Contribution of genetics and lifestyle to the risk of major cardiovascular and thromboembolic complications following COVID-19

Daniel Prieto-Alhambra (✉ daniel.prietoalhambra@ndoms.ox.ac.uk)
  1. Centre for Statistics in Medicine (CSM), Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences 2.Department of Medical Informatics, Erasmus University, Rotterdam
  https://orcid.org/0000-0002-3950-6346

Junqing Xie
  Centre for Statistics in Medicine and NIHR Biomedical Research Centre Oxford, NDORMS, University of Oxford  https://orcid.org/0000-0002-0040-0042

Yuliang Feng
  University of Oxford

Danielle Newby
  Centre for Statistics in Medicine and NIHR Biomedical Research Centre Oxford, NDORMS, University of Oxford

Bang Zheng
  London School of Hygiene and Tropical Medicine

Qi Feng
  Nuffield Department of Population Health, University of Oxford

Albert Prats-Uribe
  Centre for Statistics in Medicine, University of Oxford  https://orcid.org/0000-0003-1202-9153

Chunxiao Li
  Medical Research Council Epidemiology Unit, Institute of Metabolic Science, University of Cambridge
  School of Clinical Medicine, University of Cambridge

Nick Wareham
  MRC Epidemiology Unit  https://orcid.org/0000-0003-1422-2993

Roger Paredes
  Infectious Diseases Department & irsiCaixa AIDS Research Institute, Hospital Universitari Germans 13 Trias i Pujol

Article

Keywords:
Abstract
Clinical determinants for cardiovascular and thromboembolic (CVE) complications of COVID-19 are well-understood, but the roles of genetics and lifestyle remain unknown. We performed a prospective cohort study using UK Biobank, including 25,335 participants with confirmed SARS-CoV-2 infection between March 1, 2020, and September 3, 2021. Outcomes were hospital-diagnosed atrial fibrillation (AF), coronary artery disease (CAD), ischemic stroke (ISS), and venous thromboembolism (VTE) within 90 days post-infection. Heritable risk was represented by validated polygenic risk scores (PRSs). Lifestyle was defined by a composite of nine variables. We estimated adjusted hazard ratios (aHR) and confidence intervals (CI) using Cox proportional hazards models. In the COVID-19 acute phase, PRSs linearly predicted a higher risk of AF (aHR 1.52 per standard deviation increase, 95% CI 1.39 to 1.67), CAD (1.59, 1.40 to 1.81), and VTE (1.30, 1.11 to 1.53), but not ISS (0.92, 0.64 to 1.33). A healthy lifestyle was associated with a substantially lower risk of post-COVID-19 AF (0.70, 0.53 to 0.92), CAD (0.64, 0.44 to 0.91), and ISS (0.28, 0.12 to 0.64), but not VTE (0.82, 0.48 to 1.39), compared with an unhealthy lifestyle. No evidence for interactions between genetics and lifestyle was found. Our results demonstrated that population genetics and lifestyle considerably influence cardiovascular complications following COVID-19, with implications for future personalised thromboprophylaxis and healthy lifestyle campaigns to offset the elevated cardiovascular disease burden imposed by the ongoing pandemic.

Introduction
Cardiovascular disease is the leading cause of death globally. Recently, cardiovascular mortality and morbidity have risen further due to the direct and indirect consequences of the COVID-19 pandemic. It is expected that the repercussions and long-term sequelae of COVID-19 could further increase the cardiovascular burden to an unprecedented level.

At the individual level, preventing life-threatening cardiovascular and thromboembolic events (CVE) is crucial during the treatment of patients with COVID-19. However, a challenge remains to accurately identify individuals at sufficient risk to warrant intensive surveillance or targeted pharmacological prophylaxis. For instance, although prophylactic anticoagulation has been recommended for hospitalized patients with COVID-19, evidence for its use is vastly conflicting for more critical ICU patients and milder ambulatory patients with COVID-19.

General risk factors, including age, sex, and obesity, are recognized predictors of COVID-19 severity, including hospital admission and mechanical ventilation. However, although they are helpful in informing clinical practice, they are not specific to CVE risk, and commonly correlate with an elevated risk of bleeding. In contrast, polygenic risk scores (PRSs), a sum of genetic risk for a particular trait, have been proposed as a promising tool for precision medicine and early cardiovascular risk stratification. It is not yet known whether this previously well-captured genetic susceptibility to typical CVE will also predict the occurrence of COVID-19-related CVE during acute and post-acute illness phases.
In addition, effective public health interventions are urgently needed to reduce the population's cardiovascular burden, particularly in light of surging COVID-19 infections after the removal of most early restrictions (e.g., lockdown and social distance). The US Preventive Service Task Force updated its recommendations in 2022, promoting healthy behaviour counselling for all adults as a national strategy for primary cardiovascular prevention. However, all current clinical and public health guidelines lack insights into the potential role of healthy lifestyle modifications in alleviating COVID-19 cardiovascular complications, likely due to a paucity of evidence.

We aimed to assess the association between validated PRSs, lifestyle risk factors, and their interactions and the risk of major CVE within 90 days after COVID-19 diagnosis.

**Results**

**Population characteristics**

Table 1 shows the baseline characteristics of the whole UK Biobank eligible source population with available PRS (n = 408,395) and of participants with COVID-19 who were eligible for analyses (n = 25,335), overall and stratified by their genetic risk for AF. The average (SD) age of the COVID-19 cohort was 65.99 (8.53) years, of whom most were women (52.7%) and of White ethnicity (84.6%). The prevalence of the nine prespecified unhealthy lifestyle factors ranged from 10.8% for smoking status to 48.5% for low oily fish intake. In total, 8.6% of infected individuals lived an unhealthy lifestyle based on our composite measure (≥ 5 unhealthy lifestyle factors), a proportion generally consistent with that of the entire UK Biobank source population. All study covariates, except ethnicity, were distributed similarly across different genetic strata, illustrating independence between the selected PRS and lifestyle factors. The baseline characteristics stratified by other PRSs for CAD, ISS, and VTE were similar to those when stratified by the AF genetic score (see Supplementary Table 1–3).

After infection, 422 AF, 244 CAD, 29 ISS, and 135 VTE events occurred during the 90-day follow-up period, accounting for 1.67%, 0.96%, 0.12%, and 0.53%, of the COVID-19 cohort, respectively. The incidence rate was 6.12 per 1,000 person-years for ISS, 28.0 for VTE, 48.5 for CAD, and 86.9 for AF.

**Genetic risk and post-COVID-19-CVE**

All four PRSs followed a normal distribution (see Supplementary Fig. 1). A higher PRS for CAD, AF, or VTE was associated with an increased risk of post-COVID-19 CVE, with an HR per SD increase in the PRS of 1.52 (95% CI 1.39 to 1.67) for AF, 1.59 (1.40 to 1.81) for CAD, and 1.30 (1.11 to 1.53) for VTE (Table 2). These associations were significantly linear, with all P-values for the linear term < 0.001 and for the non-linear term > 0.05 (see Supplementary Fig. 1). There was no association between ISS-PRS and post-COVID-19 ISS (HR: 0.92 [0.64 to 1.33]).

Figure 1 depicts Kaplan-Meier cumulative incidence curves and shows that CVE increased dramatically within the first 15 to 30 days after infection. Early separation of risk was consistently observed across the
three genetic risk categories and continued to diverge over time. Compared to participants with low genetic risk, those with higher genetic risk had a significantly increased incidence of CVE post-COVID-19. The HR of high vs low genetic risk was 3.00 (2.16 to 4.15) for AF, 3.45 (2.22 to 5.34) for CAD, and 2.00 (1.15 to 3.38) for VTE.

The overall estimate of the association between each PRS and AF, CAD, and VTE was similar across a range of prespecified subgroups, despite the substantially varying baseline risks (Fig. 2). For example, the incidence rate of CAD among participants with COVID-19 aged 65 years or older was 78.03 per 1,000 person-years, almost 5-fold higher than those younger than 65 years (incidence rate: 19.64), but the HR estimate was similar (1.57 for those aged over 65 years vs 1.67 for those aged under 65 years, P-interaction = 0.61). Recent users of antithrombotics had substantially higher background risks of CAD and AF, with incidence rate of 264.99 per 1,000 person-years for users and 12.07 for non-users for CAD and incidence rates of 338.42 for users and 44.78 for non-users for AF. The association between PRS and post-COVID-19 CVE was attenuated, but remained significant (HR: 1.20 for those antithrombotic users vs 1.99 for those non-antithrombotic users for CAD, P-interaction < 0.01; 1.34 vs 1.63 for AF, P-interaction = 0.05). The genetic risk appeared to persist among fully vaccinated individuals (HR: 1.62 for breakthrough infection vs 1.50 for non-breakthrough infection for AF, P-interaction = 0.54; 1.23 vs 1.63 for CAD, P-interaction = 0.18; 1.51 vs 1.29 for VTE, P-interaction = 0.62).

Healthy lifestyle and COVID-19-related CVE

Compared to those with an unhealthy lifestyle, people adhering to healthier habits had a significantly reduced risk of AF (HR 0.70 [95% CI 0.53 to 0.92]), CAD (0.64 [0.44 to 0.91]), and ISS (0.28 [0.12 to 0.64]) within 90 days following COVID-19 infection. No association was found for post-COVID-19 VTE (0.82 [0.48 to 1.39]). With 8.6% of infected individuals having an unhealthy lifestyle, the prevented fraction of CVE outcomes for the COVID-19 population changing to a healthy lifestyle was 2.46% for AF, 2.95% for CAD, and 5.86% for ISS (Table 3). There was no significant interaction between genetic and lifestyle factors for any CVE outcomes (P-interaction > 0.05). However, there was an additive effect for the risk of AF and CAD, with an HR for high genetic risk and unhealthy lifestyle vs low genetic risk and healthy lifestyle of 5.09 [3.07 to 8.41] for AF and 5.00 [2.52 to 9.93] for CAD (Fig. 3).

Sensitivity analyses

The results of the sensitivity analysis were generally consistent with the primary analysis. Noticeably, the magnitude of PRS association was reduced for the sensitivity analysis of incident events (HR: AF 1.33 [1.11 to 1.59]; CAD 1.19 [0.87 to 1.63]) and hospital admission-related events (AF 1.31 [0.89 to 1.92]; CAD 1.21 [0.71 to 2.08]), compared to those estimates regarding any events diagnosed after infection. However, the genetic risk was even more pronounced when restricting to admission-related VTE outcome [HR 1.52 (1.14 to 2.02)].

As expected, the positive control experiment shown that PRSs were prospectively associated with the corresponding CVE event among participants without COVID-19. And the negative control experiment
found that the post-COVID-19 risk of diabetes was not associated with an increase of most of PRSs, except for AF-PRS. (Supplementary Table 4)

**Discussion**

**Key findings**

This is the first prospective cohort study of the genetic- and lifestyle-related determinants of post-COVID-19 CVE complications. We found that a higher genetic risk based on validated PRSs for CVE events was linearly associated with an increased incidence of post-COVID-19 CVE. Participants with the top 20% of PRSs had a 3-fold, 3.5-fold, and 2-fold excess risk of AF, CAD, and VTE, respectively, compared with those with the lowest 20% of PRSs. No significant association was observed for post-COVID-19 ISS. The identified genetic predisposition was persistent in subgroups of participants who were already at very high risk of CVE and were receiving antithrombotic therapy before infection and in those with a breakthrough infection after full vaccination (2 doses).

We demonstrated that a healthy lifestyle was associated with substantially lower risk of arterial events (-30% AF, -36% CAD, and −72% ISS) among participants with COVID-19. A sum of over 10% of CVE (2.46% AF, 2.95% CAD, 5.86% VTE) could have been prevented if all individuals with an unhealthy lifestyle had adopted a healthier one. We did not observe a significantly different association between healthy lifestyle and CVE risk across different genetic risk strata. Instead, high genetic risk and unhealthy behaviour had an additive effect on the increased risk of post-COVID-19 AF and CAD complications.

**Findings in context**

Our study is the first to investigate genetic determinants for COVID-19-related CVE. It differs from previous studies of PRS for adult-onset cardiovascular diseases\(^1^5^-^1^8\) as we targeted people with COVID-19 rather than the general healthy population and we evaluated the risk of acute CVE triggered by SARS-CoV-2 rather than the chronic (≥10 years) or lifetime disease risk. The strongest associations (HR per one SD increase) found in previous studies were 2.33 for AF\(^1^6\), 1.86 for CAD\(^1^6\), 1.26 for ISS\(^1^9\), and 1.27 for VTE\(^1^8\), compared with 1.52 for AF, 1.59 for CAD, 0.96 for ISS, and 1.30 for VTE found here. Our observed associations demonstrate that population genetic variations are an important contributor to developing CVE following COVID-19 and highlight underlying genetic interconnections between chronic and post-COVID-19 cardiovascular complications. However, the magnitude of gene association was reduced for all arterial-related events (CAD, AF, and ISS), suggesting that distinctive pathogenic mechanisms may be involved when arterial disorders develop during COVID-19 infection\(^2^,^2^0\), such as the virus directly mediating heart injury by entering cardiomyocytes.\(^2^1^,^2^2\) Importantly, the polygenic risk for conventional VTE was remarkably retained for COVID-19-related VTE, which echo the finding in our previous study showing that the monogenic such as factor V Leiden mutation also consistently predisposed post-COVID-19 VTE complications.\(^6\)
The role of genetics in variable COVID-19 representations is not yet well-understood. Although a large genome-wide association study (GWAS) was performed for COVID-19, it focused on SARS-COV-2-induced critical respiratory complications.\textsuperscript{23,24} We leveraged well-developed PRSs and proved that human polygenic variations affect CVE manifestations after COVID-19, beyond the severity of COVID-19 disease. There is also a lack of evidence for the potential beneficial effects of healthy lifestyles on reducing the cardiovascular disease burden in COVID-19. Although many studies have reported that lifestyle factors affected chronic cardiovascular conditions independently of individuals’ genetic background before the pandemic,\textsuperscript{15,27} little is known whether this remains the case in COVID-19-related cardiovascular complications. Some studies have observed lower risks of severe COVID-19 among people adopting more favourable behaviours.\textsuperscript{25,26} However, these studies concentrated on general health utilization outcomes such as hospital or ICU admissions, therefore limiting the specificity of the observed associations, particularly life-threatening CVE.

Clinical and public health implications

Our findings have implications for clinical responses and public health preparedness against the ongoing COVID-19 pandemic. At the individual level, compared with known general risk factors such as demographic characteristics (e.g., age and sex) and clinical risk factors (e.g., obesity and hypertension), genetic factors might inform more tailored treatment choices to prevent specific COVID-19 complications. For instance, disease-specific PRSs could help doctors identify patients with high genetic risk for arterial thrombosis who would benefit from platelet inhibitors or identify those with high genetic risk for venous thrombosis such as VTE and prioritise them for coagulation cascade suppression therapy.\textsuperscript{3,28} Such specific PRSs cannot be achieved using traditional clinical factors alone, such as age, as they were associated with both high VTE risk and high adverse events related to pharmaceutical therapy such as antithrombotics.

Over the past 10 to 15 years, global interest, efforts, and controversies have surrounded PRSs’ clinical utility for the primary prevention of non-communicable cardiovascular diseases, such as predicting 10-year risk in the general population.\textsuperscript{9,11,29} The potential of a PRS could be magnified in patients with COVID-19 as they have a substantially increased CVE risk, particularly during the initial illness. If a PRS was calculated for everyone at birth and held as part of their health records,\textsuperscript{30} it could have been used as easily as demographic determinants like age and sex to refine existing approaches to defining subgroups who are particularly vulnerable to COVID-19, possibly providing more timely, personalised shielding advice. Even a small or modest improvement in stratification accuracy might lead to a sizeable population changing their COVID-19 vulnerability category.

Although the genetic risk for post-COVID-19 CVE is inherited, our study showed that acquired healthy behaviours could offset this risk. The US Preventive Services Task Force updated its guidelines in 2022, recommending behavioural counselling for cardiovascular disease prevention for all adults aged 18 years and older.\textsuperscript{12,14} Our data support this recommendation by showing that a composite healthy lifestyle can
also benefit acute CVE outcomes after COVID-19, regardless of genetic risk. This benefit should be emphasized in future lifestyle intervention programmes during the ongoing pandemic.

Strengths and limitations

Our study benefitted from a large population-based cohort, standardized genotyping, quality-controlled genetic data, powered and validated PRS estimates, well-defined measurements of a range of lifestyle factors, PCR-confirmed COVID-19 infection, and reliable and complete linkages to cardiovascular disease outcomes, which together enable these novel findings.

However, several study limitations should be considered. Our PRS was initially built to quantify polygenic risk for any adult-onset cardiovascular disease. It may not reflect the maximum possible genetic contribution to COVID-19 cardiovascular complications, especially given the likelihood of distinct pathological mechanisms involving virus-induced cardiovascular events. Future GWAS studies explicitly designed for COVID-19-related CVE could inform the development of a bespoke PRS and improve predictive performance.

Observational studies that use routinely collected data to ascertain disease outcomes may record overdiagnoses for COVID-19 patients. The ICD records of some clinical events, such as hypertension or diabetes, immediately after COVID-19 infection could be duplicate records of historical conditions instead of a new or activated disease status. However, all of the cardiovascular disease subtypes except for CAD used as outcomes in this study appeared to be temporary and potentially life-threatening. They are unlikely to be coded for without justification in actual clinical practice. Our sensitivity analyses using only incident or hospital-admission-specific CVE also produced findings consistent with the main analysis, precluding this concern.

Demonstrating statistical significance does not guarantee that the PRS is able to provide additional predictive information up on the existing clinical factors only based cardiovascular models, as previous studies have frequently found little agreement between statistical association and predictive performance. More modelling research is urgently needed to fill this evidence gap in the contexts of COVID-19.

We used lifestyle behaviour data collected 10 years ago as a surrogate for current lifestyle habits at the time of infection, which is likely subject to misclassification and may have biased any genuine associations toward the null. Reassuringly, all participants at the time of recruitment were middle-aged or older adults whose lifestyle habits should have been well established, suggesting that their habits are likely to have remained consistent over years between recruitment and infection.

Participants in UK Biobank represent a generally healthier population than the general population of the UK and are mostly of European ancestry, which may limit our findings’ generalisability beyond this population.
Conclusions

Individuals’ genetic predisposition to cardiovascular disease, in the form of a PRS, was associated with a short-term risk of CAD, AF, and VTE complications after COVID-19, but not ISS. Additionally, the incidence of all studied arterial events was substantially lower among those adhering to a healthy lifestyle, independent of their genetic risk. These findings demonstrate that a person's genetic background underpins how COVID-19 affects their cardiovascular system, but suggest that healthy lifestyle interventions can alleviate the elevated population cardiovascular disease burden during the ongoing pandemic.

Online Methods

Data sources and COVID-19 population

UK Biobank is a large-scale, population-based prospective cohort of over 500,000 individuals aged 40 to 69 years at recruitment between 2006 and 2010 from across the United Kingdom.33 Extensive information was collected through questionnaires and physical measurements at 22 assessment centres. Affymetrix performed genotype calling based on two closely related purpose-designed arrays.34 Follow-up health outcomes were identified through linkage to various electronic health records, covering national primary and secondary care, disease and mortality registries.35 The reliability, accuracy and completeness of capturing disease conditions using this approach, especially cardiovascular diseases, have been validated in previous studies.36,37 To enable urgent research into COVID-19, additional data from Public Health England’s Second Generation Surveillance System has recently been linked to all UK Biobank participants using a bespoke algorithm.38 This information includes whether their SARS-CoV-2 infection status has been confirmed with polymerase chain reaction (PCR) testing.

In this study, we enrolled a cohort of participants from UK Biobank with COVID-19 from England who had a positive PCR testing result between March 1, 2020, and September 30, 2021. For one sensitivity analysis, we used all participants in the full UK Biobank population who were alive on March 1, 2020. Participants with missing age, sex, BMI, socioeconomic status, lifestyle factors, and genotyping information at baseline were excluded.

Polygenic risk scores

In May 2022, UK Biobank released the first version of polygenic risk scores for 28 diseases, proving their predictive ability to outperform currently published PRS (see Supplementary Methods).39 Two sets of PRS, standard and enhanced, were developed and validated. The standard score was developed based only on non-UK-Biobank populations and then calculated for all individuals in UK Biobank. The development of the enhanced score used some UK Biobank participants, with the score then computed for those remaining to avoid the risk of overfitting.40 In this study, we used the standard PRS for coronary artery disease (CAD), atrial fibrillation (AF), ischaemic stroke (ISS), and venous thromboembolic disease
(VTE) in the primary analysis and the enhanced PRS in one of the sensitivity analyses. The continuous PRS used in this study was also categorized as high-risk (fifth quintile), intermediate-risk (2–4 quintiles), and low-risk (lowest quintile) to enable comparisons with previous studies.\textsuperscript{27,41}

Definition of healthy lifestyles

We defined a composite healthy lifestyle index by combining data for nine lifestyle components based on a published study\textsuperscript{42}: smoking status, alcohol drinking, physical activity, television viewing time, sleep duration, intake of fruit and vegetable, intake of oily fish, intake of red meat, and intake of processed meat. Each lifestyle factor was assigned 0 points if healthy and 1 point if unhealthy. For example, participants with “alcohol drinking daily or almost daily” scored 1 point for the alcohol drinking element. We then summed all lifestyle factors, with a score of 0–4 classified as combined healthy lifestyle and a score of 5–9 classified as unhealthy lifestyle. More details on defining each beneficial lifestyle factor are provided in Supplemental Methods.

Post-COVID-19 CVE

Among participants with COVID-19, we defined the first infection as the index date and followed up for 90 days. We studied four major CVEs (AF, CAD, ISS, and VTE) that were frequently reported as COVID-19-related cardiovascular complications, through linkage to hospital admissions data. International Classification of Diseases 10th Revision (ICD-10) codes were used to capture clinical outcomes and are presented in Supplemental Methods. These ICD-10 codes were the same as those initially used for PRS development to minimize the impact of variation in disease phenotypes between our and previous studies.\textsuperscript{40} Data in this study were censored on September 30, 2021.

Statistical Analyses

We used the Cox proportional hazard (PH) model to assess the associations between each PRS and its corresponding CVE outcome. PH assumptions were checked based on Schoenfeld residuals and were satisfied. The hazard ratio (HR) and 95% confidence interval (CI) for a continuous PRS (per 1 standard deviation [SD] increase) were derived by adjusting for age, sex, education level (mapped to the international standard for classification of education, see Supplemental Methods), index of multiple deprivations (IMD, a continuous summary deprivation measurement used in England that contains crime, education, employment, health, housing, income, and living environment),\textsuperscript{43} genotyping batch, and the first ten principal components of genetic ancestry. To avoid overadjustment, we did not adjust for previous CVE conditions (which, for example, can mediate the genetic effects on post-COVID-19 CVE outcomes).

We calculated the PRS-CVE association and stratified it by clinically relevant features, including age (\(\geq 65\) years or < 65 years), sex, body mass index (\(\geq 30\) or < 30), ethnicity (White or other ethnic groups), recent antithrombotic medication (yes or no), setting for a positive PRC test (inpatient or outpatient), and SARS-COV-2 infection type (breakthrough infection after two-dose vaccination or non-breakthrough infection).
Multiplicative interactions between PRS and the stratification variables were tested, and $P$-values are reported. Restricted cubic splines were used to examine possible nonlinear associations for the continuous PRS. Categorical genetic risk was analyzed separately and depicted using the Kaplan–Meier method.

We used the same multivariable Cox regression model for the lifestyle factor in the overall cohort and across the PRS categories. We calculated the prevented fraction for the population (PFP) according to the formula: $\text{PFP} = \text{Pe} \times (1 - \text{HR})$, where Pe was the prevalence of the unhealthy lifestyle. We estimated the combined effect of genetics and lifestyle (six categories with low genetic risk and healthy lifestyle as reference group) on those CVE outcomes that were statistically associated with both genetic and lifestyle factors.

We performed several sensitivity analyses for genetic risk. First, we examined the associations among the sub-cohort of participants with an enhanced PRS recorded in UK Biobank. Second, we studied incident CVE by excluding participants with a historical study outcome before the index date. Third, we defined the CVE outcomes based on the primary diagnosis alone, which reflects the main cause of hospital admission. Fourth, we performed a negative control outcome experiment using the association between PRS and type 2 diabetes among the COVID-19 cohort and a positive control outcome experiment on the association between PRS and CVE among UK Biobank participants without evidence of COVID-19. The positive and negative control experiments were designed to detect residual confounding and any spurious bias related to study design, cohort building, analytic approach, or outcome ascertainment.

The analyses were performed using R software version 4.1.2. All statistical tests were 2-sided. A 95% CI that did not contain unity was considered statistically significant.

### Tables

**Table 1: Characteristics of the whole UK Biobank population and of participants with COVID-19, stratified by genetic risk of atrial fibrillation.**

1
<table>
<thead>
<tr>
<th>Participant characteristics</th>
<th>All eligible UK Biobank participants</th>
<th>Participants with COVID-19</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>408395</td>
<td>25335</td>
</tr>
<tr>
<td>Mean age, year (SD)</td>
<td>68.10 (8.10)</td>
<td>65.99 (8.53)</td>
</tr>
<tr>
<td>sex, No. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>225451 (55.2)</td>
<td>13339 (52.7)</td>
</tr>
<tr>
<td>Male</td>
<td>182944 (44.8)</td>
<td>11996 (47.3)</td>
</tr>
<tr>
<td>Ethnicity, No. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>382169 (93.6)</td>
<td>3893 (15.4)</td>
</tr>
<tr>
<td>Other ethnic groups</td>
<td>26226 (6.4)</td>
<td>21442 (84.6)</td>
</tr>
<tr>
<td>Obesity, No. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI &lt; 30</td>
<td>311205 (76.2)</td>
<td>17984 (71.0)</td>
</tr>
<tr>
<td>BMI ≥ 30</td>
<td>97190 (23.8)</td>
<td>7351 (29.0)</td>
</tr>
<tr>
<td>Education categories, No. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>73987 (18.1)</td>
<td>4863 (19.2)</td>
</tr>
<tr>
<td>4</td>
<td>110969 (27.2)</td>
<td>7999 (31.6)</td>
</tr>
<tr>
<td>5</td>
<td>45409 (11.1)</td>
<td>2613 (10.3)</td>
</tr>
<tr>
<td>6</td>
<td>20765 (5.1)</td>
<td>1194 (4.7)</td>
</tr>
<tr>
<td>7</td>
<td>157265 (38.5)</td>
<td>8666 (34.2)</td>
</tr>
<tr>
<td>Index of multiple deprivation, mean (SD)</td>
<td>17.51 (13.87)</td>
<td>20.62 (14.85)</td>
</tr>
<tr>
<td>Individual lifestyle factors, No. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never or past</td>
<td>368471 (90.2)</td>
<td>22590 (89.2)</td>
</tr>
<tr>
<td>Current</td>
<td>39924 (9.8)</td>
<td>2745 (10.8)</td>
</tr>
<tr>
<td>Alcohol intake</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 4 times week</td>
<td>325341 (79.7)</td>
<td>21062 (83.1)</td>
</tr>
<tr>
<td>Daily or almost daily</td>
<td>83054 (20.3)</td>
<td>4273 (16.9)</td>
</tr>
<tr>
<td>Physical activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥150 min per week of MIPA</td>
<td>345208 (84.5)</td>
<td>21194 (83.7)</td>
</tr>
<tr>
<td>or ≥75 min per week of VIPA</td>
<td></td>
<td>(83.5)</td>
</tr>
<tr>
<td>&lt;150 min per week of MIPA &amp; &lt;75 min per week of VIPA</td>
<td>63187 (15.5)</td>
<td>4141 (16.3)</td>
</tr>
<tr>
<td>Television viewing time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 4 h/day</td>
<td>293476 (71.9)</td>
<td>17880 (70.6)</td>
</tr>
<tr>
<td>≥ 4 h/day</td>
<td>114919 (28.1)</td>
<td>7455 (29.4)</td>
</tr>
<tr>
<td>Sleep duration</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Page 12/21
Table 2: Associations between polygenic risk scores (PRSs) and 90-day risk of cardiovascular and thromboembolic events after COVID-19.
<table>
<thead>
<tr>
<th>Outcomes</th>
<th>No. of events</th>
<th>Incidence (per 1,000 yrs)</th>
<th>Adjusted hazard ratio* (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Continuous PRS (per SD increase)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 (ref)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 (ref)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 (ref)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 (ref)</td>
</tr>
<tr>
<td>Atrial fibrillation</td>
<td>422</td>
<td>86.9</td>
<td>1.52 (1.39 to 1.67)</td>
</tr>
<tr>
<td>Coronary artery disease</td>
<td>235</td>
<td>48.5</td>
<td>1.59 (1.40 to 1.81)</td>
</tr>
<tr>
<td>Ischaemic stroke</td>
<td>29</td>
<td>6.12</td>
<td>0.92 (0.64 to 1.33)</td>
</tr>
<tr>
<td>Venous thromboembolism</td>
<td>135</td>
<td>28.0</td>
<td>1.30 (1.11 to 1.53)</td>
</tr>
</tbody>
</table>

Pys: person-years, Ref: reference, SD: standard deviation

Table 3: Hazard ratios for COVID-19-related cardiovascular and thromboembolic events and prevented fraction of the population associated with a combined healthy lifestyle.\(^1\)

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Hazard ratio (healthy lifestyle vs unhealthy lifestyle)</th>
<th>Prevented fraction of the population (%)(^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All participants with COVID-19</td>
<td>Low genetic risk</td>
</tr>
<tr>
<td>Atrial fibrillation</td>
<td>0.70 (0.53 to 0.92)</td>
<td>1.28 (0.49 to 3.30)</td>
</tr>
<tr>
<td>Coronary artery disease</td>
<td>0.64 (0.44 to 0.91)</td>
<td>1.03 (0.30 to 3.56)</td>
</tr>
<tr>
<td>Ischaemic stroke</td>
<td>0.28 (0.12 to 0.64)</td>
<td>0.25 (0.02 to 3.05)</td>
</tr>
<tr>
<td>Venous thromboembolism</td>
<td>0.82 (0.48 to 1.39)</td>
<td>0.43 (0.14 to 1.33)</td>
</tr>
</tbody>
</table>

\(^1\) 8.1% of people in the COVID-19 cohort recorded an unhealthy lifestyle. \(^2\)Prevented fraction for the population was calculated according to the formula: PFP = Pe (1-HR), where Pe was the prevalence of an unhealthy lifestyle\(^45\).

**Declarations**

Declaration of interests
DPA's department has received grant/s from Amgen, Chiesi-Taylor, Lilly, Janssen, Novartis, and UCB Biopharma. His research group has received consultancy fees from Astra Zeneca and UCB Biopharma. Amgen, Astellas, Janssen, Synapse Management Partners and UCB Biopharma have funded or supported training programmes organised by DPA's department. Roger Paredes has participated in advisory boards for Gilead, MSD, ViiV Healthcare, Theratechnologies and Lilly. His institution has received research support from Gilead, MSD, and ViiV Healthcare.

**Ethical Approval**

All participants provided written informed consent at the UKBB cohort recruitment. This study received ethical approval from the UKBB Ethics Advisory Committee (EAC) under application 65397.

**Data Sharing**

Bonafide researchers can apply to use the UK Biobank dataset by registering and applying at http://ukbiobank.ac.uk/register-apply/. Any additional summary data generated and/or analysed during the current study are available from the corresponding author upon reasonable request.

**References**


Figures
Figure 1

**Cumulative incidence of COVID-19-related cardiovascular events by the three genetic risk categories**

The upper limit of the y-axis was 0.04 for atrial fibrillation and coronary artery disease and was 0.02 for ischaemic stroke and venous thromboembolism.
Figure 2


Figure 3

Combined effect of genetic and lifestyle factors on COVID-19-related atrial fibrillation and coronary arterial disease.
HR: hazard ratio. We only examined joint effects for cardiovascular events that had significant associations with both genetic and lifestyle factors in the primary analysis.

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementary.docx