



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



# Venous or arterial thrombosis and deaths among COVID-19 cases: a European network cohort study



Edward Burn, Talita Duarte-Salles, Sergio Fernandez-Bertolin, Carlen Reyes, Kristin Kostka, Antonella Delmestri, Peter Rijnbeek, Katia Verhamme, Daniel Prieto-Alhambra

## Summary

**Background** There are few data on the incidence of thrombosis among COVID-19 cases, with most research concentrated on hospitalised patients. We aimed to estimate the incidence of venous thromboembolism, arterial thromboembolism, and death among COVID-19 cases and to assess the impact of these events on the risks of hospitalisation and death.

**Methods** We conducted a distributed network cohort study using primary care records from the Netherlands, Italy, Spain, and the UK, and outpatient specialist records from Germany. The Spanish database was linked to hospital admissions. Participants were followed up from the date of a diagnosis of COVID-19 or positive RT-PCR test for SARS-CoV-2 (index date) for 90 days. The primary study outcomes were venous thromboembolic events, arterial thromboembolic events, and death, all over the 90 days from the index date. We estimated cumulative incidences for the study outcomes. Multistate models were used to calculate adjusted hazard ratios (HRs) for the association between venous thromboembolism or arterial thromboembolism occurrence and risks of hospitalisation or COVID-19 fatality.

**Findings** Overall, 909 473 COVID-19 cases and 32 329 patients hospitalised with COVID-19 on or after Sept 1, 2020, were studied. The latest index dates across the databases ranged from Jan 30, 2021, to July 31, 2021. Cumulative 90-day incidence of venous thromboembolism ranged from 0·2% to 0·8% among COVID-19 cases, and up to 4·5% for those hospitalised. For arterial thromboembolism, estimates ranged from 0·1% to 0·8% among COVID-19 cases, increasing to 3·1% among those hospitalised. Case fatality ranged from 1·1% to 2·0% among patients with COVID-19, rising to 14·6% for hospitalised patients. The occurrence of venous thromboembolism in patients with COVID-19 was associated with an increased risk of death (adjusted HRs 4·42 [3·07–6·36] for those not hospitalised and 1·63 [1·39–1·90] for those hospitalised), as was the occurrence of arterial thromboembolism (3·16 [2·65–3·75] and 1·93 [1·57–2·37]).

**Interpretation** Risks of venous thromboembolism and arterial thromboembolism were up to 1% among COVID-19 cases, and increased with age, among males, and in those who were hospitalised. Their occurrence was associated with excess mortality, underlying the importance of developing effective treatment strategies that reduce their frequency.

**Funding** European Medicines Agency.

**Copyright** © 2022 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

## Introduction

COVID-19 can result in thrombotic disease, both in the venous and arterial circulations, due to excessive inflammation, platelet activation, endothelial dysfunction, and stasis.<sup>1</sup> Numerous studies have assessed risks of venous thromboembolism among patients hospitalised with COVID-19, with recent systematic reviews and meta-analyses identifying more than 100 such studies.<sup>2,3</sup> However, most of these studies were relatively small (with few including more than 1000 patients) and varied substantially in setting, design, and outcome definitions. Consequently, there has been substantial heterogeneity in results across previous studies. Across studies that did not use specific ultrasound screening, the prevalence of venous thromboembolic events among patients hospitalised with COVID-19 ranged from 0% to 37%.<sup>3</sup> Similar heterogeneity was also seen for arterial thromboembolism, with estimates varying from 1% to 18% for hospitalised patients.<sup>2</sup>

To date, studies of thromboembolism in patients with COVID-19 have focused almost exclusively on events among hospitalised patients, with the incidence of thromboembolism among cases of COVID-19 in general not yet known. The effect of thrombosis on prognosis and health outcomes is also yet to be described in detail. Although the prevalence of thromboembolism has been seen to be higher among patients admitted to intensive care units (ICUs),<sup>2,3</sup> the timing of these events during the disease course and their association with worse outcomes have not been fully detailed.

Patient-related factors, such as older age, being male, hypertension, diabetes, and obesity, have all been associated with an increased risk of hospitalisation and death in COVID-19.<sup>4–10</sup> Many of these factors have also previously been shown to predispose individuals to thromboembolic events in the general population.<sup>11,12</sup> The

Lancet Infect Dis 2022;  
22: 1142–52

Published Online

May 13, 2022

[https://doi.org/10.1016/S1473-3099\(22\)00223-7](https://doi.org/10.1016/S1473-3099(22)00223-7)

See [Comment](#) page 1093

Fundació Institut Universitari per a la Recerca a l'Atenció Primària de Salut Jordi Gol i Gurina, Barcelona, Spain (E Burn DPhil, T Duarte-Salles PhD, S Fernandez-Bertolin MSc, C Reyes PhD); Centre for

Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK (E Burn, A Delmestri PhD, Prof D Prieto-Alhambra PhD); Real World Solutions, IQVIA, Cambridge, MA, USA (K Kostka MPH); The Observational Health Data

Sciences and Informatics Center at The Roux Institute, Northeastern University, Portland, ME, USA (K Kostka); Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, Netherlands (Prof P Rijnbeek PhD, Prof D Prieto-Alhambra, K Verhamme PhD)

Correspondence to: Prof Daniel Prieto-Alhambra, Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford OX3 7LD, UK  
[daniel.prietoalhambra@ndorms.ox.ac.uk](mailto:daniel.prietoalhambra@ndorms.ox.ac.uk)

## Research in context

### Evidence before this study

Several systematic reviews have summarised the existing evidence base on thrombosis in COVID-19, and we did further searches of PubMed, with no language restrictions, up to Feb 1, 2022, to identify additional recently published studies. Although more than 100 studies have been published, most have focused on patients hospitalised with COVID-19. Across these, the prevalence of venous thromboembolism has been estimated to be around 9%, whereas the prevalence of arterial thromboembolism was 4%, although specific estimates varied substantially across studies. To date, little research has been published on thrombosis among outpatient COVID-19 cases.

### Added value of this study

In this study we bring together data from five European countries to study venous thromboembolism, arterial

thromboembolism, and death among COVID-19 cases.

Over 900 000 COVID-19 cases were included in the study, allowing for a detailed summary of the occurrence of these events of interest. For one data source, which had patient-level linkage to hospitalisation data, the risks in patients hospitalised with COVID-19 were also characterised.

### Implications of all the available evidence

Risks of venous and arterial thromboembolism were close to 1% for COVID-19 cases and 4% in hospitalised patients.

The occurrence of these events was associated with worse outcomes for those affected. These findings underline the widespread negative consequences associated with COVID-19 and emphasise the importance of developing effective treatment strategies that reduce the frequency of thromboembolism in patients with COVID-19.

associations between patient characteristics and risks of thromboembolic events among patients with COVID-19 have yet to be elucidated.

Study cohorts derived from routinely collected health-care data can be used to further describe thromboembolism in COVID-19, particularly when the breadth of data capture is large enough to include both outpatient and inpatient COVID-19 cases and to have sufficient sample sizes to describe the associations between patient factors and outcomes. In this study, we bring together data from across Europe to summarise the incidence of venous and arterial thromboembolism and death among people with COVID-19, describe the association between patient factors and risks of such events, and assess the association between venous thromboembolism and arterial thromboembolism and worse health outcomes for the patients affected.

## Methods

### Study design and participants

We conducted a network cohort study using routinely collected health-care data from across Europe. All datasets were mapped to the Observational Medical Outcomes Partnership common data model and the study was run in a distributed manner, with common analytical code run by each site and aggregated results returned without the need to share patient-level data between sites.<sup>13–15</sup>

Routinely collected health-care data from Germany, Italy, the Netherlands, Spain, and the UK informed the analyses. IQVIA Disease Analyzer (DA) Germany captures data from patient management software used by general practitioners and specialists practicing in ambulatory care settings, while IQVIA Longitudinal Patient Database (LPD) Italy includes anonymised patient records collected from software used by general practitioners. The Integrated Primary Care Information

(IPCI) database collects data from electronic health-care records of patients registered with general practitioners throughout the Netherlands.<sup>16</sup> The Clinical Practice Research Datalink (CPRD) Aurum database contains data contributed by general practitioners from the UK.<sup>17</sup> The Information System for Research in Primary Care (SIDIAP) is a primary care records database that covers approximately 80% of the population of Catalonia, Spain. SIDIAP was linked to Conjunto Mínimo de Datos Básicos al Alta Hospitalaria (CMBD-AH), which includes diagnosis and procedures registered during hospital admissions.<sup>18</sup>

The feasibility of running the study was also assessed for IQVIA LPD France, a computerised network of physicians, including general practitioners who contribute to a centralised database of anonymised patient electronic medical records.<sup>19</sup> At the time of assessing feasibility, this database did not capture death and had only partial capture of arterial thromboembolism and venous thromboembolism and therefore this database was omitted.

Two cohorts were defined for the primary analyses. First, a cohort of people who were diagnosed with COVID-19 or had a positive RT-PCR test result for SARS-CoV-2 were identified. The first occurrence of either was used as the index date for a given person. The second cohort was people hospitalised with COVID-19, available only in SIDIAP CMBD-AH data. These individuals had a hospitalisation with a COVID-19 diagnosis or PCR-positive test between 21 days before or up to 3 days after their hospital admission. For this cohort, date of hospital admission was taken as the index date. For the primary analyses, individuals were also required to have at least 1 year of observation time available before their index date. Cohorts were also identified for sensitivity analysis: (1) solely based on clinical diagnoses; (2) solely based on clinical diagnoses

For Information System for Research in Primary Care see <https://www.sidiap.org>

For the Integrated Primary Care Information database see <https://www.ipci.nl>

and with a broader algorithm, including less specific codes for identifying COVID-19 diagnoses; (3) solely based on RT-PCR testing; and (4) solely based on SARS-CoV-2 testing, but with antigen tests included as well as RT-PCR tests. All the aforementioned cohorts were also identified without the requirement for having a 1 year observation time before study enrolment.

Because of the particular heterogeneity in testing and diagnostic practices and data capture during the first wave of COVID-19 in Europe, here we report results for individuals with an index date on or after Sept 1, 2020. Cut-off dates for inclusion were in early 2021 and varied between the databases. Results for March 1, 2020, onwards and for March 1 to Aug 31, 2020 are reported in the appendix (pp 21–23, 25).

The protocol for this research was reviewed by the European Medicines Agency and approved by

the Independent Scientific Advisory Committee for the Medicines and Healthcare products Regulatory Agency Database Research (protocol number 20\_000211), the Institut d'Investigacio en Atencio Primaria Jordi Go Clinical Research Ethics Committee (project code: 21/007-PCV), and the IPCI governance board (application number 3/2021). Some databases used (IQVIA LPD Italy, IQVIA LPD France, and IQVIA DA Germany) in these analyses are commercially available, syndicated data assets that are licensed by contributing authors for observational research. These assets are de-identified, commercially available data products that could be purchased and licensed by any researcher. As these data are deemed commercial assets, there is no institutional review board applicable to the usage and dissemination of these result sets or required registration of the protocol with additional ethics oversight. Compliance with data

See Online for appendix

	COVID-19 diagnosis or PCR-positive test result					Hospitalised with COVID-19 (SIDIAP CMBD-AH, Spain; N=32 329)
	CPRD Aurum (UK; N=415 369)	IQVIA DA Germany (N=38 657)	IPCI (Netherlands; N=38 847)	IQVIA LPD Italy (N=25 759)	SIDIAP CMBD-AH (Spain; N=390 841)	
Earliest index date	Sept 1, 2020	Sept 1, 2020	Sept 1, 2020	Sept 1, 2020	Sept 1, 2020	Sept 1, 2020
Latest index date	Jan 30, 2021	March 31, 2021	March 3, 2021	July 31, 2021	June 22, 2021	June 22, 2021
Age, years	43 (30–56)	47 (31–60)	46 (28–59)	52 (38–65)	42 (23–56)	67 (53–79)
<20	14 264 (3.4%)	3977 (10.3%)	5118 (13.2%)	1108 (4.3%)	79 566 (20.4%)	349 (1.1%)
20–29	84 262 (20.3%)	5031 (13.0%)	5414 (13.9%)	2841 (11.0%)	48 488 (12.4%)	859 (2.7%)
30–39	83 089 (20.0%)	5784 (15.0%)	4902 (12.6%)	3160 (12.3%)	53 406 (13.7%)	1761 (5.4%)
40–49	79 232 (19.1%)	6086 (15.7%)	6497 (16.7%)	4524 (17.6%)	70 626 (18.1%)	3525 (10.9%)
50–59	78 650 (18.9%)	7637 (19.8%)	7766 (20.0%)	5511 (21.4%)	56 855 (14.5%)	5389 (16.7%)
60–69	41 324 (9.9%)	4507 (11.7%)	4724 (12.2%)	3756 (14.6%)	34 748 (8.9%)	6028 (18.6%)
70–79	18 367 (4.4%)	2320 (6.0%)	2748 (7.1%)	2638 (10.2%)	23 348 (6.0%)	6521 (20.2%)
≥80	16 181 (3.9%)	3315 (8.6%)	1678 (4.3%)	2221 (8.6%)	23 804 (6.1%)	7897 (24.4%)
Sex						
Male	186 872 (45.0%)	17584 (45.5%)	17497 (45.0%)	11 328 (44.0%)	186 912 (47.8%)	17790 (55.0%)
Female	228 497 (55.0%)	21073 (54.5%)	21350 (55.0%)	14 431 (56.0%)	203 929 (52.2%)	14 539 (45.0%)
Smoker	99 878 (24.0%)	1305 (3.4%)	4001 (10.3%)	2771 (10.8%)	68 504 (17.5%)	6891 (21.3%)
Years of previous observation time	11.9 (4.6–22.5)	6.9 (3.8–12.4)	6.8 (4.8–9.6)	9.1 (6.2–9.4)	14.8 (13.2–15.1)	15.0 (14.8–15.1)
Comorbidities						
Autoimmune disease	6925 (1.7%)	2570 (6.6%)	416 (1.1%)	1403 (5.4%)	6269 (1.6%)	1105 (3.4%)
Antiphospholipid syndrome	202 (<0.1%)	0	0	0	190 (<0.1%)	27 (0.1%)
Thrombophilia	484 (0.1%)	81 (0.2%)	0 (0.0%)	0 (0.0%)	422 (0.1%)	71 (0.2%)
Asthma	47 723 (11.5%)	4815 (12.5%)	2755 (7.1%)	2823 (11.0%)	24 202 (6.2%)	2309 (7.1%)
Atrial fibrillation	7268 (1.7%)	960 (2.5%)	841 (2.2%)	941 (3.7%)	10 657 (2.7%)	3776 (11.7%)
Malignant neoplastic disease	15 927 (3.8%)	2958 (7.7%)	2488 (6.4%)	2395 (9.3%)	22 857 (5.8%)	6043 (18.7%)
Diabetes	32 026 (7.7%)	5153 (13.3%)	2638 (6.8%)	2417 (9.4%)	31 297 (8.0%)	8415 (26.0%)
Obesity	15 538 (3.7%)	5419 (14.0%)	1142 (2.9%)	387 (1.5%)	69 233 (17.7%)	11 358 (35.1%)
Heart disease	26 336 (6.3%)	8608 (22.3%)	3171 (8.2%)	5038 (19.6%)	42 332 (10.8%)	11 238 (34.8%)
Hypertensive disorder	55 073 (13.3%)	12 231 (31.6%)	4964 (12.8%)	8210 (31.9%)	64 768 (16.6%)	15 640 (48.4%)
Renal impairment	16 401 (3.9%)	2176 (5.6%)	979 (2.5%)	894 (3.5%)	18 782 (4.8%)	6877 (21.3%)
Chronic obstructive pulmonary disease	6998 (1.7%)	3473 (9.0%)	774 (2.0%)	739 (2.9%)	10 685 (2.7%)	3790 (11.7%)
Dementia	5465 (1.3%)	1673 (4.3%)	262 (0.7%)	269 (1.0%)	7327 (1.9%)	1883 (5.8%)

(Table 1 continues on next page)

	COVID-19 diagnosis or PCR-positive test result					Hospitalised with COVID-19 (SIDIAP CMBD-AH, Spain; N=32 329)
	CPRD Aurum (UK; N=415 369)	IQVIA DA Germany (N=38 657)	IPCI (Netherlands; N=38 847)	IQVIA LPD Italy (N=25 759)	SIDIAP CMBD-AH (Spain; N=390 841)	
(Continued from previous page)						
Medication use (4–183 days before index date)						
Non-steroidal anti-inflammatory drugs	11 435 (2.8%)	5192 (13.4%)	3188 (8.2%)	5816 (22.6%)	55 929 (14.3%)	5415 (16.7%)
Cox2 inhibitors	289 (0.1%)	476 (1.2%)	217 (0.6%)	610 (2.4%)	2039 (0.5%)	279 (0.9%)
Systemic corticosteroids	5299 (1.3%)	938 (2.4%)	1425 (3.7%)	2577 (10.0%)	9248 (2.4%)	2209 (6.8%)
Antithrombotic and anticoagulant therapies	4958 (1.2%)	2026 (5.2%)	1430 (3.7%)	2496 (9.7%)	8063 (2.1%)	2592 (8.0%)
Lipid modifying agents	9542 (2.3%)	1939 (5.0%)	1519 (3.9%)	2323 (9.0%)	6612 (1.7%)	1633 (5.1%)
Antineoplastic and immunomodulating agents	2869 (0.7%)	242 (0.6%)	1831 (4.7%)	566 (2.2%)	3624 (0.9%)	435 (1.3%)
Hormonal contraceptives for systemic use	11 426 (2.8%)	381 (1.0%)	1768 (4.6%)	354 (1.4%)	5057 (1.3%)	208 (0.6%)
Tamoxifen	97 (<0.1%)	17 (<0.1%)	32 (0.1%)	17 (0.1%)	116 (<0.1%)	11 (<0.1%)
Sex hormones and modulators of the genital system	15 667 (3.8%)	481 (1.2%)	2145 (5.5%)	668 (2.6%)	6212 (1.6%)	306 (0.9%)
Immunoglobulins	0	0	12 (<0.1%)	14 (0.1%)	236 (0.1%)	60 (0.2%)
Summary variables						
One or more conditions of interest*	73 036 (17.6%)	12 925 (33.4%)	6749 (17.4%)	6557 (25.5%)	111 113 (28.4%)	20 630 (63.8%)
One or more medications of interest†	31 246 (7.5%)	6263 (16.2%)	6207 (16.0%)	7210 (28.0%)	66 929 (17.1%)	7290 (22.5%)
One or more conditions or medications of interest*†	97 723 (23.5%)	16 593 (42.9%)	11 532 (29.7%)	11 289 (43.8%)	155 747 (39.8%)	23 271 (72.0%)
Data are n (%) or median (IQR). CPRD Aurum=Clinical Practice Research Datalink Aurum database. IPCI=Integrated Primary Care Information database. IQVIA DA Germany=IQVIA Disease Analyzer Germany database. IQVIA LPD Italy=IQVIA Longitudinal Patient Database Italy. SIDIAP CMBD-AH=Information System for Research in Primary Care Conjunto Mínimo de Datos Básicos al Alta Hospitalaria data. *Conditions of interest: autoimmune disease, antiphospholipid syndrome, thrombophilia, asthma, atrial fibrillation, malignant neoplastic disease, diabetes, obesity, or renal impairment. †Medications of interest: non-steroidal anti-inflammatory drugs, Cox2 inhibitors, systemic corticosteroids, hormonal contraceptives, tamoxifen, and sex hormones and modulators of the genital system.						

Table 1: Patient characteristics

use agreement terms, which stipulate how these data can be used and for what purpose, is sufficient for the licensing commercial entities. Further inquiry related to the governance oversight of these assets can be made with the respective commercial entity. Because the study was based on routinely collected data, further patient consent was not required. The study protocol was registered in the European Union electronic register of post-authorisation studies (number EUPAS40414).

### Study outcomes and covariates

The primary study outcomes were venous thromboembolic events, arterial thromboembolic events, and death, all over 90 days from the index date. Venous thromboembolism was identified by diagnostic codes for either pulmonary embolism or deep vein thrombosis. Arterial thromboembolic events were identified as a composite of an acute myocardial infarction or acute ischemic stroke. Death was identified in all databases except for IQVIA DA Germany, in which it is not reliably captured. Code lists used for outcome ascertainment have been previously reported.<sup>20</sup>

The demographics of study participants are summarised, along with previous health conditions and medication use. Health conditions and medications of interest were prespecified based on discussions with key stakeholders and after reviewing the literature, and approved at protocol stage by the study funders. Comorbidities were identified

on the basis of prespecified Systemized Nomenclature of Medicine codes in all available history before the index date, and consisted of autoimmune disease, antiphospholipid syndrome, thrombophilia, asthma, atrial fibrillation, malignant neoplastic disease, diabetes, obesity, heart disease, hypertensive disorder, renal impairment, chronic obstructive pulmonary disease, and dementia. Previous medication use was characterised using RxNorm codes and based on prescriptions or dispensations in the 4–183 days before the index date. Medications extracted were antithrombotic and anticoagulant therapies, non-specific non-steroidal anti-inflammatory drugs, Cox-2 inhibitors, systemic corticosteroids, lipid modifying agents, antineoplastic and immune-modulating agents, hormonal contraceptives for systemic use, tamoxifen, and sex hormones and modulators of the genital system. Smoking (ever vs never) was also assessed based on diagnostic codes and observations recorded any time before the index date.

### Statistical analysis

The 90-day cumulative incidence of study outcomes was estimated, overall and stratified by age and sex. Cumulative incidence was estimated using the Kaplan-Meier approach for IQVIA DA Germany, where deaths were not reliably captured. To account for the competing risk of death, we used cumulative incidence functions for the other four databases.

For IQVIA see <https://www.iqvia.com>

	Events	Cumulative incidence*
<b>Venous thromboembolism</b>		
COVID-19 diagnosis or PCR-positive test result		
CPRD Aurum (UK)	930	0.27% (0.26–0.29)
IQVIA DA Germany	110	0.44% (0.36–0.53)
IQVIA LPD Italy	56	0.27% (0.21–0.35)
IPCI (Netherlands)	60	0.21% (0.16–0.27)
SIDIAP CMBD-AH (Spain)	3519	0.80% (0.77–0.83)
Hospitalised with COVID-19		
SIDIAP CMBD-AH (Spain)	3262	4.52% (4.37–4.68)
<b>Arterial thromboembolism</b>		
COVID-19 diagnosis or PCR-positive test result		
CPRD Aurum (UK)	165	0.06% (0.05–0.07)
IQVIA DA Germany	42	0.18% (0.12–0.23)
IQVIA LPD Italy	14	0.06% (0.04–0.11)
IPCI (Netherlands)	28	0.10% (0.07–0.15)
SIDIAP CMBD-AH (Spain)	3476	0.79% (0.77–0.82)
Hospitalised with COVID-19		
SIDIAP CMBD-AH (Spain)	2174	3.08% (2.96–3.21)
<b>Death†</b>		
COVID-19 diagnosis or PCR-positive test result		
CPRD Aurum (UK)	4792	1.35% (1.31–1.39)
IPCI (Netherlands)	344	1.08% (0.96–1.20)
IQVIA LPD Italy	364	1.66% (1.49–1.83)
SIDIAP CMBD-AH (Spain)	7609	1.99% (1.95–2.03)
Hospitalised with COVID-19		
SIDIAP CMBD-AH (Spain)	4649	14.61% (14.22–15.00)
Data are n or % (95% CI). CPRD Aurum=Clinical Practice Research Datalink Aurum database. IPCI=Integrated Primary Care Information database. IQVIA DA Germany=IQVIA Disease Analyzer Germany database. IQVIA LPD Italy=IQVIA Longitudinal Patient Database Italy. SIDIAP CMBD-AH=Information System for Research in Primary Care Conjunto Mínimo de Datos Básicos al Alta Hospitalaria data. *Cumulative incidence at 90 days for venous thromboembolism and arterial thromboembolism was calculated based on Kaplan-Meier for IQVIA DA Germany and on cumulative incidence functions for other databases. †Death is not reliably captured in IQVIA DA Germany.		
<b>Table 2: 90-day cumulative incidence of venous thromboembolism, arterial thromboembolism, and fatality in COVID-19 cases</b>		

For the analytical code see <https://www.github.com/oxford-pharmacoepi/coagulopathyInCovid19>

Cause-specific Cox models were used to describe the association between prespecified explanatory variables of interest and risks of venous thromboembolism, arterial thromboembolism, and death. The association between age and outcomes was studied stratified by sex and using restricted cubic splines to account for non-linearity where their use led to a lower Akaike information criterion compared with a linear fit. The association with sex was estimated with models adjusted for age. Finally, the association between comorbidities or medications and outcomes was estimated from unadjusted and age-adjusted and sex-adjusted models. Models were fit only where at least 40 outcomes were observed.

The effect of venous thromboembolism and arterial thromboembolism on COVID-19 outcomes was explored using a multistate model. Multistate models allow for a consideration of individuals' progression to multiple

events of interest, extending on competing risk models by also describing transitions to intermediate events,<sup>21</sup> and we previously developed such a model to describe patient outcomes during the first wave of COVID-19 in Catalonia, Spain.<sup>4</sup> We used a similar framework, where individuals start in a general population and COVID-19-free state and then can progress to states of outpatient COVID-19 diagnosis or PCR test positive, hospitalised with COVID-19, and death. The model was informed using data from SIDIAP CMBD-AH as this dataset was the only one to capture information on all of the states. The starting population was people in the SIDIAP CMBD-AH who, as of Sept 1, 2020, had at least 1 year of previous history captured in the database, had no previous history of COVID-19 or positive test for SARS-CoV-2, and no history of venous thromboembolism or arterial thromboembolism in the 1 year before. Venous thromboembolism and arterial thromboembolism were then assessed, separately, as time-dependent exposures for the following transitions: (1) from outpatient COVID-19 diagnosis or PCR test positive to hospitalised with COVID-19, (2) from outpatient COVID-19 diagnosis or PCR test positive to death (without a COVID-19 hospitalisation in between), and (3) from being hospitalised with COVID-19 to death. As above, models were adjusted for age and sex. The framework used for the multistate modelling is depicted in the appendix (p 1). All analytical code is available at GitHub.

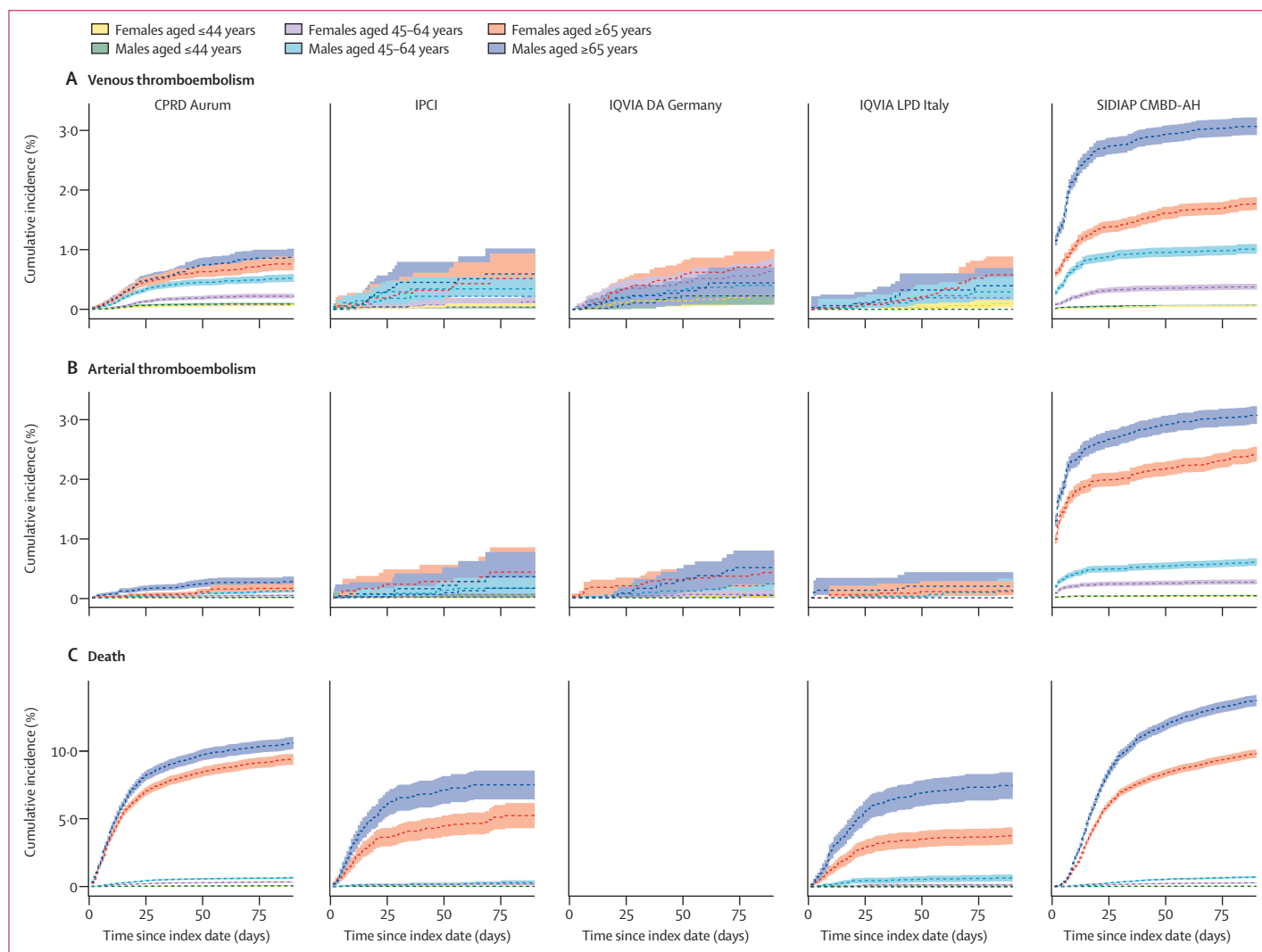
This study is registered with the EU Electronic Register of Post-Authorisation Studies (EUPAS40414).

### Role of the funding source

The funder provided feedback on the study protocol (study design, data collection, and data analysis) but had no role in the data interpretation or the decision to submit the manuscript for publication.

### Results

In total, 909 473 people with an index date on or after Sept 1, 2020, were included (415 369 from CPRD Aurum [the UK], 38 657 from IQVIA DA Germany, 38 847 from IPCI [the Netherlands], 25 759 from IQVIA LPD Italy, and 390 841 from SIDIAP CMBD-AH [Spain]). In addition, a cohort of 32 329 patients hospitalised with COVID-19 were identified from SIDIAP CMBD-AH. The median age of people with a COVID-19 diagnosis or PCR-positive test result ranged from 42 years (IQR 23–56) in SIDIAP CMBD-AH to 52 years (38–65) in IQVIA LPD Italy. In all databases, more COVID-19 cases were female than male (table 1). The proportion of individuals with a comorbidity or previous use of medication of interest ranged from 24% (97 723 of 415 369) in CPRD Aurum to 44% (11 289 of 25 759) in IQVIA LPD Italy. Those hospitalised with COVID-19 had a median age of 67 years (53–79), the majority were male, and 23 271 (72%) of 32 329 had a comorbidity or had previously used a medication of interest (table 1). The latest index date



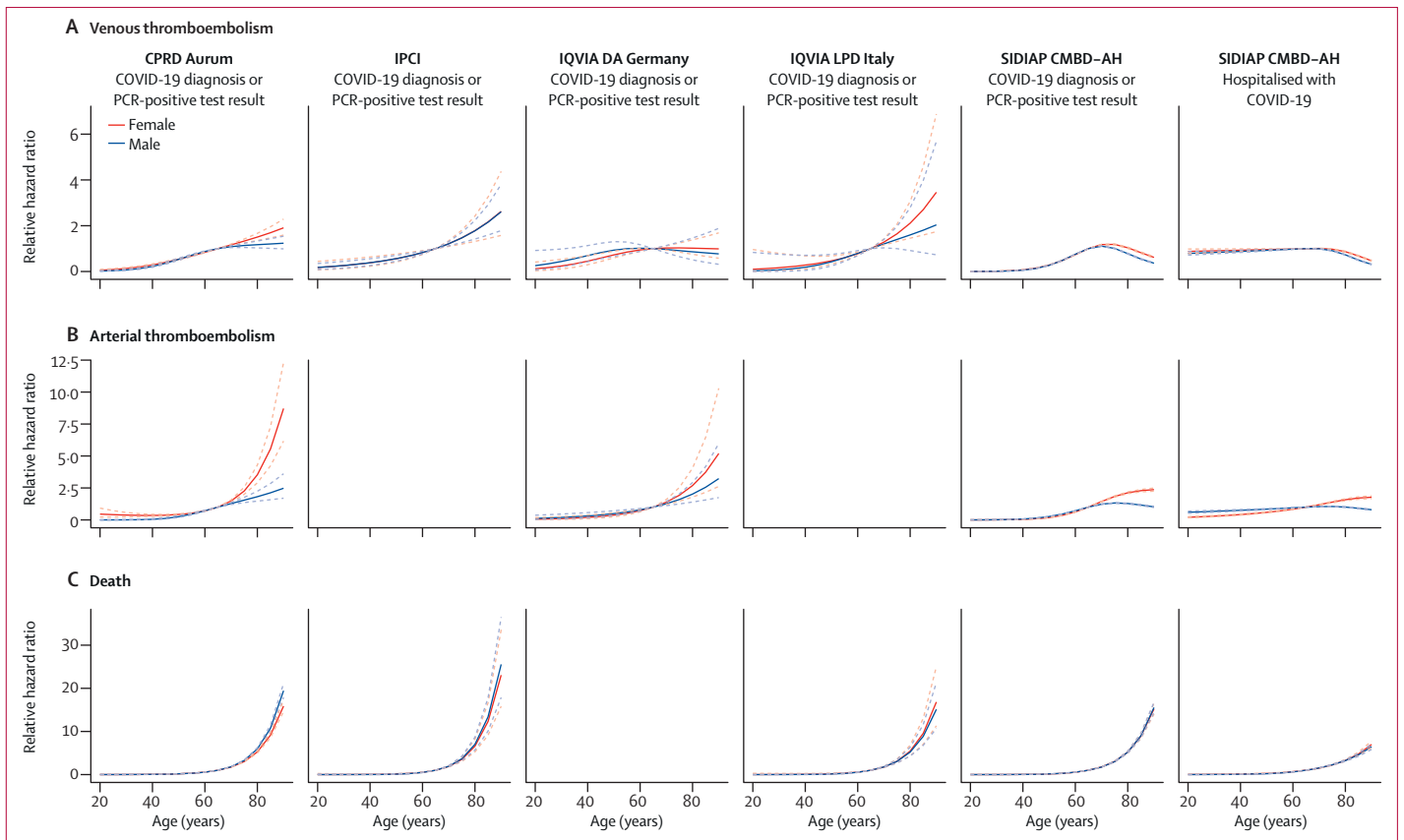
**Figure 1: Cumulative incidence of venous thromboembolism, arterial thromboembolism, and death in COVID-19 cases**

Data are stratified by age and sex. Estimates (solid lines) are presented with 95% CIs (dashed lines). Index date refers to the date of first COVID-19 diagnosis or positive RT-PCR test result. CPRD Aurum=Clinical Practice Research Datalink Aurum database. IPCI=Integrated Primary Care Information database. IQVIA DA Germany=IQVIA Disease Analyzer Germany database. IQVIA LPD Italy=IQVIA Longitudinal Patient Database Italy. SIDIAP CMBD-AH=Information System for Research in Primary Care Conjunto Mínimo de Datos Básicos al Alta Hospitalaria data.

observed in a database ranged from Jan 30, 2021, in CPRD Aurum to July 31, 2021, for IQVIA LPD Italy.

The 90-day cumulative incidence of venous thromboembolism among people with a COVID-19 diagnosis or PCR-positive test ranged from 0.21% (95% CI 0.16–0.27) in IPCI to 0.80% (0.77–0.83) in SIDIAP CMBD-AH, and was higher at 4.52% (4.37–4.68) for those hospitalised. For arterial thromboembolism, 90-day cumulative incidence ranged from 0.06% (0.05–0.07) in CPRD Aurum and 0.06% (0.04–0.11) in IQVIA LPD Italy to 0.79% (0.77–0.82) in SIDIAP CMBD-AH, and increased to 3.08% (2.96–3.21) among those hospitalised. Meanwhile, 90-day fatality among COVID-19 cases ranged between 1.08% (0.96–1.20) in IPCI to 1.99% (1.95–2.03) in SIDIAP CMBD-AH. 90-day fatality rose to 14.61% (14.22–15.00) among those hospitalised (table 2).

Incidence of venous thromboembolism, arterial thromboembolism, and death among people with a COVID-19 diagnosis or positive PCR test were generally higher among those aged 65 years and older (figure 1, appendix pp 2–13). For those hospitalised with COVID-19, although risks of arterial thromboembolism and death were also higher in older ages, this pattern was not seen for venous thromboembolism (appendix p 24). In models with age as the sole explanatory variable, risks of death were highest among those at the oldest ages in all databases in which deaths were captured (figure 2). Similarly, risks of arterial thromboembolism were increased for higher ages, but this increase was to a far lesser degree than for death. For venous thromboembolism, risks were higher as age increased, but peaked to then plateau or fall at around age 70 years in some databases. In models adjusted for



**Figure 2: Association of age with risks of venous thromboembolism, arterial thromboembolism, and death in COVID-19 cases, stratified by sex**

Data are estimates (solid lines) with 95% CIs (dashed lines), relative to age 65 years. Too few outcomes were observed to fit models for arterial thromboembolism for IPCI and IQVIA LPD Italy. CPRD Aurum=Clinical Practice Research Datalink Aurum database. IPCI=Integrated Primary Care Information database. IQVIA DA Germany=IQVIA Disease Analyzer Germany database. IQVIA LPD Italy=IQVIA Longitudinal Patient Database Italy. SIDIAP CMBD-AH=Information System for Research in Primary Care Conjunto Mínimo de Datos Básicos al Alta Hospitalaria data.

age, being male was associated with an increased risk of arterial thromboembolism and death for all databases (figure 3). Risks of venous thromboembolism were higher in male than in female individuals in CPRD Aurum, IPCI, and SIDIAP CMBD-AH, but not in IQVIA DA Germany or IQVIA LPD Italy (figure 3).

Most of the prespecified medications and comorbidities were associated with an increased risk of venous thromboembolism, arterial thromboembolism, and death among people with a COVID-19 diagnosis or positive PCR test in unadjusted models (figure 4). However, adjustment for age and sex attenuated most of the observed associations (figure 4). For arterial thromboembolism, consistent increases in risk were seen among those with a history of heart disease (in all databases) or hypertensive disorder (in all except IQVIA DA Germany in the adjusted model). Associations were more consistent for risk of death: corticosteroid use, diabetes, and renal impairment were all, for example, associated with an increased risk of death across most, if not all, databases.

Venous thromboembolism events in the outpatient setting (after COVID-19 diagnosis or PCR-positive test

but before hospitalisation) were associated with worse outcomes, with age-sex-adjusted hazard ratios (HRs) of 1.36 (95% CI 0.95–1.96) for hospitalisation and 4.42 (3.07–6.36) for death without hospitalisation. Inpatient venous thromboembolism (on or after date of hospital admission) was also associated with an increased risk of death (1.63 [1.39–1.90]; appendix p 14). Outpatient arterial thromboembolism was associated with an increased risk of death without hospitalisation (3.16 [2.65–3.75]) but not with risk of hospitalisation (1.05 [0.89–1.25]). Similarly, arterial thromboembolism on or after the date of hospital admission was associated with an adjusted HR of 1.93 (1.57–2.37) for death. Baseline characteristics for people at each transition in the multistate model are detailed in the appendix (pp 16–20).

In general, individuals identified before Sept 1, 2020, were older than those included from Sept 1, 2020, onwards (appendix p 22), and risk of death was higher in individuals identified before Sept 1, 2020, than in those identified from Sept 1, 2020, onwards (appendix p 25). Although for most databases results were similar regardless of how COVID-19 cases were identified, in CPRD Aurum risks were higher among those with a



clinical diagnosis than in those with a positive SARS-CoV-2 test (appendix p 26). Requiring a year of previous history had minimal effect on results (appendix p 27). All study results can be further explored in an online application.

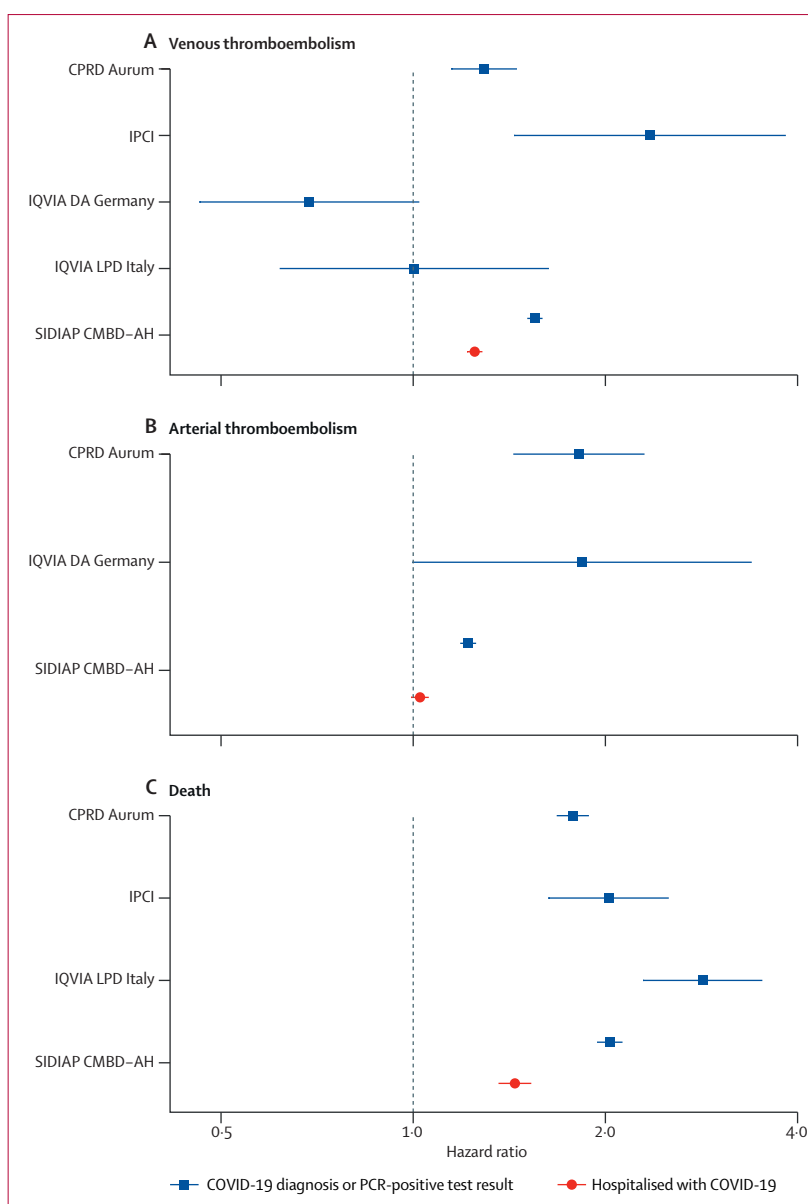
## Discussion

In this study we have estimated the incidence of venous thromboembolism and arterial thromboembolism among people with COVID-19. For people with a positive PCR test or a diagnosis of COVID-19, 90-day cumulative incidence ranged from 0.2% to 0.8% for venous thromboembolism and 0.1% to 0.8% for arterial thromboembolism. The highest estimates came from a database with patient-level linkage to hospital records. Incidence of these events increased to 4.5% for venous thromboembolism and 3.1% for arterial thromboembolism in those hospitalised with COVID-19. 90-day fatality was between 1.1% and 2.0% among COVID-19 cases and increased to 14.6% for those hospitalised.

Although fatality was much higher for those people at oldest ages, the increase in risks of arterial thromboembolism for higher ages was less pronounced and risks of venous thromboembolism appeared to peak at around age 70 years in some databases. After adjusting for age, men were generally at an increased risk of venous thromboembolism, arterial thromboembolism, and death. In addition, various comorbidities and previous medications were associated with an increased risk of venous thromboembolism, arterial thromboembolism, and death in unadjusted models, but estimates were generally attenuated after age and sex adjustment.

The occurrence of venous thromboembolism and arterial thromboembolism were both associated with worse COVID-19-related health outcomes for patients. The occurrence of venous thromboembolism before hospitalisation with COVID-19 was associated with increased risks of hospital admission and death (without an intermediate hospitalisation) following a COVID-19 diagnosis, and venous thromboembolism and arterial thromboembolism occurring after hospitalisation with COVID-19 were also associated with increased inpatient fatality.

Previous research predating the advent of COVID-19 indicated that risks of venous thromboembolism and arterial thromboembolism are heightened after respiratory infections.<sup>22,23</sup> Numerous studies have been undertaken to describe the prevalence of these events in COVID-19; however, these have predominantly focused on patients hospitalised with COVID-19, and even for this population estimates have varied substantially.<sup>2,3</sup> Overall, across studies that did not only include patients in the ICU, the prevalence of venous thromboembolism during admission was around 9% and the prevalence of arterial thromboembolism was 4%.<sup>2</sup> In our study, among people hospitalised with COVID-19, the 90-day

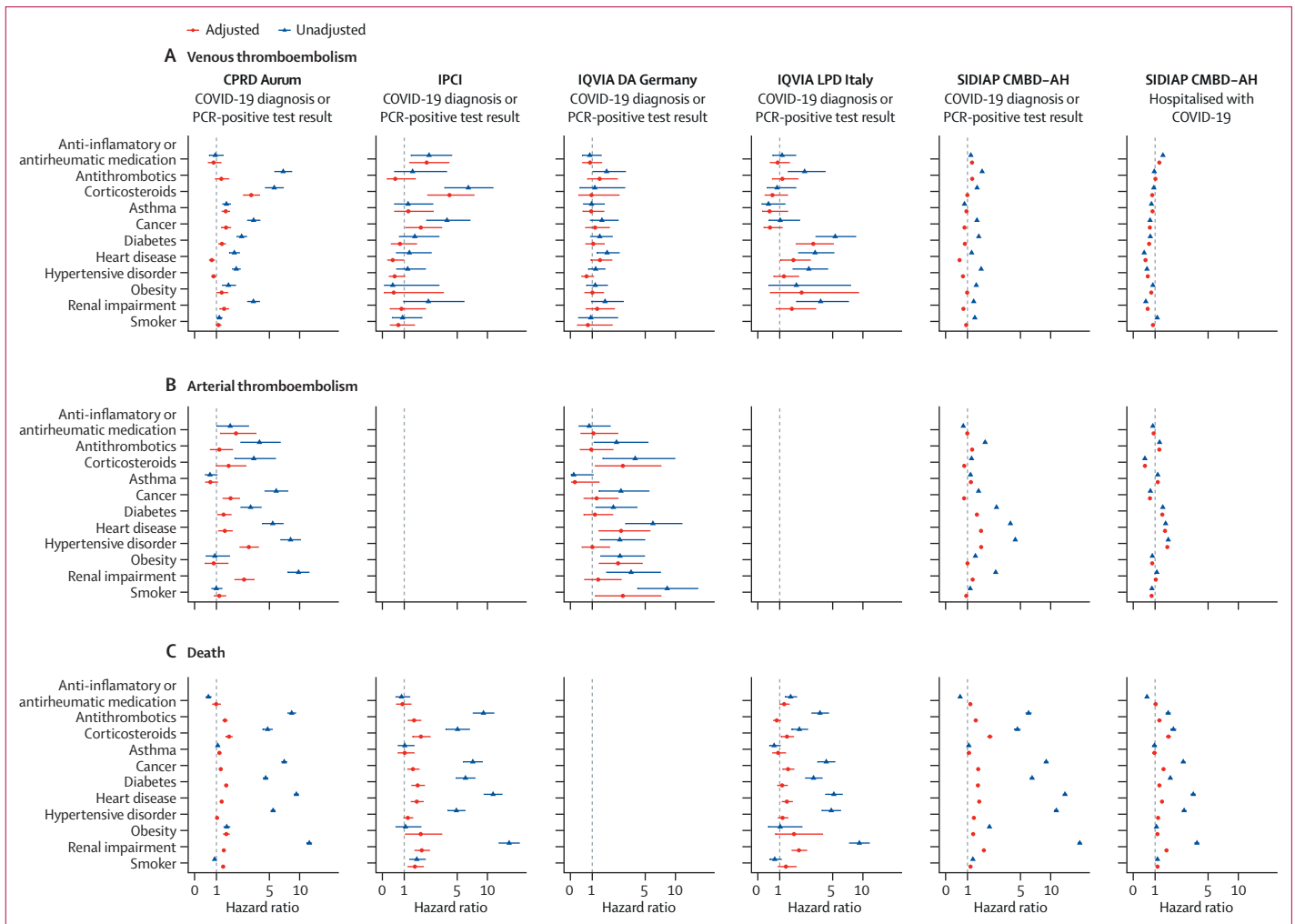


**Figure 3: Association of male sex (compared with female sex) with risks of venous thromboembolism, arterial thromboembolism, and death in COVID-19 cases, adjusted for age**

Horizontal lines represent 95% CIs. Too few outcomes were observed to fit models for arterial thromboembolism for IPCI and IQVIA LPD Italy. CPRD Aurum=Clinical Practice Research Datalink Aurum database. IPCI=Integrated Primary Care Information database. IQVIA DA Germany=IQVIA Disease Analyzer Germany database. IQVIA LPD Italy=IQVIA Longitudinal Patient Database Italy. SIDIAP CMBD-AH=Information System for Research in Primary Care Conjunto Mínimo de Datos Básicos al Alta Hospitalaria data.

cumulative incidence of venous thromboembolism was estimated to be 4.5% and of arterial thromboembolism was estimated to be 3.1%. Data on the incidence of these events among COVID-19 cases in general are scarce. As would be expected, in our study we found rates among all COVID-19 patients to be lower than among those hospitalised. Considering the vast numbers of people who have been infected with SARS-CoV-2, these relatively small risks translate into large numbers of people

For the **study data** see <https://livedataoxford.shinyapps.io/Covid19Thrombosis/>



**Figure 4: Unadjusted and age-sex-adjusted hazard ratios for association of selected medications and comorbidities with venous thromboembolism, arterial thromboembolism, and death** Horizontal lines represent 95% CIs. Too few outcomes were observed to fit models for arterial thromboembolism for IPCI and IQVIA LPD Italy. CPRD Aurum=Clinical Practice Research Datalink Aurum database. IPCI=Integrated Primary Care Information database. IQVIA DA Germany=IQVIA Disease Analyzer Germany database. IQVIA LPD Italy=IQVIA Longitudinal Patient Database Italy. SIDIAP CMBD-AH=Information System for Research in Primary Care Conjunto Mínimo de Datos Básicos al Alta Hospitalaria data.

being affected. Estimates were higher for the dataset with hospital linkage (SIDIAP CMBD-AH), probably indicating underreporting in datasets based solely on primary care records. It should also be noted that people with COVID-19 but who did not have an interaction with the health system were not captured.

In general, risks of venous thromboembolism and arterial thromboembolism can be expected to increase with age.<sup>24</sup> However, among COVID-19 patients we see, at least in some databases, risks of venous thromboembolism to peak around age 70 years. This is probably largely explained by the substantial competing risk of death for those with COVID-19, which is much increased with older age. Men have previously been associated with an increased risk of venous thromboembolism.<sup>25</sup> Similarly, men have also been seen to have an increased risk of arterial thromboembolism.<sup>26,27</sup> In our study we similarly found men to be at an increased risk of

both arterial thromboembolism and, in most databases, venous thromboembolism. Men were also associated with an increased risk of death among those with COVID-19 in our study, consistent with previous research.<sup>5,8,9</sup>

We have assessed the association of various comorbidities and previous medication use with risks of venous thromboembolism, arterial thromboembolism, and death among people with COVID-19. Whereas most of these were associated with increased risk of all three outcomes in unadjusted models, age and sex adjustment attenuated most of these associations. Findings for fatality were generally consistent across databases, with previous use of corticosteroids and comorbidities such as diabetes and heart disease associated with an increased risk of death. These findings are consistent with previous research.<sup>7-10</sup> For venous thromboembolism and arterial thromboembolism, age-adjusted and sex-adjusted estimates were somewhat

mixed. For arterial thromboembolism, heart disease and hypertension were associated with an increased risk in people with COVID-19 in almost all analyses, consistent with their associations in the general population.<sup>24</sup> However, only a few of these factors appeared associated with increased risks of venous thromboembolism among those with COVID-19.

This study relies on routinely collected health-care data, predominantly from outpatient records. While this has allowed for large study populations to be identified and included, the absence of linkage to hospital data is a limitation for all databases other than SIDIAP CMBD-AH. Absence of hospital linkage can be expected to lead to some degree of incompleteness for the overall study populations and has meant that outcomes for those hospitalised could be described only for SIDIAP CMBD-AH. With over 30 000 hospitalised patients included, this has allowed for a detailed analysis of risks of venous thromboembolism and arterial thromboembolism among people in such a setting.

As seen with other health conditions, analyses of so-called risk factors in COVID-19 are fraught with difficulties. Collider bias due, for example, to restricting analyses to people who have been diagnosed or hospitalised with COVID-19 can result in associations that do not exist in the general population or even reversing the sign of existing associations.<sup>28</sup> Meanwhile, modelling approaches that involve the mutual adjustment for various factors of interest in the absence of a causal framework can lead to the so called table 2 fallacy, as seen, for example, in interpretations around smoking and mortality in COVID-19.<sup>29,30</sup> Finally, unresolved confounding by indication precludes the causal interpretation around the observed associations between medicine use and risks of study outcomes. In our study, COVID-19 cases were identified based on clinical diagnoses or positive PCR tests. There is a risk of collider bias if, for example, PCR testing is focused on specific at-risk populations. Although this was the case during the first wave of the pandemic when the availability of testing was scarce, we report results here from Sept 1, 2020, onwards, at which point testing for SARS-CoV-2 was far more widely available. In addition, in our study, in which the research questions were descriptive in nature, we have used a modelling that at most adjusted for age and sex.

Drawing on routinely collected data from across Europe, in this study we have further described the epidemiology of venous thromboembolism and arterial thromboembolism in COVID-19. Although risks are relatively low overall, they increase with age, among men, and in those with certain risk factors. When they occur, venous thromboembolism and arterial thromboembolism are associated with worse outcomes, including hospitalisation and fatality. The absolute burden of thrombosis is high given the vast numbers of people who have been infected, and the consequences of these events to their health. The prevention of venous

thromboembolism and arterial thromboembolism remains, therefore, at the forefront in the management of severe COVID-19. More data are needed on the risk-benefit of anticoagulation among patients with less severe disease and younger patients who are often treated in outpatient settings.

#### Contributors

EB led the data analysis and wrote the initial draft of the manuscript with DP-A. EB, TD-S, CR, and SF-B had access to and verified the Information System for Research in Primary Care (SIDIAP) data; EB, AD, and DP-A had access to and verified the Clinical Practice Research Datalink (CPRD) data; PR had access to and verified the Integrated Primary Care Information (IPCI) data; and KK had access to and verified the data from IQVIA Longitudinal Patient Database (LPD) France, IQVIA LPD Italy, and IQVIA Disease Analyzer (DA) Germany. EB and DP-A were responsible for the decision to submit for publication. All authors were involved in the study conception and design, interpretation of the results, and the preparation of the manuscript.

#### Declaration of interests

DP-A's research group has received research grants from the European Medicines Agency, the Innovative Medicines Initiative, Amgen, Chiesi, and UCB Biopharma; consultancy fees from Amgen, UCB Biopharma, and AstraZeneca; and speaker fees from Astellas, Amgen, and UCB Biopharma. From the time of study design to analysis, KK was an employee of IQVIA. KK reported receiving funding from the National Institutes of Health National COVID Cohort Collaborative. IQVIA received funding from the University of Oxford on behalf of the Bill & Melinda Gates Foundation for the standardisation of data from IQVIA LPD Italy to a common data model for analysis, and the use of IQVIA DA Germany data for COVID-19-related research. TDS acknowledges receiving financial support from the Instituto de Salud Carlos III (ISCIII; Miguel Servet 2021: CP21/00023). All other authors declare no competing interests.

#### Data sharing

For patient privacy it is not possible to share the patient-level data that informed the study. All aggregated study results have been made available at <https://livedataoxford.shinyapps.io/Covid19Thrombosis/>.

#### Acknowledgments

This study was funded by the European Medicines Agency in the form of a competitive tender (Lot ROC No EMA/2017/09/PE). The views expressed in this Article are the personal views of the authors and may not be understood or quoted as being made on behalf of or reflecting the position of the European Medicines Agency or one of its committees or working parties. We acknowledge Prof Johan Van der Lei (Medical Informatics, Erasmus University Medical Centre, Rotterdam, Netherlands) for the overall management of this research grant.

#### References

- Bikdeli B, Madhavan MV, Jimenez D, et al. COVID-19 and thrombotic or thromboembolic disease: implications for prevention, antithrombotic therapy, and follow-up: JACC state-of-the-art review. *J Am Coll Cardiol* 2020; **75**: 2950–73.
- Tan BK, Mainbourg S, Friggeri A, et al. Arterial and venous thromboembolism in COVID-19: a study-level meta-analysis. *Thorax* 2021; **76**: 970–79.
- Nopp S, Moik F, Jilma B, Pabinger I, Ay C. Risk of venous thromboembolism in patients with COVID-19: a systematic review and meta-analysis. *Res Pract Thromb Haemost* 2020; **4**: 1178–91.
- Burn E, Tebé C, Fernandez-Bertolin S, et al. The natural history of symptomatic COVID-19 during the first wave in Catalonia. *Nat Commun* 2021; **12**: 777.
- Petrilli CM, Jones SA, Yang J, et al. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study. *BMJ* 2020; **369**: m1966.
- Reilev M, Kristensen KB, Pottegård A, et al. Characteristics and predictors of hospitalization and death in the first 11122 cases with a positive RT-PCR test for SARS-CoV-2 in Denmark: a nationwide cohort. *Int J Epidemiol* 2020; **49**: 1468–81.

- 7 Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* 2020; **395**: 1054–62.
- 8 Docherty AB, Harrison EM, Green CA, et al. Features of 20133 UK patients in hospital with COVID-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. *BMJ* 2020; **369**: m1985.
- 9 Gupta S, Hayek SS, Wang W, et al. Factors associated with death in critically ill patients with coronavirus disease 2019 in the US. *JAMA Intern Med* 2020; **180**: 1436–47.
- 10 Williamson EJ, Walker AJ, Bhaskaran K, et al. Factors associated with COVID-19 death in 17 million patients. *Nature* 2020; **584**: 430–36.
- 11 Lowe GDO. Common risk factors for both arterial and venous thrombosis. *Br J Haematol* 2008; **140**: 488–95.
- 12 Previtali E, Bucciarelli P, Passamonti SM, Martinelli I. Risk factors for venous and arterial thrombosis. *Blood Transfus* 2011; **9**: 120–38.
- 13 Voss EA, Makadia R, Matcho A, et al. Feasibility and utility of applications of the common data model to multiple, disparate observational health databases. *J Am Med Inform Assoc* 2015; **22**: 553–64.
- 14 Hripcsak G, Duke JD, Shah NH, et al. Observational health data sciences and informatics (OHDSI): opportunities for observational researchers. *Stud Health Technol Inform* 2015; **216**: 574–78.
- 15 Overhage JM, Ryan PB, Reich CG, Hartzema AG, Stang PE. Validation of a common data model for active safety surveillance research. *J Am Med Inform Assoc* 2012; **19**: 54–60.
- 16 de Ridder MAJ, de Wilde M, de Ben C, et al. Data resource profile: the Integrated Primary Care Information (IPCI) database, the Netherlands. *Int J Epidemiol* 2022; published online Feb 19. <https://doi.org/10.1093/ije/dyac026>.
- 17 Herrett E, Gallagher AM, Bhaskaran K, et al. Data resource profile: clinical practice research datalink (CPRD). *Int J Epidemiol* 2015; **44**: 827–36.
- 18 Recalde M, Rodríguez C, Burn E, et al. Data resource profile: the Information System for Research in Primary Care (SIDIAP). *Int J Epidemiol* 2022; published online April 13. <https://doi.org/10.1093/ije/dyac068>.
- 19 Jouaville SL, Miotti H, Coffin G, Sarfati B, Meihoc A. Validity and limitations of the Longitudinal Patient Database France for use in pharmacoepidemiological and pharmacoconomics studies. *Value Health* 2015; **18**: A18.
- 20 Burn E, Li X, Kostka K, et al. Background rates of five thrombosis with thrombocytopenia syndromes of special interest for COVID-19 vaccine safety surveillance: incidence between 2017 and 2019 and patient profiles from 38.6 million people in six European countries. *Pharmacoepidemiol Drug Saf* 2022; **31**: 495–510.
- 21 Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Stat Med* 2007; **26**: 2389–430.
- 22 Clayton TC, Thompson M, Meade TW. Recent respiratory infection and risk of cardiovascular disease: case-control study through a general practice database. *Eur Heart J* 2008; **29**: 96–103.
- 23 Clayton TC, Gaskin M, Meade TW. Recent respiratory infection and risk of venous thromboembolism: case-control study through a general practice database. *Int J Epidemiol* 2011; **40**: 819–27.
- 24 Lowe GDO. Common risk factors for both arterial and venous thrombosis. *Br J Haematol* 2008; **140**: 488–95.
- 25 Roach REJ, Lijfering WM, Rosendaal FR, Cannegieter SC, le Cessie S. Sex difference in risk of second but not of first venous thrombosis: paradox explained. *Circulation* 2014; **129**: 51–56.
- 26 Millett ERC, Peters SAE, Woodward M. Sex differences in risk factors for myocardial infarction: cohort study of UK Biobank participants. *BMJ* 2018; **363**: k4247.
- 27 Appelros P, Stegmayr B, Terént A. Sex differences in stroke epidemiology: a systematic review. *Stroke* 2009; **40**: 1082–90.
- 28 Griffith GJ, Morris TT, Tudball MJ, et al. Collider bias undermines our understanding of COVID-19 disease risk and severity. *Nat Commun* 2020; **11**: 5749.
- 29 Williamson EJ, Walker AJ, Bhaskaran K, et al. Factors associated with COVID-19 death in 17 million patients. *Nature* 2020; **584**: 430–36.
- 30 Westreich D, Edwards JK, van Smeden M. Comment on Williamson et al. The table 2 fallacy in a study of COVID-19 mortality risk factors. *Epidemiology* 2021; **32**: e1–2.