OpenSAFELY Protocol: Short-term safety of COVID-19 vaccines in England 2020/21 – acute venous thromboembolic events

Authors:

Anna Schultze*,[1,3] Jemma L Walker*,[2,3,4] Caroline E Morton,[1,5] Alex J Walker,[1,5] Helen J Curtis,[1,5] Julia Stowe,[4] Katherine Donegan,[6] Amir Mehrkar,[1,5] Christopher T Rentsch,[1,3] Elizabeth J Williamson,[1,3] Rosalind M Eggo,[1,3] Will Hulme,[1,5] Arina Tamborska,[7,8] Bhagteshwar Singh,[7,9] Rohini Mathur,[1,3] John Tazare,[1,3] Sebastian Bacon,[1,5] Daniel J Grint,[1,3] Krishnan Bhaskaran,[1,3] Simon Davy,[1,5] Laurie Tomlinson,[1,3] Chris Bates,[1,8] Harriet Forbes,[1,3] Kevin Wing,[1,3] Caroline Minassian,[1,3] Emily Nightingale,[1,3] Dave Evans,[1,5] Brian D Nicholson,[1,5] Peter Inglesby,[1,5] Brian MacKenna,[1,5] Kathryn E Mansfield,[1,3] George Hickman,[1,5] Amelia CA Green, Viyaasan Mahalingasivam,[1,3] Jonathan Cockburn,[1,10] Jess Morley,[1,5] Angel YS Wong,[1,3] Richard Croker,[1,5] John Parry,[1,10] Frank Hester,[1,10] Sam Harper,[1,8] Tom Solomon,[7,8] Stephen JW Evans,[1,3] Nick Andrews,[2,3,4] Ian J Douglas[1,3], Liam Smeeth,*[1,2,3] Ben Goldacre*,[1,5] Helen I McDonald*[1,2,3]

* Equal contributions

1. OpenSAFELY Collaborative
2. NIHR Health Protection Research Unit (HPRU) in Vaccines and Immunisation
3. London School of Hygiene and Tropical Medicine
4. Public Health England
5. The DataLab, Nuffield Department of Primary Care Health Sciences, University of Oxford
6. Medicines and Healthcare products Regulatory Agency (MHRA)
7. National Institute for Health Research Health Protection Research Unit (HPRU) in Emerging and Zoonotic Infections, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool
8. Division of Neurology, Walton Centre NHS Foundation Trust, Liverpool, L9 7LJ, UK
9. Tropical & Infectious Diseases Unit, Royal Liverpool University Hospital, Liverpool, L7 8XP, UK
10. TPP, Leeds

Version: 2.0

Date: 17 May 2021

Contents

Background

Safe vaccination against the coronavirus SARS-CoV-2 is vital for an effective response to the COVID-19 pandemic. New COVID-19 vaccines have demonstrated reassuring safety profiles in large clinical trials before authorisation for use by regulators.[1-3] However, even large clinical trials have limited power to detect rare safety events. In addition, when vaccinating a large population, baseline rates of disease mean that vaccination will occasionally coincide with the onset of illness even if the disease is not caused by vaccination. Vaccine confidence and uptake may be undermined by media stories of negative experiences following COVID-19 vaccination, whether these are genuine adverse effects of vaccination, or not.[4, 5] To monitor safety and maintain appropriate public confidence, regulators continue vaccine safety pharmacovigilance after authorisation for use.[6] Vaccine safety surveillance is designed to detect and strengthen potential safety signals, but does not adjust for confounding. Any potential safety signals detected require further investigation to assess the likelihood of a causal relationship between the vaccine and any potential adverse event.

On March 11 2021, the Danish authorities temporarily paused delivery of COVID-19 vaccination while investigating a report of a fatal blood clot in a person who had received the AstraZeneca COVID-19 adenovirus vector vaccine.[7] Germany and Austria also reported several cases of unusual thrombotic events and thrombocytopenia following vaccination.[12] A new syndrome of thrombosis with thrombocytopaenia has been described, thought to be mediated by antiplatelet-4 antibodies.[8] This has prompted further investigation of venous thromboembolism events following vaccination. A Danish-Norwegian registry study compared rates of cardiovascular and haemostatic events among healthcare workers vaccinated against COVID-19 to expected rates in the general population, standardised for age and sex.[9] They observed increased rates of venous thromboembolic events in the 28 days after vaccination, and in particular higher than expected rates of cerebral venous thrombosis. However, they were unable to adjust for potential confounders such as oral contraceptive use, and healthcare access and diagnosis of venous thromboembolic events may be better among healthcare workers than the general population.

In the United Kingdom, more than 33 million individuals have received at least one dose of COVID-19 vaccine, with two COVID-19 vaccines widely used: Pfizer BioNTech messenger RNA vaccine since 8 December 2020 and additionally AstraZeneca adenovirus vector vaccine since 4 January 2021.[10] Vaccine eligibility in the UK has been prioritised by older age, healthcare work, and underlying health conditions (**Appendix A**).[10] Universal healthcare with electronic health records offer an opportunity for a large observational study of COVID-19 vaccination and venous thromboembolic events, to investigate whether the excess rates of venous thromboembolism among Danish and Norwegian healthcare workers compared to the general population is also observed following vaccination among the general population vaccine priority groups in the UK.

This study aims to investigate the association of vaccination with Pfizer and AstraZeneca COVID-19 vaccines with venous thromboembolism in England, from the start of the national vaccination programme on 8 December 2020 to the report of potential safety concerns on 11 March 2021.

Objectives

The primary objective is to investigate the associations of each of Pfizer BioNTech and AstraZeneca COVID-19 vaccines with any venous thromboembolism (VTE) among COVID-19 vaccine priority groups in England, from the start of the national vaccination programme on 8 December 2020 until to the reporting of potential safety concerns on 11 March 2021. The study will also look at the specific outcomes of pulmonary embolism and cerebral venous thrombosis, explore for effect modification by age, and compare the associations of vaccination with any venous thromboembolism for AstraZeneca vaccine compared to Pfizer BioNTech COVID-19 vaccine.

A secondary objective is to investigate the extent to which self-controlled case series (SCCS) and cohort study designs differ in their estimations of any associations of COVID-19 vaccination with venous thromboembolic events, to aid interpretation of findings from other observational vaccine safety studies.

Methods

Data source

We will use data from general practice (GP) records, obtained from the GP software provider TPP, which represent approximately 40% of the population in England, linked to: (1) the NHSE/NHSX hospital inpatient activity data sets from Secondary Uses Service (SUS) data extracts; (2) Office for National Statistics (ONS) death data, including cause of death; (3) NHSE data of healthcare worker status collected at vaccination (available only for vaccinated individuals); and (4) the NHSE/NHSX Second Generation Surveillance System (SGSS) data on SARS-CoV-2 test results to ascertain a history of COVID-19 infection.

The data will be accessed, linked and analysed through OpenSAFELY, a new data analytics platform created on behalf of NHS England to address urgent questions relating to the epidemiology and treatment of COVID-19.[11] For more information on the OpenSAFELY data platform, please see **Appendix B**.

Study Design

Primary analysis for each outcome will use a self-controlled case series (SCCS) design. Secondary analysis will use a cohort study design with both historical and concurrent comparator groups.

SCCS is a case-only design which compares incidence of events within the same individual across different time periods. An SCCS design is suited to the study of acute outcome events and transient effects of precisely-timed exposures. It is particularly valuable for observational studies of vaccine safety, since people who receive vaccines typically differ from those who are eligible but unvaccinated, by characteristics which are unavailable in routinely recorded electronic health records, such as attitudes to health and health-care, health behaviours and healthcare access. This can result in confounding of observational studies of vaccine safety (the 'healthy vaccinee' effect).[12] Since self-controlled case series compare incidence of events within the same individual, the SCCS design implicitly controls for confounding by fixed individual characteristics, even for characteristics which are unmeasured.[13]

The SCCS design requires several assumptions, including that the exposure is not (appreciably) affected by outcome events.[13] If this assumption is violated, the design is vulnerable to reverse causality bias. This could conceivably be the case after the first safety report of thrombosis in March 11 2021, as people with a history of thrombosis may defer or decline vaccination, or be given particular brands in preference to AstraZeneca, and so our study is censored at March 2021. It may also be confounded by time-varying characteristics if these cannot be adequately adjusted for, such as secular trends in VTE (likely due to healthcare pressures during the pandemic affecting diagnosis and recording), if there is a time-dependent correlation with the timing of delivery of the national vaccination programme.

Secondary analysis will comprise a cohort study with both historical and concurrent comparator groups. A cohort study design can be a useful complement to investigate the robustness of findings from SCCS analysis in safety studies, since they have a different profile of biases.[14] The cohort study will not be susceptible to reverse causality, but will be susceptible to confounding by both fixed and time-varying characteristics (particularly the 'healthy vaccinee effect'). A cohort study with a concurrent unvaccinated comparator group can control for any confounding by trends over time in VTE, but may be particularly susceptible to the healthy vaccinee effect when vaccine uptake is high and unvaccinated individuals are increasingly atypical. For new vaccination programmes, the 'healthy vaccinee' effect may be reduced by selecting a historical unvaccinated comparator group (individuals who meet vaccine eligibility criteria during a recent time period before vaccination was available).[15] However, as with the self-controlled case series, a cohort study with a historical comparator group is susceptible to confounding by trends over time in VTE comparing pre- and post-vaccination periods, as well as potentially differential levels of case ascertainment between pandemic and non-pandemic time periods.

The strengths and limitations of each design are summarised in **Table 1**. The overall effect of the combined bias and confounding patterns in each study design is difficult to predict, and one of the aims of this study is to compare the estimates from each study design to aid interpretation of other studies of COVID-19 vaccine and venous thromboembolism.

Table 1: Strengths and limitations of the selected study designs

Design	Key strengths	Key limitations
Primary analysis: Self-controlled case series	● Not susceptible to selection bias or confounding by time-invariant characteristics (such as fixed attitudes or access to healthcare).	● Susceptible to reverse causality (if outcome affects the probability of exposure). ● May be susceptible to confounding by trends over time in VTE diagnoses (in analyses using a pre-vaccination comparison period) if there is a time-dependent correlation with the timing of the vaccination programme.
Cohort study (historical comparator group)	● Not susceptible to reverse causality. ● Confounding by 'healthy vaccinee effect' attenuated (compared to cohort study with concurrent comparator group)	● Somewhat susceptible to between person confounding due to 'healthy vaccinee effect' (compared to SCCS). ● Susceptible to confounding by trends over time in the outcome

	since the comparator group is not selected on basis of vaccination status, only vaccine eligibility criteria.	pre- compared to post-vaccination period and time-varying COVID-19 force of infection.
Cohort study (concurrent comparator group)	● Not susceptible to reverse causality. ● Less susceptible to confounding by trends over time in the outcome and time-varying COVID-19 force of infection.	● Susceptible to between person confounding due to 'healthy vaccinee effect'. ● High vaccine uptake could deplete comparator group particularly among older adults who were targeted first for vaccination and have high coverage. The healthy vaccinee effect could have a strong confounding effect and estimates of effect modification by age may be particularly affected.

Study Population

Eligible individuals will be aged 16-105 years, registered for at least a year on 1 July 2020 as permanent residents at a practice using TPP software, and eligible for COVID vaccination within JCVI vaccine priority groups 2-6 at baseline due to underlying health conditions, or age 65 years or over (JCVI priority group criteria are in **Appendix A**). Inclusion and exclusion criteria are summarised in **Table 2**. The eligible study population is restricted to individuals with at least one year's continuous registration prior to the study start date to ensure that outcome events represent new incident outcomes rather than retrospective records of past events.[16] For consistency with national definitions of vaccine eligibility groups, age will be categorised based on expected age on 31 March 2021. Vaccine eligibility due to underlying health conditions will be assessed using the PRIMIS vaccine eligibility specification at baseline on 1 July 2020.

Patients with missing age, or a recorded age under 16 years or over 105 years, missing sex, or missing postcode (used to identify care home residence) will be excluded. Individuals who receive a Moderna or unspecified COVID-19 vaccine brand will be considered no longer eligible for vaccination with AstraZeneca or Pfizer BioNTech vaccines, and excluded from the vaccination date. People who are indicated as receiving two or more types of vaccines on the same date will be excluded from that date. To ensure that only new incidences of venous thromboembolism are recorded as outcomes during the study period, we will exclude individuals with a diagnosis of venous thromboembolism recorded in the year prior to 1 July 2020. National guidance on vaccination of pregnant women was initially cautious, and numbers of pregnant women vaccinated are expected to be low during the study period, and so women with evidence of pregnancy during the 9 months prior to baseline will also be excluded.

The national vaccination programme began on 8 December 2020. The study start date is 1 July 2020, selected to provide some pre-vaccination comparison period while minimising changes over time in ascertainment of VTE during the study period as far as possible. During the period March 2020 to June 2020 primary care contacts and hospitalisations for many non-COVID conditions were considerably depressed compared to the same period in 2019, but were considerably recovered for many acute severe physical conditions by the start of July 2020.[17, 18] The study end date will be

the earliest of March 11 2021 (on which safety concerns were announced, which may have changed outcome ascertainment among vaccinees), or three weeks prior to the latest available SUS data on the date of data extraction, to allow for reasonably complete recording of hospitalisations.[22, 23]

Table 2: Summary of study population inclusion and exclusion criteria

	Inclusion / Exclusion Criteria
Inclusion criteria	Continuously registered in TPP for at least one year before the start of the study period on 1 July 2020 AND ● either aged 65–105 years on 31 March 2021 ● or an underlying health condition in the PRIMIS specification of the vaccine 'at risk' group assessed on 1 July 2021
Exclusion criteria	Missing age, or a recorded age that is under 16 years or over 105 years on 31 March 2021, missing sex, or missing postcode or received another COVID-19 vaccine (Moderna or unspecified) or evidence of pregnancy in the 9 months prior to 1 July 2020 or a record of venous thromboembolism in the year prior to 1 July 2020

Study Measures

Exposure

The primary exposure of interest is the first dose of brand-specific vaccination against COVID-19, with each of AstraZeneca or Pfizer BioNTech vaccine compared separately to absence of vaccination in primary analysis. First dose vaccination with AstraZeneca will be compared to first dose vaccination with Pfizer BioNTech COVID-19 vaccine as a secondary analysis.

Vaccination status will be determined from the 'vaccination table' in the TPP primary care record, categorised by vaccine type from SNOMED and dictionary of medicines and devices coding as Pfizer BioNTech, AstraZeneca, Moderna (for censoring vaccine eligibility) and unspecified brand (for censoring vaccine eligibility).[19] The details of all administered vaccinations are automatically sent to every person's GP record and the National Immunisation Management Service.

Both COVID-19 vaccines currently in use in the UK require a two-dose regimen, although national guidance advised there could be up to 12 weeks' interval between the first and second dose. In primary analysis we will disregard second doses. Individuals will be considered vaccinated from the first dose of vaccine, and the vaccine brand will be based on the first dose given. For secondary analysis, a second dose will be defined as a second record of vaccination at least 21 days after the first dose, and will be assumed to be the same brand as the first dose unless otherwise specified.

Outcomes

Our primary study outcome is any incident venous thromboembolic event (VTE), ascertained from primary care, hospital admission, or mortality records. As secondary outcomes, we will look specifically at cerebral venous thrombosis (disproportionately represented in clinical reports of VTE following vaccination) and at pulmonary embolism (since this is a severe acute event which may be less subject to ascertainment bias and misclassification than venous thromboembolism in general).

Our venous thromboembolism ICD-10 codelist (**Table 3**) was based on a validated codelist for VTE, supplemented with a validated codelist for cerebral venous thrombosis. The VTE codelist was validated among pregnant and postpartum women in the Clinical Practice Research Datalink with Hospital Episode Statistics linkage by Sultan et al., who found that a diagnosis code alone over-estimated incident VTE among pregnant and post-partum women, unless supported by evidence of anticoagulant therapy within 90 days of the event or death within 30 days.[20] Our study follow-up time is too short to include this requirement for all, but we will describe death and anticoagulant therapy following VTE events. We may overestimate VTE incidence but this is unlikely to be differential by vaccination status. If the overestimation found by Sultan et al. was partly due to recording of historical events prompted by assessment of VTE risk during pregnancy or peripartum, this would be less applicable to our general study population. The ICD-10 codes for cerebral venous thrombosis have been found to have a high positive predictive value for the diagnosis in UK hospital records when validated using radiological evidence.[21] The Read version 3 codelist for primary care records was developed to match the concepts in the ICD-10 codelist.

A single episode of VTE may result in more than one code being recorded at multiple points of follow up. A diagnostic code within 365 days of a previous diagnostic code (in any data source) will be considered a continuation of the same episode. Individuals with a diagnosis code (in any data source) within 365 days before the index date will be excluded, and individuals will not be eligible for

an incident episode for 365 days after a previous diagnostic record of VTE. In practice, this will result in follow up to the first outcome event during the study period.

Table 3: Codelists for venous thromboembolism

Codelist	
Primary care (CTV3)	https://codelists.opensafely.org/codelist/opensafely/vte-classified-codes/08b94f7d/ excluding codes for a history of VTE (14A8. 14A81 14AC. XaBMc XaBMc XaBMd XaZHi XE0q9)
Secondary care and mortality (ICD-10)	https://codelists.opensafely.org/codelist/opensafely/venous-thromboembolism-past-by-type-secondary-care-and-mortality-data/4402fca7/ excluding I870 post thrombotic syndrome

Covariates

Risk factors for VTE were considered as potential confounders if they could affect vaccine eligibility, uptake, or timing. Factors associated with vaccination prioritisation or coverage were also considered as potential confounders if they could affect investigation or diagnosis of VTE. Identification of potential confounders is detailed in **Appendix C**.

Recent hospital admission will be time-updated during the study. Age will be categorised based on expected age on 31 March 2021 for consistency with national vaccine eligibility. All other covariates **in Table 4** will be considered fixed and assessed at the baseline date of 1 July 2020, although we will explore this in sensitivity analysis (below).

Table 4: Covariates to be considered as potential confounders

Covariate	**Definition**
Age	Age on 31 March 2021
Sex	Male or female, excluding individuals with missing data
Residential care	Resident in a care home for older people (categorised as care home or nursing home)[22]
Area	Geographical area as categorised by Sustainability and Transformation Partnership (STP)
Ethnicity	Five categories, obtained from 16 (White, Black, South Asian, Mixed, Other) assessed from primary care records and supplemented with HES APC records https://codelists.opensafely.org/codelist/opensafely/ethnicity/2020-04-27/
Deprivation	Index of multiple deprivation (IMD) quintile based on lower super output area
Haematological cancer	Haematogical cancer (categorised as first ever code <1 year, 2-5 years, >5 years before 1 July 2020) https://codelists.opensafely.org/codelist/opensafely/haematological-cancer/2020-04-15/
Non-haematological cancer	Non-haematological cancer (categorised as first ever code <1 year, 2-5 years, >5 years before 1 July 2020) https://codelists.opensafely.org/codelist/opensafely/cancer-excluding-lung-and-haematological/2020-04-15/ https://codelists.opensafely.org/codelist/opensafely/lung-cancer/2020-04-15/

Previous VTE	Any history of VTE prior to 1 July 2019 (individuals with a record of VTE between 1 July 2019 and 30 June 2020 will be excluded). Identified from primary care records: https://codelists.opensafely.org/codelist/opensafely/vte-classified-codes/08b94f7d/ or secondary care records: https://codelists.opensafely.org/codelist/opensafely/venous-thromboembolism-past-by-type-secondary-care-and-mortality-data/4402fca7/
Severe obesity	Based on latest Body Mass Index (BMI) and classified as 40+ kg/m^2, consistent with vaccine prioritisation guidance. Individuals with missing BMI measurements will be classified as being not severely obese (BMI less than 40).
Chronic heart disease	Any chronic cardiac disease https://codelists.opensafely.org/codelist/opensafely/chronic-cardiac-disease/2020-04-08/ including atrial fibrillation or flutter https://codelists.opensafely.org/codelist/opensafely/atrial-fibrillation-or-flutter/2020-07-30/
COPD	Any history of chronic obstructive pulmonary disease https://www.opencodelists.org/codelist/opensafely/current-copd/2020-05-06/
Other chronic respiratory disease	Any history of other chronic respiratory diseases that are an indication for COVID-19 vaccination, including fibrosing lung disease, bronchiectasis and cystic fibrosis (not including asthma) https://www.opencodelists.org/codelist/opensafely/other-respiratory-conditions/2020-05-14/
Chronic liver disease	Any history of chronic liver disease https://codelists.opensafely.org/codelist/opensafely/chronic-liver-disease/2020-06-02/
Stroke	Any history of stroke, categorised as - recent (first ever diagnosis code within the 3 months prior to 1 July 2020) - historical (any diagnosis code prior to 1 April 2020) https://codelists.opensafely.org/codelist/opensafely/stroke-updated/2020-06-02/
Other chronic neurological conditions	Chronic neurological disease which would be an indication for vaccine prioritisation, and also likely to cause leg paresis or limited mobility: hemiplegia or quadriplegia; motor neurone disease; cerebral palsy; MS/neuromyelitis optica; Parkinson's disease.
SLE/Behçet's disease	Any diagnosis of systemic lupus erythematosus or Behçet's disease.
Inflammatory bowel disease	Any diagnosis of inflammatory bowel disease https://www.opencodelists.org/codelist/opensafely/inflammatory-bowel-disease/2020-04-07/
Chronic kidney disease	Any history of dialysis or kidney transplant https://www.opencodelists.org/codelist/opensafely/chronic-kidney-disease/2020-04-14/ or Chronic kidney disease stage 3-5 based on latest estimated glomerular filtration rate at baseline
Smoking	Smoking status (current, former, never)

Combined oral contraceptive pill	Current use will be defined using primary care prescription records (at least one prescription in the six months before 1 July 2020), and categorised as: - new users (first ever prescription within 6 months prior to 1 July 2020) - recently started user (first ever prescription 12-6 months prior to 1 July 2020) - longer-term users (any prescription before 1 July 2019)
Hormone replacement therapy	Current use will be defined using primary care prescription records (at least one prescription in the six months before 1 July 2020), and categorised as: - new users (first ever prescription within 6 months prior to 1 July 2020) - recently started user (first ever prescription 12-6 months prior to 1 July 2020) - longer-term users (any prescription before 1 July 2019)
ACE inhibitor, angiotensin receptor blocker, calcium channel blocker	Current use will be defined as any primary care prescription in the six months before 1 July 2020, and each class (ACEi, ARB and calcium channel blocker) considered separately https://www.opencodelists.org/codelist/opensafely/ace-inhibitor-medications/2020-05-19/ https://www.opencodelists.org/codelist/opensafely/angiotensin-ii-receptor-blockers-arbs/2020-05-19/ https://www.opencodelists.org/codelist/opensafely/calcium-channel-blockers/2020-05-19/
NSAIDs, antiplatelet medications, oral anticoagulant	Current use will be defined as any primary care prescription in the six months before 1 July 2020, and each class (NSAID, antiplatelet medication, oral anticoagulant) considered separately
Time-updated covariates	
Recent hospital admission (infectious or non-infectious)	Any overnight hospital admission with discharge within the previous 3 months, time-updated through the study period, with separate variables for - admissions with infection (admissions in which the primary or non-primary admission diagnoses included any diagnosis of infection) - non-infectious admissions (admissions in which the primary or non-primary admission diagnoses did not include any diagnosis of infection)

For analysis with the specific outcome of cerebral venous thrombosis we will also consider recent meningitis/encephalitis as a potential confounder, time-updated through the study.

Three additional covariates will be used in descriptive analysis or sensitivity analyses.

- We will describe the number of VTE events that were followed by death (from any cause) within 28 days or anticoagulation therapy within 90 days, by vaccination status. Death will be ascertained from ONS mortality records, and anticoagulation therapy from a primary care prescription for an oral or parenteral anticoagulant (including warfarin, heparin or direct acting oral anticoagulants).
- We will describe a history of COVID-19 in the individual or (separately) in a household member within the previous four weeks or three months prior to VTE, according to vaccination status at the time of VTE. COVID-19 status will be defined using positive test result from SGSS.
- We will describe the number of pregnancy outcomes (livebirths, stillbirths, miscarriage and abortions) during the study using primary care records of these events.

Statistical Analysis

Descriptive

We will describe the type of venous thromboembolic event by site, classified as: deep vein thrombosis (DVT)/pulmonary embolism (PE); cerebral venous thrombosis; portal vein thrombosis; and other sites/unspecified), overall, and stratified by vaccination status (AstraZeneca, Pfizer BioNTech, unvaccinated).

To explore the potential misclassification of incident venous thromboembolism, we will describe the number of VTE events that were followed by death (from any cause) within 28 days or anticoagulation therapy (within 90 days), by vaccination status and VTE type. Death will be ascertained from ONS mortality records, and anticoagulation therapy from a primary care prescription for an oral or parenteral anticoagulant (including warfarin, heparin or direct acting oral anticoagulants).

To describe the extent to which a history of COVID-19 or household exposure to COVID-19 may explain any association between vaccination and VTE, we will describe a history of COVID-19 (in the individual or in a household member prior to VTE (categorised as 0-4, or 5-12 weeks prior), according to vaccination status and stratified by calendar time (months).

Self-controlled case series

The primary analysis will use a self-controlled case series design, which uses conditional Poisson regression to calculate incidence rate ratios comparing time periods for the same individual, with time windows defined relative to the individual's vaccination date. A separate analysis will be conducted for each of AstraZeneca and Pfizer BioNTech COVID-19 vaccine brands.

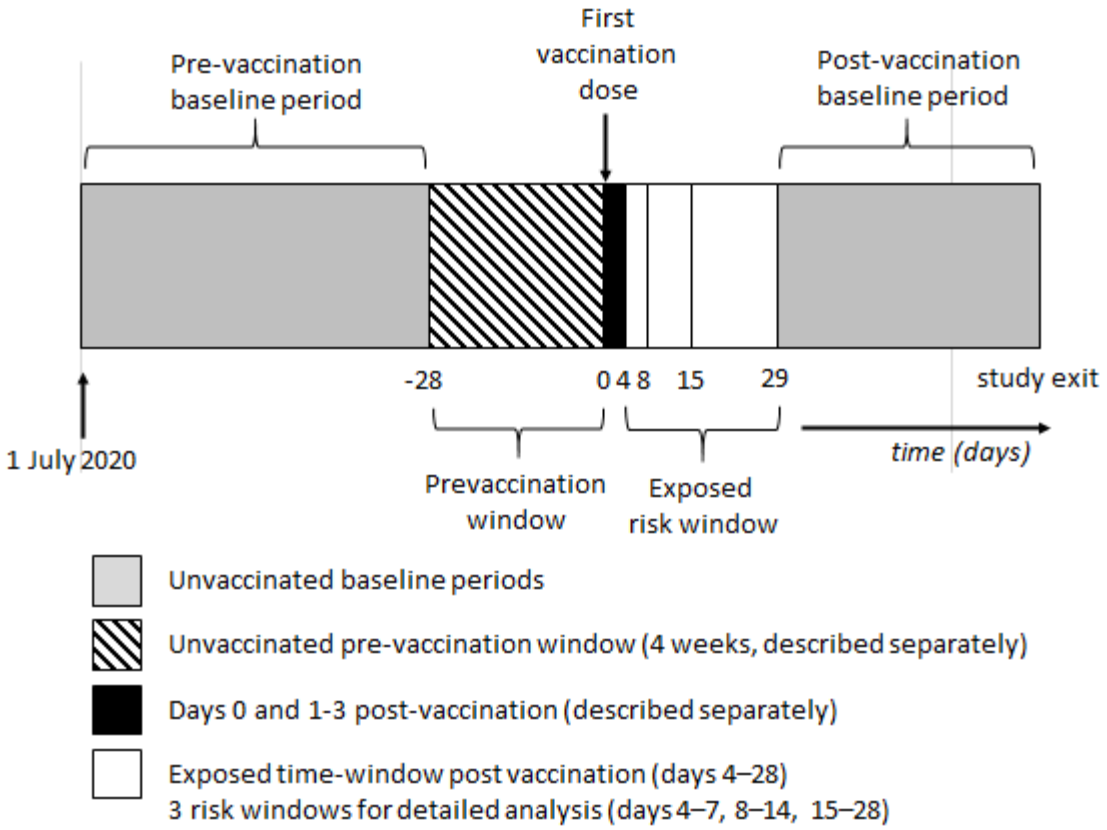
Eligible individuals will be included in self-controlled case series analyses if they have experienced both the exposure (brand-specific first dose of COVID-19 vaccination) and the outcome of interest during the study period.

SCCS designs assume that occurrence of an event (venous thromboembolism) should not alter the probability of subsequent exposure (COVID-19 vaccination).[13, 23] Although the study population are all within JCVI priority groups 1-6 and so eligible for vaccination during the study, admission to hospital with an incident VTE may prompt or defer vaccination. Conversely, vaccination is recommended to be deferred for 4 weeks following COVID-19 infection, which can cause VTE. To minimise reverse causation from vaccine deferral or expedition, we will describe separately a pre-exposure window of 4 weeks (and explore this in sensitivity analysis described below). We will also exclude the day of vaccination from the risk window, to exclude opportunistic recording of outcomes on attendance for vaccination.[24]

The main period of interest for post-vaccination exposure is 4–28 days after first dose of vaccination. The exposed time period will be compared to the unexposed period which runs from 1 July 2020 to the start of the pre-exposure window, and restarts 29 days after first vaccination dose until the end of follow up. The incidence of VTE is not expected to be constant over time following vaccination (particularly as prevention of COVID-19 related VTE would be expected to develop over this time period), so we will break the exposed period into risk windows of 1-3, 4-7, 8-14, and 15-28 days. We will describe the number of events on the day of vaccination separately to exclude opportunistic recording on the day of vaccination.

The main analysis will be an unadjusted analysis using conditional Poisson regression. We will then explore for time-varying confounding by time-updated hospital admission status. All other potential confounders are considered fixed individual characteristics within the short time frame of the study, and innately controlled for by the study design. We will additionally explore for effect modification by age, grouped as 16–39, 40-64 years and 65–105 years.

Figure 1: Illustration of time-windows for self-controlled case series (main analysis)



SCCS design assumes that study outcome events do not influence the length of observation period. This assumption may be violated by high mortality following both study outcomes. We will describe the number of individuals who died within 28 days of the outcome event and check for event-dependent censoring of observation time by describing time from event to end of observation in two histograms according to whether observation times are censored or not.[23] Depending on findings we will consider whether to exclude all individuals who died within 28 days of their outcome event (if there are a small number of deaths), or use an extension of the SCCS model to take the distribution of intervals from event to end of observation into account.[25]

The SCCS design assumes that recurrent outcome events are independent, that is, the chance of a second event is not influenced by having a first event.[13] Unprovoked VTE has a high recurrence rate, but the study definition of an incident event means that recurrence cannot occur within the study period.

Sensitivity Analyses

We will conduct sensitivity analyses to assess the robustness of the primary analysis to our assumptions.

Using only pre- or post-vaccination time as a comparison period

The main analysis will compare the risk window of 4-28 days after first vaccination to both pre- and post-vaccination time. Comparison to the pre-vaccination period before the vaccine programme began is susceptible to confounding by secular trends in VTE diagnosis. Rates of VTE may be lower at 28+ days post-vaccination, as vaccine may prevent VTE as a complication of COVID-19. Both comparisons may therefore over-estimate the association of vaccination with VTE, but to different extents. In sensitivity analysis we will compare the risk window of 4-28 days separately to each of the pre- and post-vaccination comparison periods to explore this.

Varying pre-exposure excluded window

We will describe the interval between vaccination and outcome for each outcome, using a histogram (an 'exposure-centred interval plot') to assess the suitability of the length of the pre-exposure window for each outcome, and may conduct sensitivity analyses varying the length of the pre-exposure window in response to these.[23] In addition, depending on the effect of acute illness on deferral or non-uptake of vaccination, exclusion of the pre-exposure window may in fact introduce bias. We will therefore conduct a sensitivity analysis in which the pre-exposure window is not excluded from the baseline comparison period.

Unspecified or conflicting vaccine brand

We will describe the number of individuals with unspecified brand of vaccination, and individuals who receive a second dose of vaccine of a different brand to the first dose. Depending on numbers, we may conduct sensitivity analyses to additionally include individuals with unspecified brand in each brand-specific analysis, and censoring follow up at second vaccination dose if the second dose of vaccine is a different vaccine brand to the first dose.

Excluding healthcare workers

Venous thromboembolism may result in time off work and even prevent return to work, which could introduce reverse causation in the SCCS analysis among healthcare workers eligible for vaccination due to their work. We will therefore conduct a sensitivity analysis excluding healthcare workers, using a record of healthcare worker status collected at the time of vaccination.

Cohort study

A complementary analysis will use a cohort study design with concurrent and historical comparator groups. We will analyse the association of each vaccine with VTE separately, using a Poisson regression model with lexis expansions for time post vaccination.

For the cohort study, eligible individuals who are vaccinated will be matched to up to 10 concurrent and 10 historical controls by age (80+, 75-79, 70-74, 65-69, and <65 years to reflect JCVI priority group categories) and area (STP). Matching on STP has been selected for roughly comparable timing of eligibility to be vaccinated, since the timing of vaccine roll-out to each priority group varied locally but matching too closely on this (for example by general practice) could exacerbate the 'healthy vaccinee' effect. For the concurrent comparator group this will also result in a roughly comparable timing of vaccine eligibility relative to local health pressures and force of COVID-19 infection which may both drive trends over time in VTE recording.

Individuals will be eligible to be concurrent controls from the latest of 8 December 2020 or 365 days after their latest diagnosis of VTE prior to 8 December 2020, and the index date will be the

vaccination date. Individuals will be eligible to be historical controls from the latest of 1 July 2020 or 365 days after a diagnosis of VTE, and the index date will be the vaccination date for the vaccinated individual, and 160 days prior to the vaccination date (the difference between 1 July and 8 December) for the historical control. Individuals will exit the study at the earliest of: study outcome, death, leaving the practice, or the study end date. Vaccinated individuals will be eligible to become controls until their vaccination date. Follow up of controls will be censored at vaccination (with any brand of vaccine, including Moderna or unspecified brand).

We will explore for potential confounding by age, sex, care home residence, deprivation, and additionally by relevant underlying health conditions and medications at baseline (**Table 4**).

To explore for potential confounding by ethnicity, we will assess the association of vaccination with VTE with and without adjustment for ethnicity in a complete case analysis. If there is evidence of confounding, we will analyse the final model as a complete case analysis, excluding individuals with missing ethnicity.

Finally, we will explore for effect modification by age, grouped as 16-39, 40-64 and 65-105 years.

Sensitivity analyses

Restricted to vaccination in 2021

The initial vaccination programme in December targeted individuals at highest risk and in particular settings, such as hospital inpatients. The additional deployment of the AstraZeneca vaccine in early January, increasing vaccine supply and reducing cold chain requirements, was followed by vaccination of less highly-selected individuals. We will therefore restrict to individuals vaccinated after 1 January 2021 for a sensitivity analysis with lower power and shorter duration of post-vaccination follow up, but which may be less subject to confounding by unmeasured characteristics.

Time-updating confounders

We will assess underlying health conditions and medication use at baseline, since numbers of new diagnoses are expected to be small during the study period. To check this assumption, we will describe the number of individuals newly diagnosed with each included underlying health condition during the study (for the vaccinated group and concurrent comparator group), according to vaccination status at the time. For each medication we will describe the number of individuals who receive a prescription during the study for each included medication but were not assessed as receiving the medication at baseline, and whether these were first time users, again according to vaccination status at the first prescription during the study. If any covariates are found to have both a strong confounding effect when assessed at baseline, and a considerable incidence of new onset during the study, we will investigate the effect of time-updating this covariate.

Similarly, we are unable to reliably identify the timing of current pregnancies to time-update pregnancy status during the study. We will describe the number of pregnancy outcomes (livebirths, stillbirths, miscarriage and abortions) during the study using primary care records. Depending on numbers, we will consider the need for sensitivity analysis excluding individuals with any record of a completed pregnancy during the study.

Length of follow-up for the concurrent comparator group

We will describe length of follow up for concurrent controls by age group. For age groups with high coverage during the study and depletion of controls, we will consider the need for sensitivity analysis censoring follow up of these age groups to reduce the healthy vaccinee effect.

Quantitative bias analysis

Depending on results, we may also consider quantitative bias analysis to describe the potential difference in confounding between analysis with concurrent and historical comparator groups, and to explore the potential extent of differential outcome misclassification between the comparison groups using concurrent and historical comparator groups.

Secondary analyses

For both SCCS and cohort study designs, we will undertake the following secondary analyses.

Relative safety of the two different vaccines

When two products are in use as alternatives for the same population, a head-to-head comparison can reduce confounding by indication.[14] In the UK, the Pfizer BioNTech COVID-19 mRNA vaccine was available sooner than the AstraZeneca COVID-19 adenovirus vector vaccine, and was delivered to the highest priority individuals, including targeting by setting such as hospital inpatients. During the time in which both vaccinations have been available, the different cold chain requirements and batch sizes have resulted in the delivery of different vaccination brands by setting, which is likely to result in confounding of a head-to-head comparison, varying by location due to different local delivery plans. In addition, a head-to-head comparison of vaccine brands may miss any safety issues common to both vaccine brands. Despite its potential advantages for reducing confounding by indication, a head-to-head comparison of the vaccine brands has therefore not been selected for main analysis.

As a secondary analysis for both SCCS and the cohort study designs we will use a ratio-of-ratios approach to compare the relative safety of the two vaccine brands, with the *a priori* hypothesis that the AstraZeneca COVID-19 vaccine will have a greater association with increased venous thromboembolism incidence than the Pfizer BioNTech COVID-19 vaccine, due to a combination of lower effectiveness to prevent COVID-19 (and hence VTE as a complication) and possibly any adenovirus platform-specific vaccine safety issues. This analysis will be limited to vaccinations from 1 January 2020 (the period from which both vaccines were available), and will be interpreted with caution due to the likelihood of confounding by indication.

Second dose of vaccination

Depending on the number of eligible individuals who received a second dose of vaccination during follow up, we will consider the second dose of vaccination as a second exposure. In cohort study analysis, vaccine exposure will be categorised as unvaccinated, first dose, or second dose received. In SCCS analysis, days 4-28 post second dose of vaccination will comprise a separate risk window, censoring post-vaccination follow-up from the first dose if needed.

Future analysis

Future analyses may explore changes in outcome recording following the public announcement of a possible safety concern regarding blood clots following vaccination on March 11 2021, but this is outside the scope of this study at present due to limited follow up time currently available after March 11 2021 in secondary care data.

Software and Reproducibility

Data management will be performed using Python and Google BigQuery, with analysis carried out using Stata 16.1 MP. The codelists, and code for data management and analysis will be archived online.

Administrative

Information Governance

NHS England is the data controller; TPP is the data processor; and the key researchers on OpenSAFELY are acting on behalf of NHS England. This implementation of OpenSAFELY is hosted within the TPP environment, which is accredited to the ISO 27001 information security standard and is NHS IG Toolkit compliant;[26, 27] patient data have been pseudonymized for analysis and linkage using industry standard cryptographic hashing techniques; all pseudonymized datasets transmitted for linkage onto OpenSAFELY are encrypted; access to the platform is through a virtual private network (VPN) connection; the researchers hold contracts with NHS England and only access the platform to initiate database queries and statistical models; all database activity is logged; and only aggregate statistical outputs leave the platform environment following best practice for anonymization of results such as statistical disclosure control for low cell counts.[28] The OpenSAFELY research platform adheres to the data protection principles of the UK Data Protection Act 2018 and the EU General Data Protection Regulation (GDPR) 2016. In March 2020, the Secretary of State for Health and Social Care used powers under the UK Health Service (Control of Patient Information) Regulations 2002 (COPI) to require organizations to process confidential patient information for the purposes of protecting public health, providing healthcare services to the public and monitoring and managing the COVID-19 outbreak and incidents of exposure.[29] Together, these provide the legal bases to link patient datasets on the OpenSAFELY platform. GP practices, from which the primary care data are obtained, are required to share relevant health information to support the public health response to the pandemic, and have been informed of the OpenSAFELY analytics platform.

Ethics

We received ethics approval to conduct the data linkage and analyses by the London - City & East Research Ethics Committee on the 2nd of April 2020 (REC reference: 20/LO/0651) and LSHTM Ethics Board (ref 21863). No further ethical or research governance approval was required by the University of Oxford but copies of the approval documents were reviewed and held on record.

Funding

This work was supported by the NIHR Health Protection Research Unit in Vaccines and Immunisation and the Medical Research Council MR/V015737/1. TPP provided technical expertise and infrastructure within their data centre pro bono in the context of a national emergency. HIM, JLW, and NA are funded by the National Institute for Health Research (NIHR) Health Protection Research Unit in Vaccines and Immunisation, a partnership between Public Health England and LSHTM.
EW was supported by MRC project grant MR/S01442X/1. BG's work on better use of data in healthcare more broadly is currently funded in part by: NIHR Oxford Biomedical Research Centre, NIHR Applied Research Collaboration Oxford and Thames Valley, the Mohn-Westlake Foundation, NHS England, and the Health Foundation; all DataLab staff are supported by BG's grants on this work.

LS reports grants from Wellcome, MRC, NIHR, UKRI, British Council, GSK, British Heart Foundation, and Diabetes UK outside this work. JPB is funded by a studentship from GSK. AS is employed by LSHTM on a fellowship sponsored by GSK. KB holds a Sir Henry Dale fellowship jointly funded by Wellcome and the Royal Society. AYSW holds a fellowship from BHF. RM holds a Sir Henry Wellcome fellowship. ID holds grants from NIHR and GSK. RM holds a Sir Henry Wellcome Fellowship funded by the Wellcome Trust. HF holds a UKRI fellowship. RME is funded by HDR-UK and the MRC. HK receives funding from the PENDA grant from FCDO.

The views expressed are those of the authors and not necessarily those of the NIHR, NHS England, Public Health England or the Department of Health and Social Care. Funders had no role in the study design, collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

Conflict of Interests

Members of the OpenSAFELY collaboration employed by Oxford University are listed in the author affiliations, above. Public Health England Immunisation and Countermeasures Division has provided vaccine manufacturers with post-marketing surveillance reports which the companies are required to submit to the UK Licensing Authority in compliance with their Risk Management Strategy, and a cost recovery charge is made for these reports.

BG has received research funding from the Laura and John Arnold Foundation, the NHS National Institute for Health Research (NIHR), the NIHR School of Primary Care Research, the NIHR Oxford Biomedical Research Centre, the Mohn-Westlake Foundation, NIHR Applied Research Collaboration Oxford and Thames Valley, the Wellcome Trust, the Good Thinking Foundation, Health Data Research UK (HDRUK), the Health Foundation, and the World Health Organization; he also receives personal income from speaking and writing for lay audiences on the misuse of science. IJD has received unrestricted research grants and holds shares in GlaxoSmithKline (GSK).

Appendix A: National COVID-19 vaccination eligibility

COVID-19 vaccinations were prioritised according to advice from the Joint Committee on Vaccination and Immunisation (**Table A1**).

Table A1: Priority groups for vaccination advised by the Joint Committee on Vaccination and Immunisation (reproduced from the Green Book provisional guidance, 5 December 2020 update)[10]

Priority group	Risk group
1	Residents in a care home for older adults Staff working in care homes for older adults
2	All those 80 years of age and over Frontline Health and social care workers
3	All those 75 years of age and over
4	All those 70 years of age and over Clinically extremely vulnerable individuals (not including pregnant women and those under 16 years of age)
5	All those 65 years of age and over
6	Adults aged 16 to 65 years in an at-risk group*
7	All those 60 years of age and over
8	All those 55 years of age and over
9	All those 50 years of age and over

* 'at-risk' groups comprise individuals with any of: chronic respiratory disease including severe asthma, chronic heart disease including atrial fibrillation, peripheral vascular disease, a history of venous thromboembolism, chronic kidney disease, chronic liver disease, stroke, transient ischaemic attack, cerebral palsy, Down's syndrome, or other chronic neurological conditions which may compromise respiratory function, diabetes mellitus, immunosuppression due to disease or treatment including any history of haematological malignancy, morbid obesity, severe mental illness, adult carers, and younger adults in residential care settings.

Appendix B: The OpenSAFELY data platform

OpenSAFELY provides a secure software interface allowing the analysis of pseudonymised primary care patient records from England in near real-time within the I vendor's highly secure data centre, avoiding the need for large volumes of potentially disclosive pseudonymised patient data to be transferred off-site. This, in addition to other technical and organisational controls, minimises any risk of re-identification. Similarly pseudonymised datasets from other data providers are securely provided to the I vendor and linked to the primary care data. Descriptions of OpenSAFELY have been previously published,[30] and more information can be found on https://opensafely.org/.

Primary care records retrieved from the TPP SystmOne electronic health record system include diagnoses (Read 3 CTV3 and SNOMED coded), prescriptions (dictionary of medicines and devices, dm+d coded), basic sociodemographics and vital signs for 22 million individuals – approximately 40% of the English population. Data extracted by SystmOne have previously been used in medical research as part of the ResearchOne dataset.[31]

All data is held in a secure research environment hosted by TPP, which is a Tier 3 data centre, accredited to NHS Digital standards for centrally hosted clinical systems (ISO 27001 standard and IG Toolkit version 2). We received ethics approval to conduct the data linkage and analyses by the London – City & East Research Ethics Committee on the 2nd of April 2020 (REC reference: 20/LO/0651) and LSHTM Ethics Board (ref 21863). No further ethical or research governance approval was required by the University of Oxford but copies of the approval documents were reviewed and held on record.

Appendix C: Consideration of potential confounders as covariate

Risk factors for venous thromboembolism (VTE) were considered as potential confounders if they were also either criteria for vaccine prioritisation, or were factors known to be associated with variation in vaccination coverage. Factors associated with vaccination prioritisation or coverage were also considered as potential confounders if they could affect investigation or diagnosis of VTE.

Risk factors for venous thromboembolism

Key risk factors for VTE are summarised in **Table C1**.[32-34] Around 50% of venous thromboembolism (VTE) are provoked by a transient risk factor (such as recent hospital admission), 20% occur among people with active cancer, and the rest are unprovoked. Risk factors for VTE which were criteria for vaccine prioritisation include: recent hospital admission, increasing age, care home residence, recent cancer (which will be divided into haematological and non-haematological), history of VTE, severe obesity, chronic heart disease, chronic respiratory disease, chronic liver disease and chronic neurological conditions likely to cause leg paresis or severely reduced mobility. Specific risk factors for cerebral venous thrombosis (CVT) were also considered, and meningitis and encephalitis were identified as likely to affect timing of vaccination. Some potential confounders are not routinely recorded in electronic health records and so may be a source of residual confounding. These include short- and long-term reduced mobility, intravenous drug use, and current (ongoing) pregnancies.

Venous thromboembolism is also a complication of COVID-19 infection. In this study COVID infection was considered a potential mediator of the association of interest. Estimates adjusted for COVID-19 infection could be difficult to interpret. We therefore describe COVID-19 in the period before VTE rather than adjusting for it.

Factors associated with vaccine coverage or timing of vaccination

Vaccine coverage varies by age, sex, ethnicity, and deprivation, each of which may also affect healthcare attendance and access to VTE investigation and diagnosis during the pandemic. Timing of vaccination rollout by priority group varied among local health services, and local health service pressures may also delay or reduce ascertainment and recording of VTE (for example, if there is a delay to SUS coding), and so local area was also considered a potential confounder.

Secular trends in VTE diagnosis

There have been marked secular trends in recorded VTE incidence over the COVID-19 pandemic.[18] These could be driven by either or both of local healthcare pressures resulting in changes in attendance, investigation, diagnosis and recording over the pandemic and changing incidence of COVID as a risk factor for VTE. Both are likely to vary locally. In the cohort study with a concurrent comparator group, this could be controlled for by matching vaccinated to unvaccinated individuals by STP. In the SCCS and cohort study with a historical comparator group, this was not an option. We considered using local COVID-19-related emergency department attendance rates as a proxy for the confounding effect of the COVID-19 pandemic on healthcare pressures and exposure to SARS-CoV2. However, as COVID-19 is a mediator of any association between COVID vaccination and VTE, this would complicate interpretation of the association. Instead, we will compare findings from the cohort study using concurrent and historical comparator groups, and may use quantitative bias analysis to explore the likely impact of secular trends in VTE recording on the estimates with a historical comparator group.

Table C1: Risk factors for VTE and likely or known association with COVID-19 vaccination [37]

Risk factor for VTE	Association with vaccination	Include as potential confounder
Permanent (or risk is unlikely to be return to normal once present)		
Age	Criterion for vaccine prioritisation	✔
Ethnicity	Associated with vaccine coverage	✔
Deprivation	Associated with vaccine coverage	✔
Nursing or care home residence	Criterion for vaccine prioritisation, and vaccine programme included specific delivery for residential care settings	✔
History of VTE	Criterion for vaccine prioritisation	✔
Obesity	Severe obesity was a criterion for vaccine prioritisation	✔
Central vein catheterisation or transvenous pacemaker	No, except by virtue of underlying medical conditions particularly chronic cardiac disease	chronic cardiac disease
Neurologic disease with leg paresis	Criterion for vaccine prioritisation	✔
Congestive heart failure	Criterion for vaccine prioritisation	✔
Atrial fibrillation	Criterion for vaccine prioritisation	✔
Respiratory failure	Chronic respiratory disease was a criterion for vaccine prioritisation	✔
Varicose veins	No	no
Family history of venous thromboembolism	No	no
Smoking	No, may be associated with uptake	✔
Long-term immobility	Likely to be associated with access to timely vaccination, but not routinely fully recorded in the electronic health records	not available
SLE/ Behçet's disease	Commonly treated with immunosuppressive medications which were a criterion for vaccine prioritisation	✔
Inflammatory bowel disease	No, but often treated with immunosuppressive medications which were a criterion for vaccine prioritisation	✔
Renal disease	Criterion for vaccine prioritisation	✔
Antiphospholipid syndrome and other prothrombotic conditions	No, unless history of VTE	no
Medium term (level of risk unlikely to change during the course of the study period if present at baseline)		
Cancer	Indication for vaccination if recent immunosuppressing treatment such as chemotherapy, or if any history ever of haematological cancer	✔
Lower limb fracture or joint replacement	No, other than hospitalisation may have expedited vaccination if already eligible and admitted with an incident event, while an event in the recent past may reduce mobility, affecting access to services for less timely vaccination	hospital admission
Iron deficiency anaemia	No	no
Thyroid disease	No	n0
Hormone replacement therapy	No, but may be related to uptake by association with social factors such as occupation	✔
Combined oral contraceptive pill	No, but may be related to uptake by association with social factors such as occupation	✔
Pregnancy and postpartum	During the study period, vaccination advice for pregnant women evolved. Ascertainment of pregnancies using electronic health records relies to a considerable extent on completed pregnancies (eg. records of a delivery) and ongoing pregnancies are likely to be under-ascertained from electronic health records.	exclude and describe
NSAIDs, antiplatelets and anticoagulants (preventive)	No, but the indications for these medications include criteria for vaccine prioritisation, and they may also affect investigation and ascertainment of thrombotic events	✔
ACE inhibitors, angiotensin receptor blockers, calcium channel blockers (protective)	No, but the indications for these medications include criteria for vaccine prioritisation	✔
Transient level of risk expected to vary within the study period		

Hospital admission	Vaccine delivery was prioritised to inpatients in December and this may have expedited vaccination for eligible individuals (>80 years)	✔
Lumbar puncture	No, other than hospitalisation may have expedited vaccination if already eligible	hospital admission
Short term bed rest	May have reduced access for timely vaccination but not routinely recorded in electronic health records	not available
COVID-19	A potential mediator for the association of vaccination with VTE. Small if any association of previous COVID-19 infection with vaccine uptake among eligible individuals in early priority groups.	no, but describe
Dehydration	Not recorded in electronic health records	not available

References

1. Medicines and Healthcare products Regulatory Agency (MHRA). *Regulation 174 Information for UK healthcare professionals on COVID-19 Vaccine Pfizer BioNTech*. 2020 10 December 2020 10 December 2020]; Available from: https://web.archive.org/web/20201207114708/https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/941452/Information_for_healthcare_professionals.pdf
2. Medicines and Healthcare products Regulatory Agency (MHRA). *Regulation 174 Information for UK healthcare professionals on COVID-19 Vaccine AstraZeneca*

2020 29 December 2020]; Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/948334/Information_for_UK_healthcare_professionals_on_COVID-19_Vaccine_AstraZeneca.pdf.
3. Medicines and Healthcare products Regulatory Agency (MHRA). *Summary of Product Characteristics for COVID-19 Vaccine Moderna*. 2021 19/04/2021 [cited 2021 30/04/2021]; Available from: https://www.gov.uk/government/publications/regulatory-approval-of-covid-19-vaccine-moderna/information-for-healthcare-professionals-on-covid-19-vaccine-moderna.
4. Larson, H.J., et al., *Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: a systematic review of published literature, 2007-2012.* Vaccine, 2014. **32**(19): p. 2150-9.
5. Larson, H.J., et al., *Tracking the global spread of vaccine sentiments: the global response to Japan's suspension of its HPV vaccine recommendation.* Hum Vaccin Immunother, 2014. **10**(9): p. 2543-50.
6. Medicines and Healthcare products Regulatory Agency (MHRA), *MHRA Board Meeting held in public: Item 5 Assurance Report - MHRA Review of COVID-19 Vaccine Submission Readiness*. 2020.
7. Medicines and Healthcare products Regulatory Agency (MHRA). *MHRA response to Danish, Norwegian and Icelandic authorities' action to temporarily suspend the AstraZeneca COVID-19 vaccine*. 2021 30/04/2021]; Available from: https://www.gov.uk/government/news/mhra-response-to-danish-authorities-action-to-temporarily-suspend-the-astrazeneca-covid-19-vaccine
8. Scully, M., et al., *Pathologic Antibodies to Platelet Factor 4 after ChAdOx1 nCoV-19 Vaccination.* New England Journal of Medicine, 2021.
9. Pottegård, A., et al., *Arterial events, venous thromboembolism, thrombocytopenia, and bleeding after vaccination with Oxford-AstraZeneca ChAdOx1-S in Denmark and Norway: population based cohort study.* BMJ, 2021. **373**: p. n1114.
10. Public Health England, *Immunisation against Infectious Disease (the Green Book). Chapter 14a: COVID-19. (Provisional guidance subject to MHRA approval of vaccine supply)*.

11. NHS England. *Coronavirus (COVID-19) Research Platform*. Available from: https://www.england.nhs.uk/contact-us/privacy-notice/how-we-use-your-information/covid-19-response/coronavirus-covid-19-research-platform/.
12. Remschmidt, C., O. Wichmann, and T. Harder, *Frequency and impact of confounding by indication and healthy vaccinee bias in observational studies assessing influenza vaccine effectiveness: a systematic review.* BMC Infect Dis, 2015. **15**: p. 429.
13. Petersen, I., I. Douglas, and H. Whitaker, *Self controlled case series methods: an alternative to standard epidemiological study designs.* BMJ, 2016. **354**: p. i4515.
14. Maclure, M., et al., *When should case-only designs be used for safety monitoring of medical products?* Pharmacoepidemiol Drug Saf, 2012. **21 Suppl 1**: p. 50-61.
15. Donegan, K., B. King, and P. Bryan, *Safety of pertussis vaccination in pregnant women in UK: observational study.* Bmj, 2014. **349**: p. g4219.
16. Lewis, J.D., et al., *The relationship between time since registration and measured incidence rates in the General Practice Research Database.* Pharmacoepidemiol Drug Saf, 2005. **14**(7): p. 443-51.
17. Mansfield, K.E., et al., *COVID-19 collateral: Indirect acute effects of the pandemic on physical and mental health in the UK.* medRxiv, 2020: p. 2020.10.29.20222174.
18. Walker, A.J., et al., *Changes in the rate of cardiometabolic and pulmonary events during the COVID-19 pandemic.* medRxiv, 2021: p. 2021.02.17.21251812.
19. MacKenna, B., et al., *Trends, regional variation, and clinical characteristics of COVID-19 vaccine recipients: a retrospective cohort study in 23.4 million patients using OpenSAFELY.* medRxiv, 2021: p. 2021.01.25.21250356.
20. Abdul Sultan, A., et al., *The incidence of first venous thromboembolism in and around pregnancy using linked primary and secondary care data: a population based cohort study from England and comparative meta-analysis.* PLoS One, 2013. **8**(7): p. e70310.
21. Handley, J.D. and H.C. Emsley, *Validation of ICD-10 codes shows intracranial venous thrombosis incidence to be higher than previously reported.* Health Inf Manag, 2020. **49**(1): p. 58-61.
22. Schultze, A., et al., *Identifying Care Home Residents in Electronic Health Records - An OpenSAFELY Short Data Report [version 1; peer review: awaiting peer review].* Wellcome Open Research, 2021. **6**(90).
23. Whitaker, H.J., et al., *Investigating the assumptions of the self-controlled case series method.* Stat Med, 2018. **37**(4): p. 643-658.
24. Stowe, J., et al., *Bell's palsy and parenteral inactivated influenza vaccine.* Hum Vaccin, 2006. **2**(3): p. 110-2.
25. Farrington, C.P., et al., *Self-Controlled Case Series Analysis With Event-Dependent Observation Periods.* Journal of the American Statistical Association, 2011. **106**(494): p. 417-426.
26. NHS Digital. *Data Security and Protection Toolkit*. 2020 October 6, 2020]; Available from: https://digital.nhs.uk/data-and-information/looking-after-information/data-security-and-information-governance/data-security-and-protection-toolkit.
27. NHS Digital. *BETA - Data Security Standards*. 2020 October 6, 2020]; Available from: https://digital.nhs.uk/about-nhs-digital/our-work/nhs-digital-data-and-technology-standards/framework/beta---data-security-standards.
28. NHS Digital. *ISB1523: Anonymisation Standard for Publishing Health and Social Care Data*. 2020 October 6, 2020]; Available from: https://digital.nhs.uk/data-and-information/information-standards/information-standards-and-data-collections-including-extractions/publications-and-notifications/standards-and-collections/isb1523-anonymisation-standard-for-publishing-health-and-social-care-data
29. Secretary of State for Health - UK Government. *Coronavirus (COVID-19): notification to organisations to share information*. 2020 October 6, 2020]; Available from:

https://www.gov.uk/government/publications/coronavirus-covid-19-notification-of-data-controllers-to-share-information.
30. Williamson, E.J., et al., *OpenSAFELY: factors associated with COVID-19 death in 17 million patients.* Nature, 2020.
31. Clegg, A., et al., *Development and validation of an electronic frailty index using routine primary care electronic health record data.* Age and Ageing, 2016. **45**(3): p. 353-360.
32. Heit, J.A., F.A. Spencer, and R.H. White, *The epidemiology of venous thromboembolism.* J Thromb Thrombolysis, 2016. **41**(1): p. 3-14.
33. Duffett, L., L.A. Castellucci, and M.A. Forgie, *Pulmonary embolism: update on management and controversies.* BMJ, 2020. **370**: p. m2177.
34. National Institute for Health and Care Excellence (NICE). *Clinical Knowledge Summary. Pulmonary embolism: What are the risk factors?* Clinical Knowledge Summaries 2020 [cited 2021 30/04/2021]; Available from: https://cks.nice.org.uk/topics/pulmonary-embolism/background-information/risk-factors/.