittaras, T., Panayiotakopoulos, G. & Politou, M. Oral Contraceptives
566 and HRT Risk of Thrombosis. *Clin. Appl. Thromb. Hemost.* **24**, 217–225 (2018).
567 31. *Medical Eligibility Criteria for Contraceptive Use*. (World Health Organization).
568 32. Fahed, A. C. *et al.* Polygenic background modifies penetrance of monogenic variants for tier 1
569 genomic conditions. *Nat. Commun.* **11**, 3635 (2020).
570 33. Hynninen, Y., Linna, M. & Vilkkumaa, E. Value of genetic testing in the prevention of coronary
571 heart disease events. *PLoS One* **14**, e0210010 (2019).
572 34. Mujwara, D. *et al.* Integrating a Polygenic Risk Score for Coronary Artery Disease as a Risk-

573 Enhancing Factor in the Pooled Cohort Equation: A Cost-Effectiveness Analysis Study. *J. Am.*
574 *Heart Assoc.* **11**, e025236 (2022).
575 35. Lambert, S. A., Abraham, G. & Inouye, M. Towards clinical utility of polygenic risk scores. *Hum.*
576 *Mol. Genet.* **28**, R133–R142 (2019).
577 36. Emmerich, J. *et al.* Combined effect of factor V Leiden and prothrombin 20210A on the risk of
578 venous thromboembolism--pooled analysis of 8 case-control studies including 2310 cases and
579 3204 controls. Study Group for Pooled-Analysis in Venous Thromboembolism. *Thromb.*
580 *Haemost.* **86**, 809–816 (2001).
581 37. Rosendaal, F. R. *et al.* Geographic distribution of the 20210 G to A prothrombin variant.
582 *Thromb. Haemost.* **79**, 706–708 (1998).
583 38. Folsom, A. R. *et al.* A prospective study of venous thromboembolism in relation to factor V
584 Leiden and related factors. *Blood* **99**, 2720–2725 (2002).
585 39. Bosler, D., Mattson, J. & Crisan, D. Phenotypic Heterogeneity in Patients with Homozygous
586 Prothrombin 20210AA Genotype. A paper from the 2005 William Beaumont Hospital
587 Symposium on Molecular Pathology. *J. Mol. Diagn.* **8**, 420–425 (2006).
588 40. Elkattawy, S., Alyacoub, R., Singh, K. S., Fichadiya, H. & Kessler, W. Prothrombin G20210A Gene
589 Mutation-Induced Recurrent Deep Vein Thrombosis and Pulmonary Embolism: Case Report and
590 Literature Review. *J Investig Med High Impact Case Rep* **10**, 23247096211058486 (2022).
591 41. Smith, K. J., Monsef, B. S. & Ragni, M. V. Should female relatives of factor V Leiden carriers be
592 screened prior to oral contraceptive use? *Thromb. Haemost.* **100**, 447–452 (2008).
593 42. McDaid, A. *et al.* Risk prediction of developing venous thrombosis in combined oral
594 contraceptive users. *PLoS One* **12**, e0182041 (2017).

595
596

597

Tables

	Users of oral contraceptive	Never users of oral contraceptive	P-value *
Number (%)	193371	51049	-
Venous thromboembolism events, N (%)	7687	3169	3.78E-105
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 when initiated oral contraceptive, median (Q1–Q3)	21 (18–24)	-	-
Age when discontinued oral contraceptive, median (Q1–Q3)	30 (26–37)	-	-
Duration of oral contraceptive use, median (Q1–Q3)	9 (4–15)	-	-
First two 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 (%)	34273	11373	5.521E-122
Had bilateral oophorectomy, N (%)	14856	4893	8.659E-45
Townsend deprivation index, median (Q1–Q3)	-2.30 (-3.7 - 0.12)	-2.17 (-3.62 - 0.46)	< 2.2e-16
FVL carriers, N (%)	8682	2303	0.8345
PTM carriers, N (%)	4119	1125	0.3059
Delivery, N (%)	133150	32931	< 2.2e-16
Smoking, N (%)	55479	11773	1.599E-141
Polygenic Risk score	-0.03 (-0.66 - 0.63)	-0.02 (-0.65 - 0.64)	0.32

Table 1: Characteristics of the entire study participants. Numbers (N) are given as median (Q1= first quartile; Q3= third quartile) for continuous data and total number and percentage for binary data. * Mann-Whitney U Test for quantitative traits and Pearson $\chi 2$ test for binary traits, without considering any potential confounding.

Figure legends

Figure 1: Discriminatory ability of VTE polygenic risk scores among female participants of the UKB. Receiver operating characteristic curves assess the discriminative power of different significant models. The grey dot line with an area under the curve of 50 % is used as reference.

Figure 2: Odds ratio estimates for each VTE polygenic risk score decile. The first decile was used as reference to the others. The odds ratio and 95 % confidence interval were estimated using logistic regression. Each point indicates the odds ratios and the bar is lower and upper 95 % confidence interval for each odds ratio. In the upper part of the figure, the density distribution plot of VTE cases versus controls is shown.

Figure 3: Time-varying risk of venous thromboembolism according to PRS, FVL, presence of PTM, and oral contraceptive use. Hazard ratios (HRs) for venous thromboembolism were calculated for both the first two years of oral contraceptive use (shown as blue squares) in the upper part of each group and for the remaining years of oral contraceptive use (shown as red circles) in the lower part of each comparison group. The HRs for venous thromboembolism and their error bars, which indicate 95% confidence intervals, are shown in the HR (95% CI) column. The rate per 1,000 person-years (RPY) column shows the rate of venous thromboembolism per 1,000 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 polygenic risk score (1st PRS), who were not carriers of both FVL and PTM and who had never used oral contraceptives. In the analysis of the risk of FVL/PTM, we used non-user and non-carriers 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 four principal components. Abbreviations: OC, oral contraceptive; PRS, polygenic risk score; CI, confidence interval; RPY, rate per 1,000 person-years.

Supplementary Figure legend

Supplementary Figure 1: Workflow for inclusion and exclusion of UKB participants.

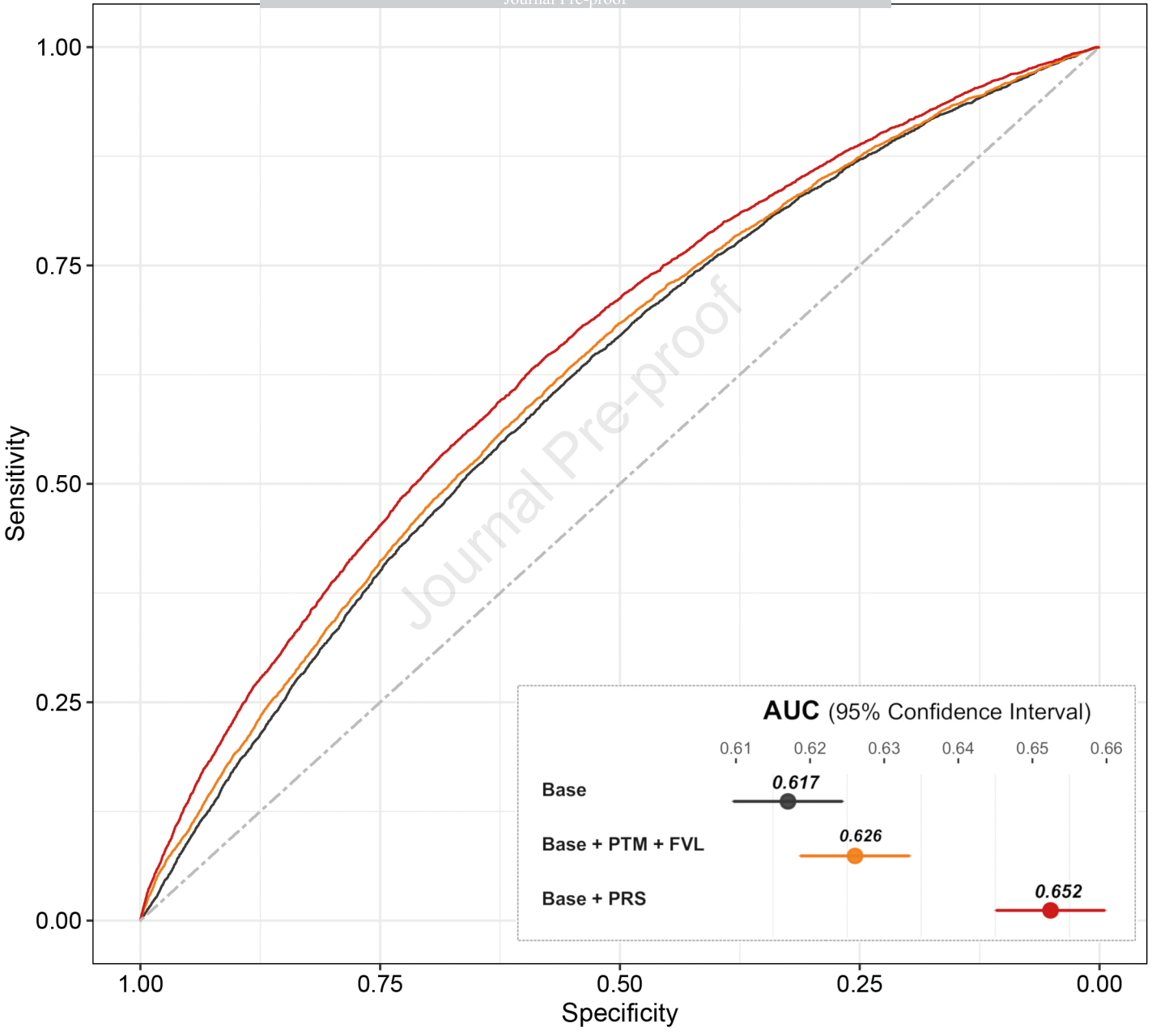
Supplementary Tables legends:

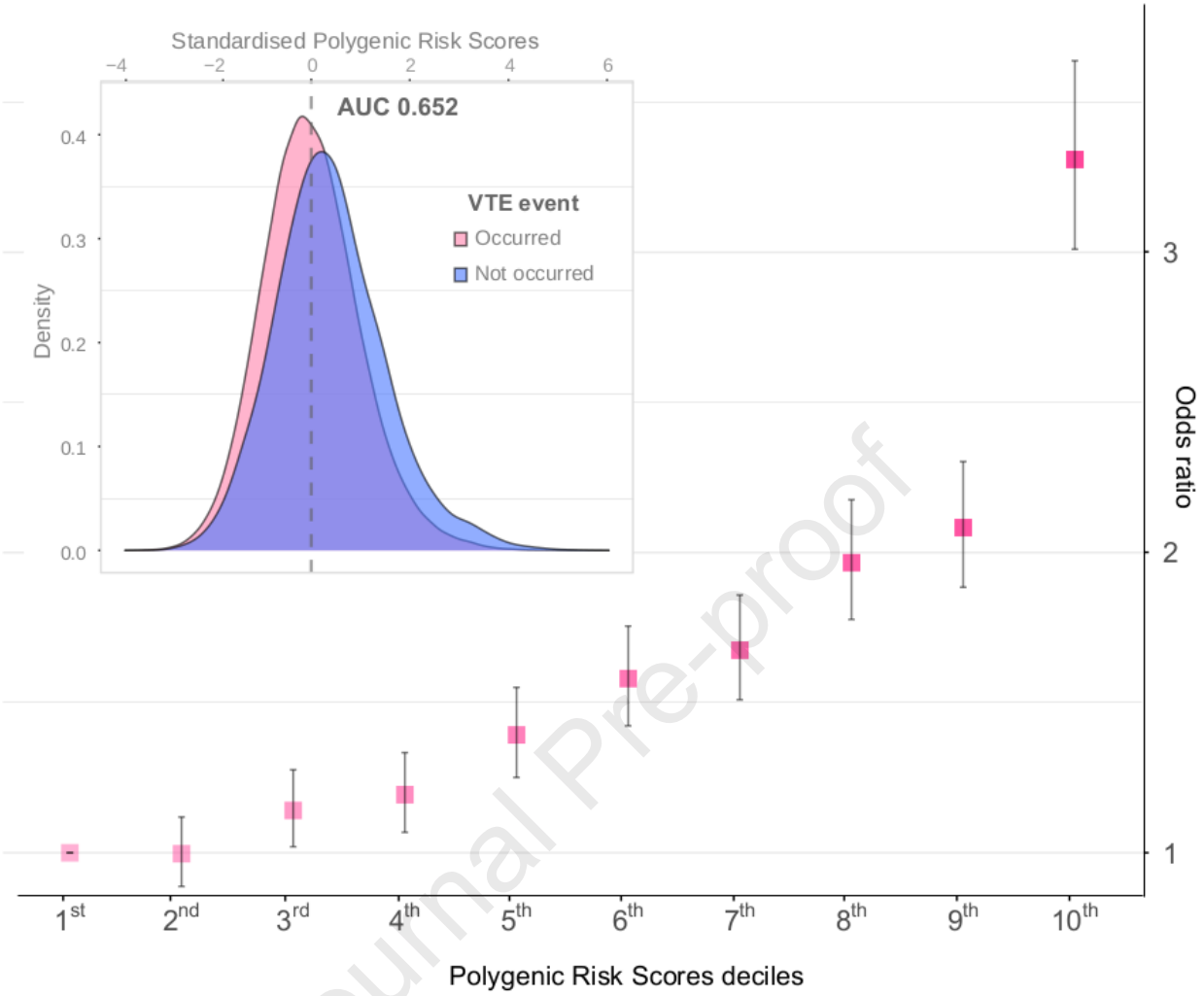
Supplementary Table 1: Descriptive statistics for the different genetic risk groups and oral contraceptive time of use analyzed in this study. Numbers are given as median (Q1= first quartile; Q3= third quartile) for continuous data, and total number and percentage for binary data. * Mann-Whitney U Test for quantitative traits and Pearson $\chi 2$ test for binary traits, without considering any potential confounding.

Supplementary Table 2: Categories and number of FVL, PTM carriers and polygenic risk score in the study cohort separated by oral contraceptive use and for 1st and 10th polygenic risk score deciles. For each category, the total number of participants and the number of venous thromboembolism events (in parenthesis) are reported. Abbreviation: NA, not present.

Supplementary Table 3: Risk of venous thromboembolism according to the presence of Polygenic risk score, PTM, and FVL in all women and in never users. All models were adjusted for body mass index, year of birth, smoking status, pregnancy, Townsend deprivation index, and the first four principal components.

Supplementary Table 4: Area under the curve. The area under curve (AUC) calculated for the Base model (including age and genetic principal components), Base plus FVL and PTM, and Base plus Polygenic risk score. The AUC is based on a logistic regression model with the coefficients for age, FVL, PTM and Polygenic risk score estimated from the UKB data.





Comparison groups	HR (95% CI)	RPY
1st PRS, never OC, no carriers	Reference	-
10th PRS	6.35 (4.98 - 8.09)	2.42
	2.12 (1.81 – 2.49)	1.76
non-carriers of FVL - PTM	5.82 (4.48 - 7.91)	2.11
	1.94 (1.06 – 2.35)	1.5
FVL carriers	8.78 (6.12 - 12.6)	3.1
	3.12 (2.5 – 3.88)	2.32
PTM carriers	10.58 (7.48 – 14.97)	2.85
	3.64 (2.58 - 5.13)	1.76
FVL + PTM carriers	14.8 (9.28 - 23.6)	9.17
	4.93 (3.16 – 7.7)	2.88
Never OC, no carriers	Reference	-
FVL carriers	5.73 (5.31 - 6.17)	2.51
	2 (1.86 - 2.16)	1.9
PTM carriers	5.23 (4.67 – 5.87)	3.02
	1.76 (1.57 - 1.97)	1.35
FVL + PTM carriers	9 (6.07 - 13.34)	6.07
	3.39 (2.38 – 4.83)	3.52