PASS Protocol	
Active substance	AZD1222
Product reference	D8111R00010
Version number	1.0
Date	15-June-2021

An Assessment of a relationship between the exposure to COVID-19 vaccines and risk of thrombotic thrombocytopenia syndrome

ATTEST study

(Association of the risk for Thrombotic Thrombocytopenia Syndrome and Exposure To COVID-19 vaccines)

Marketing Authorisation Holder(s)

Marketing authorisation holder(s)	[redacted]
MAH contact person	[redacted]

[redacted]

PASS INFORMATION

Title	*An assessment of a relationship between the exposure to COVID-19 vaccines and risk of thrombotic thrombocytopenia syndrome*
Protocol version identifier	1.1
Date of last version of protocol	15/June/2021
EU PAS register number	Study not yet registered
Active substance	*AZD12222*
Medicinal product	Vaxevria
Product reference	005675
Procedure number	EMEA/H/C/005675/MEA 005
Marketing authorisation holder(s)	[redacted]
Joint PASS	*No*
Research question and objectives	*To evaluate an association between COVID-19 vaccine exposure and thromboembolic events occurring with thrombocytopenia (thrombotic thrombocytopenia syndrome; TTS).*
Country of study	*United Kingdom (UK)*
Author(s)	[redacted]

[redacted]

1. TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

2. LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
Acute thrombosis	AT
Body mass index	BMI
Cerebral venous sinus thrombosis	CVST
Commissioning Data Set	CDS
Computerised medical record	CMR
COVID-19 Hospitalization in England Surveillance System	COVID-19 SARI-Watch/CHESS
Data Access Request Service	DARS
Data and Connectivity: COVID-19 Vaccines Pharmacovigilance network	DaCVaP
Deep vein thrombosis	DVT
Direct oral anticoagulants	DOACs
General practitioners	GP
GPES data for pandemic planning and research	GDPPR
High Dependency Units	HDU
Hormone replacement therapy	HRT
Hospital Episode Statistics	HES
Index of Multiple Deprivation	IMD
Integrated research application system	IRAS
Intensive Care Units	ICU
Joint Committee on Vaccination and Immunisation	JCVI
Low weight molecular heparin	LWMH
Medicines & Healthcare products Regulatory Agency	MHRA
Middle East respiratory syndrome	MERS-CoV
Myocardial infarction	MI
National Health Service	NHS
Nested case-control study	NCCS
Office for National Statistics	ONS
Oxford Royal College of General Practitioners Clinical Informatics Digital Hub	ORCHID
Oxford-Royal College of General Practitioners	RCGP
Platelet factor 4	PF4
Principal Investigator	PI

[redacted]

Abbreviation or special term	Explanation
Public Health England	PHE
Pulmonary embolism	PE
Relative incidence	RI
REporting of studies Conducted using Observational Routinely-collected Data	RECORD
Research and Surveillance Centre	RSC
Second Generation Surveillance System	SGSS
Self-controlled case series	SCCS
Strengthening the Reporting of Observational Studies in Epidemiology	STROBE
Thrombocytopenia	TP
Thromboembolism	TE
Trusted research environment	TRE
United Kingdom	UK
Venous thromboembolism	VTE

3. RESPONSIBLE PARTIES

AstraZeneca Responsible Parties

Role	Name, title, qualifications
Scientific lead, epidemiology	
Operational lead	
Study physician	
Statistician	
Pharmacovigilance	
Pharmacovigilance	
Pharmacovigilance	

EU, European Union; PV, pharmacovigilance; QPPV, Qualified Person Responsible for Pharmacovigilance.

[redacted]

4. ABSTRACT

Background/Rationale: A very rare syndrome of thrombosis associated with low platelets has been reported in a few cases of recent exposure to COVID-19 vaccine. No causal association with COVID-19 vaccination has yet been established. As of 15 April 2021, over 29.4 million doses of the AstraZeneca vaccine had been administered in EU/EEA countries.[1] Medicines & Healthcare products Regulatory Agency (MHRA) had requested for all cases of thrombosis or thrombocytopenia occurring within 28 days of coronavirus vaccine to be reported via the online yellow card system.[1] This thrombotic thrombocytopenia syndrome seems to be affecting patients of [redacted] and [redacted]; at present there is no clear signal of risk factors. Up to and including 31 March 2021, the MHRA had received 79 United Kingdom (UK) reports of blood clotting cases alongside low levels of platelets following the use of the COVID-19 Vaccine AstraZeneca. The United Kingdom (UK) is uniquely placed to study this area because of its registration-based primary care system, and a unique identifier umber links primary care to secondary care data. Additionally, vaccination is well advanced maximising population wide vaccine exposure.

Objectives: To evaluate an association between COVID-19 vaccine exposure and thromboembolic events occurring with thrombocytopenia (thrombotic thrombocytopenia syndrome; TTS).

Methods:

Study design: Two primary study designs will be considered, a case control study and a self-controlled case series (SCCS). A cohort analysis will be considered, in addition or as an alternative to either of the primary study designs, pending feasibility assessment of the follow-up time.

Data Source(s): We will conduct retrospective studies using linked secondary databases in England. Data for the definitive study accessed through the NHS Digital Trusted Research Environment (TRE), providing national data coverage. Primary care data will be linked with vaccination, hospitalization, COVID-19 test results, mortality data. Initial exploratory analyses will be conducted using the [redacted]

Study Population: Subjects of interest are people who [redacted]. However, we will required access to data from patients, in the integrated health records of [redacted] at the start of each study period. We need to make comparisons with unvaccinated, and our analysis needs information about levels of infection across the whole population.

Variables: Demographic, socioeconomic descriptors, clinical diagnosis, vaccines, potential confounders, thrombotic phenomena, COVID-19 infection, and other characteristics as applicable.

[redacted]

Statistical Analysis: Counts of cases will be reported by [redacted]. The study period will start on December 02, 2020, when the COVID-19 vaccine was first approved in the UK and will end approximately by end in May or at the end of data availability.

For the case-control design, all cases of TTS will be matched with control using risk-set matching. A case or matched control will be considered exposed if a COVID-19 vaccination will be within risk interval prior to the diagnosis date of the matched case.

For the SCCS, estimates of relative incidence (RI), in all risk intervals will be reported. RI will also be reported in the Joint Committee for Vaccinations and Immunisation's (JCVI's) vaccination risk and age groups. Seasonal effects are not deemed relevant. Cumulative hazard plots for gap times will be provided, and we will investigate if the inclusion of pre-exposure risk periods affects incidence of adverse events of interest associated with vaccination using a sequence of values and exploring whether RI is sensitive to the choice of pre-exposure period.

For the retrospective cohort design, the incidence of TTS after receiving COVID-19 vaccine will be estimated and we will compare this incidence with that occurring in an unvaccinated comparator group. We will look at the concurrent unvaccinated and retrospective comparator groups.

5. AMENDMENTS AND UPDATES

Table 1 Amendments and updates

Number	Date	Section of study protocol	Amendment or update	Reason
1			None	

[redacted]

6. MILESTONES

Table 2 Study milestones

Milestones	**Planned dates**
Study protocol approved	15 June 2021
Clinical study report approved	
Operational information	**Details**
Budget holder(s), including cost-sharing	[redacted]
Delivery model	[redacted]
Planned data re-use	Not applicable
Approach towards patient centricity (e.g. engagement with patient groups related to study)	Not applicable
International coordinating investigator or executive steering committee	[redacted]

7. RATIONALE AND BACKGROUND

The current COVID-19 pandemic constitutes a public health emergency unprecedented in the last century. From a small cluster of cases initially identified in 2019 in Wuhan, China,[2] the disease has spread around the globe. As of 4 April 2021, there have been more than 130 million confirmed cases and more than 2.8 million deaths globally.[3] SARS-CoV-2 (the viral agent that causes COVID-19) shares more than 79% of its sequence with SARS-CoV and 50% with the coronavirus responsible for Middle East respiratory syndrome (MERS-CoV).[4] It is believed that evolution of the pandemic will further vary across countries, affected in part by different containment strategies ranging from extreme lockdown to relative inaction. As a result, there may be regional waves of the disease and pockets of deeply affected populations. Globally, governments have acknowledged that an effective vaccine against COVID-19 constitutes a major public health need and may be the only way to guarantee a safe and sustained exit strategy from human movement restrictions while avoiding escalating mortality rates across populations. Accelerated development of safe and effective vaccines and treatments are currently underway and focus on adults. Various vaccines against COVID-19 based on different technologies are now in clinical development around the world or have been recently authorized, with some being rolled in middle or high-income countries. Many of the vaccines display optimal immune

[redacted]

response and protection following 2 vaccinations, a prime and a boost vaccination separated by several weeks. As of 15 April 2021, over 29.4 million doses of the AstraZeneca vaccine had been administered in EU/EEA countries. [1]

A rare syndrome of thrombosis associated with low platelets has been reported in a few cases of recent exposure to COVID-19 vaccine. No causal association with COVID-19 vaccination has yet been established. The cases are unusual because, despite the thrombocytopenia, there is progressive thrombosis, primarily venous, with a high preponderance of cerebral venous sinus thrombosis. Some arterial thrombotic events have also been noted. Testing typically reveals low fibrinogen and very raised D-dimer levels above the level typically expected in venous thromboembolism. Antibodies to platelet factor 4 (PF4) have been identified, hence there are similarities to heparin-induced thrombocytopenia despite the absence of prior exposure to heparin treatment. This syndrome seems to be affecting patients of all ages and both genders; at present there is no clear signal of risk factors. However, there is little data regarding occurrence and risk factors of thromboembolism (TE) with thrombocytopenia (TCP) or its relationship with prior COVID-19 infection or COVID-19 vaccination. Several papers reported case series of thrombosis co-occurring with thrombocytopenia, also called thrombotic thrombocytopenia syndrome (TTS) or vaccine induced thrombotic thrombocytopenia.[5] These represented relatively small numbers of patients without clear characterization and without a comparison of the risk of events among non-vaccinated populations.

The Medicines and Healthcare products Regulatory Agency (MHRA) had requested for all cases of thrombosis or thrombocytopenia occurring within 28 days of coronavirus vaccine to be reported via the online yellow card system.[6] Up to and including 31 March 2021, the MHRA had received 79 United Kingdom (UK) reports of blood clotting cases alongside low levels of platelets following the use of the COVID-19 Vaccine AstraZeneca: 44 of the 79 cases were of cerebral venous sinus thrombosis (CVST) with thrombocytopenia 35 of the 79 cases were of thrombosis in other major veins with thrombocytopenia. These 79 cases occurred in 51 women and 28 men, aged from 18 to 79 years.[7] As of Apr 9, 5 cases of arterial thrombosis, 374 cases of deep vein thrombosis, 3 cases of jugular vein thrombosis, 7 cases of pelvic venous thrombosis, 1 subclavian vein thrombosis, 345 cases of thrombosis, one case of vena cava thrombosis, 21 cases of venous thrombosis and 22 cases of venous thrombosis limb, have being reported in the in the EudraVigilance database.[8]

Benefits are well-established while risks are very low, so it should continue to be used. In addition, recently drafted cases definitions being posed by the Brighton Collaboration currently suggest the following diagnoses are also of interest in defining probable cases: deep vein thrombosis (DVT), pulmonary embolism (PE), stroke, and myocardial infarction (MI).[9]

The UK is one of the first countries that are introducing mass vaccination campaign for COVID-19 and has completed vaccinating the older half of the adult population starting from the oldest age groups, by May 2021vaccien was being offered to people [redacted].[10] Three COVID-19 vaccines were licensed and are being used including the: Moderna, BioNTech/Pfizer, and Oxford/Astrazeneca vaccines. Vaccination with the BioNTech/Pfizer vaccine started in December 2020 and with the Oxford/AstraZeneca vaccine started in early January 2021. Also, The UK is uniquely placed to study this area because of its registration-based primary care system (one patient registers with a single general practitioner), and a unique identifier National Health Service (NHS) number links primary care to secondary care data.

All general practices have had electronic links to their pathology systems since 2004, so that haematology tests (including platelet counts) and SARS-CoV-2 test results from any location nationally are recorded in the patient's computerised medical record (CMR). However, although SARS-CoV-2 testing is now readily available, that was not the case at the start of the pandemic. [redacted] However, prior to this there is a strong association with clinicians making a clinical diagnosis of COVID-19 and mortality.[9] Additionally, vaccination is well advanced maximising population wide vaccine exposure. [redacted]

8. OBJECTIVE

To evaluate an association between COVID-19 vaccine exposure and the outcome of thrombosis with thrombocytopenia syndrome (TTS).

9. RESEARCH METHODS

9.1 Study design-General Aspects

A case control study and a self-controlled case series (SCCS) will be designed and run, with a cohort analysis being considered in addition or as an alternative primary study design, pending feasibility assessment of the follow-up time available. The observation period will generally start on 02 December 2020 (approval date of the first COVID-19 vaccine) to the TTS event date or end of data availability, estimated mid May 2021 or another censoring event. There are several considerations to take when selecting the final study design/s. Firstly, some assumptions of the SCCS could be violated: 1) the occurrence of an event of interest must not alter the probability of subsequent exposure; and 2) the occurrence of the event of interest must not censor or affect the observation period. This is because in case of TTS occurring after the first dose of vaccine, the second dose could not be administered and/or because of a possible high

[redacted]

fatality rate associated with TTS. Observation period will be censored at the time of fatal events. In conclusion, the case-control design is likely to be more appropriate as the main analysis and SCCS is proposed as a sensitivity analysis for the non-fatal cases using the extended method for curtailed post-event exposure.

The matched case-control study will be conducted to determine the likelihood of those presenting with thrombotic thrombocytopenia (TTS) events within a risk window of 28 days of receiving COVID 19 vaccination. Patient who fulfils the definition of thrombotic thrombocytopenia based on a distribution of TTS spontaneously reported cases for COVID-19 vaccine AstraZeneca (sensitivity windows of 42 days as risk window for many immune-mediated vaccine safety events and as per EMA recommendation. We will conduct exploratory analyses on the dataset with the definitive analysis conducted on the national data held in the NHS Digital TRE, to identify people who presented with thrombotic, and thrombocytopenia events over a period (02 December 2020 to end of data availability, estimated mid May 2021) following COVID-19 vaccination. Electronic records of cases will be checked for any previous recorded history of conditions consisting of TTS; several wash-out periods will be considered including prior 90 days or prior for 12 months prior to the date of the index episode. Those without a previous record will be deemed to have experienced an incident diagnosis of a TTS on that date. Each incident case will be matched by date of TTS event, . Diagnosis dates of the cases will be considered the index dates for the controls. Records of both cases and controls will then be checked for the 28 days (sensitivity period of42 days) before the index date for a previous ChAdOx1 and BNT162b2 or other COVID-19 vaccination. .

For the self-controlled case series (SCCS), only individuals with a TTS event recorded during the defined study time are included in the analysis. Event and exposure history will be obtained from the linked dataset; events primarily extracted from hospital data. We will control for potential delaying of vaccination by the occurrence of an event by using the pre-vaccination risk period -7 to -1 days. Furthermore, since venous thromboembolism (VTE) is a likely indication for administration of a vitamin K antagonist (warfarin) or direct oral anticoagulants (DOAC), SCCS models with both vaccination and anticoagulation as exposures are potentially subject to event-dependent exposures, including recent infection. In this case we will perform a subgroup analysis of those with historical anticoagulation (defined by a prescription for

anticoagulation 90-days prior to the beginning of the study use versus those with no such history. We may also use this as a time varying covariate in our SCCS analysis). Another subgroup analysis will include prior exposure to COVID-19 infection. The observation period will commence 42 days before the first administration of a COVID vaccine in England, this results in the date 26th October 2020. The SCCS risk periods will be defined as days -7--1, 0-2, and 3-28 (our later analysis will extend to 42 days) where day 0 is the day of vaccination. This post vaccine period proposed is hypothesised to capture the time period within which most VTE events occur. A control period after 1st dose of 42 days will be considered. Given that most vaccinated in the UK were given the two-vaccine disease with about 3-month (90 days) time interval, we will check, but will not expect a high probability of overlapping of risk period after 2nd dose.

For the retrospective cohort design, the incidence of TTS after receiving COVID-19 vaccine will be estimated and we will compare this incidence with that occurring in an unvaccinated comparator group. A cohort of concurrent matched subjects will be selected. The comparison cohort will comprise subjects who have not yet received any COVID-19 vaccine. . The rationale for matching on calendar date is to control for the temporal variation in the SARS-CoV-2 circulation (COVID-19 infection may affect risk of some AESIs) and expected changes in vaccinee characteristics over time. Propensity scores (PS) will be used to address confounding, balancing multiple covariates such as prior COVID-19 infection, comorbidities, and markers of health care utilisation between vaccinated and unvaccinated cohorts. Relative risk will be the main measure of effect, which will be estimated using incidence rate ratios (IRRs) or hazard ratios (HRs), depending on the analysis and outcome. The proposal to use concurrent, rather than historical, comparator (unvaccinated) subjects is because they are at risk for COVID-19, and COVID-19 may function as an important covariable in the analyses of several of the safety endpoints. The potential challenge with this approach is that during the pandemic, unvaccinated subjects in the comparator group may themselves become vaccinated over time. The rationale for not selecting historical pre-pandemic subjects for primary comparison pertains to concerns over non-comparability. The background rate of TTS may have changed during the pandemic because of COVID-19 or because of less intensive ascertainment due to access issues or health-seeking behaviours.

9.2 Variables and Data Sources

Table 3 lists the groupings of variables available for this study by data source. Exposure data are described in the Vaccinations category. Outcome data are described in the primary care, secondary care use and mortality (as recorded in primary care data). The rest of the categories contain data on potential confounding variables and effect modifiers.

Table 3 Data items/variables and data sources

Data category	Data item	Data source
	splenectomy	
	In addition, we will look to identify:	

[redacted]

Data category	Data item	Data source
	1.	
Vaccinations	Vaccine type	
	Vaccine dose	
	Vaccination date	
Death	Data of death	

9.2.1 Exposure of interest

First dose o accines recorded in RCGP RSC. Uncertain vaccine type after 1/1/2021 when both vaccines were available, will be excluded. Uncertain vaccines administered during December will be assigned to the People will be excluded with a second dose within the follow-up window.

9.2.2 Outcomes of interest

Thrombotic thrombocytopenia syndrome (TTS) case definition (general population)[6]

The precise definition of thrombotic thrombocytopenia syndrome (TTS) is challenging. We will start with a broad definition. Thrombocytopenia defined either by diagnosis or platelets recording within 7 days of Thromboembolism dg. These definitions and the planned sensitivity will be iteratively adapted after reviewing available data. For example, we may rather than use

a fixed platelet count threshold use a drop of platelet count below a certain threshold or change the time interval between TE and TCP.

Sensitivity Definitions include:

a. [redacted]

Operationally, see **Figure 1**, a patient will have a new* diagnostic code for thrombocytopenia (TCP) OR a new* platelet count [redacted], see Diagram:

a. Within 7 days from thromboembolism event, i.e., from 7 days prior to and up to 7 days after index event (based on observation that spontaneously reported TTS cases are largely based on diagnoses during hospitalisation)
b. (sensitivity period) within _7/+42 days from thromboembolism event (reflecting on strengthening the association between TE and TCP events)

*'new' TCP is defined by excluding any prior TCP diagnosis during prior 12 months OR platelet count 10 to less than 100 (150 for sensitivity) $x10^9$/L during prior 3 months.

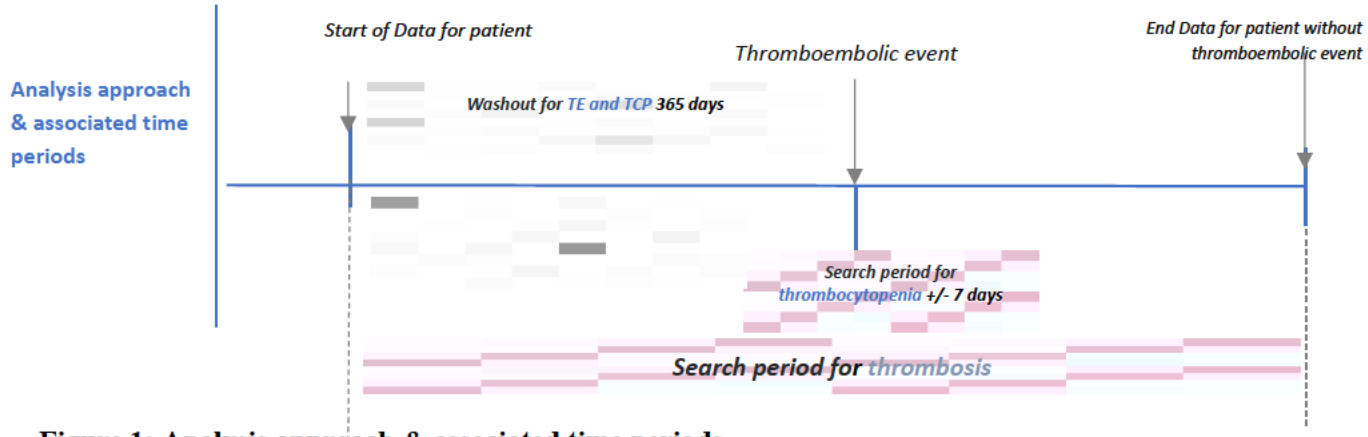

Figure 1: Analysis approach & associated time periods

9.3 Data sources

This study will employ linked England's national databases for data on [redacted] The datasets use unique identifiers allowing linkage across the datasets and avoidance of duplicates. We will assemble the required datasets for the analysis in the [redacted] for our initial feasibility analyses, then run the final analysis with the national data held in the [redacted].

[redacted]

provides access to data. Through this service it is possible to request access to several secondary data assets collected as part of routine care and commissioning activities in the . With appropriate approval, these data assets are linkable to the primary care data in), or the national dataset made available in their own). Data assets available through which may be used for this study include:

- GPES data for pandemic planning and research (GDPPR): central collection of GP patient data for COVID-19 purposes (fortnightly collection – All GP practices in England). only, the rest of these data will also be linked to data.

- – includes 2 associated datasets including COVID-19 Vaccination Status and COVID-19 Adverse Reaction.

- Hospital Episode Statistics (HES): this is the transformed data, initially part of the Commissioning Data Set (CDS), covering patients attending accident and emergency units, admitted for treatment, or attending outpatient clinics at hospitals in England. Statistical controls have been applied to the HES products. (Time lag 8-12 weeks).

- COVID-19 Second Generation Surveillance System (SGSS) - Demographic and diagnostic information from laboratory test reports for patients tested for COVID-19 in England only. It currently includes the first positive results from pillar 1 (swab testing in and hospital labs and pillar 2 (swab testing for the wider population) (Time lag – less than 1 week).

- Covid-19 UK Non-hospital Antigen Testing Results (Pillar 2) data includes a range of COVID19 test results, including . This is broadly like SGSS, but only covers Pillar 2 data, however, contains the full result set – i.e., all positive, negative, and null results.

- COVID-19 Hospitalization in England Surveillance System (COVID-19 SARI-Watch/CHESS): Epidemiological data on COVID-19 infection in persons requiring hospitalization in Intensive Care Units (ICU) or High Dependency Units (HDU). It records all demographic, risk factor, treatment and outcome information for patients admitted to hospital with a confirmed COVID-19 diagnosis (Time lag – less than 1 week).

- [redacted]

[redacted][12] is the secure data processing environment for [redacted] of GPs Clinical Informatics Digital Hub to extract and aggregate data from primary care computer systems and other parts of the health system for cross-sectional, cohort and case-control studies.[10] This [redacted]. The network has been expanded considerably during the SARS-CoV-2 pandemic. It pseudonymized data are linked to the range of data sets described above by [redacted], though usage must be approved on a study-by-study basis. It is one of the freshest (either daily or twice weekly data extracts) with data quality as good as it gets in primary care.

9.4 Study size

To be further developed using feasibility counts

In Interim, Table 1 shows the probability that the upper bound of the 95% confidence interval (CI) around the observed relative risk will be below 1.5, 2.0, 2.5, and 3.0 for various study sizes and TTS rates. These calculations assume a 1-to-1 ratio of vaccinated to comparator person-time and that the true IRR between those vaccinated and those unvaccinated is 1.0.

[redacted].

Table 4 Probability That the Upper 95% Confidence Limit of the Observed IRR Will Be Below 1.5, 2, 2.5, and 3 for Various Study Sizes of [redacted] Assuming True RR is 1 and Ratio of Vaccinated to Comparator Person-time is 1 to 1 and Duration of Risk Interval is 42 days

Estimated background rate per 100,000 PY	People vaccinated	Upper confidence limit of IRR			
		1.5	2.0	2.5	3.0
2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[redacted]

Note: calculations assume a true IRR of 1.0 and a ratio of 1 to 1 exposed to unexposed person-years. Calculations also assume that each person contributes to 84 days (42-day risk interval x 2 doses). Calculations follow the method of. IRR, incidence rate ratio; PY, person-years; RR, risk ratio. [13]

9.5 Data management

All statistical analyses will use the R statistical software, use of the SCCS library version 1.2 will be made.

9.6 Data analysis

9.6.1 Statistical Methods – General Aspects

A detailed Statistical Analysis Plan (SAP) will be developed for this analysis. The SAP may modify the plans outlined in the protocol; any major modifications of primary endpoint definitions or their analyses would be reflected in a protocol amendment. For all study designs, P-values will be quoted to two decimal places, unless they are less than 0.001 (whereby the p-value will be given as <0.001) or between <0.005 and >0.001, in which case they will be stated to three decimal places. Parameter estimates will be reported with 95% confidence intervals. All statistical tests are two sided and with a 5% significance level. All statistical analyses will use the R statistical software.[14]

9.6.2 Cohort design

[redacted]

Crude IRs and 95% CIs of each of the TTS will be estimated for [redacted] and comparator cohort. These analyses will primarily combine person-time at risk after dose 1 and dose 2 (if it was received). Exploratory analyses will report the IRs by specific dose, i.e. dose 1, dose 2. As TTS is a relatively new concept, the validity of risk windows, will be evaluated through cumulative incidence estimated using 1 − Kaplan-Meier survival curves starting at the index date.

For comparative design using (potentially, subject of decision when outputting cohort description) retrospective cohort analysis, we propose using time-to-event analyses. Cox regression models with robust estimation of the variance will be used to estimate crude and adjusted hazards ratios and 95% CIs. Using Cox regression models allows the flexibility to

[redacted]

potentially adjust for time-dependant covariates, such as intercurrent COVID-19 infection. Adjustment for confounding will occur using PS, either through PS matching or by analytic methods involving stratification or weighting, which will be specified in the SAP. If important covariates remain imbalanced after final selection of the cohorts for comparative analyses, additional approaches for confounding adjustment will be considered, including double-robust methods (ie, including the covariate in the outcome model as a separate term) or through analyses that restrict the population to subjects with more similar baseline characteristics (eg, subjects without substantial comorbidity).

9.6.3 Self-controlled case series (SCCS)

For comparative analysis using the SCCS approach, conditional Poisson regression will be used to estimate IRRs and 95% CIs. Estimates of relative incidence (RI), in all risk intervals and by age groups will be reported. . Exploratory analyses may stratify by dose, depending on available sample size. Time-invariant confounders will be inherently adjusted for using the SCCS design; however, time-varying confounders (e.g. recent prior infection) will be included as covariates in regression models. [5]

9.6.4 Matched case-control

For matched case-control analysis, the sampled risk sets will be represented matched strata, which will be included in conditional logistic regression models to estimate odds ratios (OR) for the risk of the event in the 42-day risk period. . Confounding may to be due to time varying confounders that will be much harder to adjust for than in our case-control study. We will also have confounding due to the differential availability of the vaccines, their storage requirements and intervals between doses (initially short and applied mainly to one brand) and then lengthened. We may be able to adjust for those that are measured and complete; as well as recognising others may be unmeasured (such as occupation). We will stratify exposure by dose as an exploratory analysis.

9.7 Subgroup analyses

A subgroup analysis will be conducted Further subgroup analysis will be considered . Further subgroup analysis will be of those

9.8 Sensitivity analyses

To reduce the potential for ascertainment bias, we will carry out a sensitivity analysis with a censoring date prior to the first media reports of possible [redacted] (i.e., March 07, 2021). We will consider exploring the use of different time intervals in the SCCS analysis following administration of the vaccine to define risk intervals to determine robustness of model estimates. Cases newly exposed to low weight molecular heparin (LWMH) direct oral anticoagulants (DOACs) or Warfarin in the post vaccination period, will be reported to explore if there are potentially missed cases.

9.9 Exploratory analyses

Artificial intelligence (AI) methods offer an alternative to conventional medical statistical approaches, through permitting potentially complex interactions between input covariates; permitting non-linear associations between inputs and risk; providing estimates of uncertainty in output classifications; in providing patient-specific estimates of risk (rather than population-based estimates); and in an improved ability to handle missing or artefactual data. AI-based approaches (enumerated below) will be compared with the conventional statistical approaches (enumerated below), using the same input data, for (i) predicting outcomes listed above and (ii) determining whether phenotypical groups may be defined using the input covariates that correspond to different risk models. This comparison will be undertaken using standard classification metrics (AUC, sensitivity-specificity, precision-recall, F1-score, etc.).

These AI-based approaches will include the major classes of models: ensemble methods (random forests, XGBoost, gradient-boosted trees), kernel methods (support vector machines), and deep learning methods (of varying architectures). Visualisation (via tSNE / UMAP algorithms) will be performed to facilitate interpretation of the outputs of the AI models. Feature selection and feature relevance will be undertaken to determine the primary risk factors learned by each of the AI models, for comparison against those risk models determined using conventional statistical methods. We emphasise that these are associations, and not causal, relationships – directly comparable to those of conventional medical statistical approaches.

9.10 Falsification Analysis

We will consider conducting a falsification analysis (negative controls) using a fictional date of exposure that is two months prior to receiving the first vaccine dose in order to estimate the influence of variables that are not directly associated with immunological effects of the vaccine (e.g., behaviour changes associated with the rollout programme) on our estimates.

9.11 Missing data

[redacted]

[redacted]

9.12 Bias

Possible bias may include but not limited to:

- Population differences vaccination cohort vs general population may limit validity of matching.
- [redacted]
- Death before event of TTS occurs.
- Outcome misclassification: Outcome definition relies on the accuracy of codes and algorithms to identify outcomes and the data available in each data source.
- Uncertainty about risk periods may lead to misclassification and potential attenuation of risk estimates.

9.13 Strengths and limitations of the research methods

Strengths:

The use of the national registries of all vaccination, COVID-19 testing, hospitalization, and mortality in England allow for the largest possible sample size and a near complete assessment of exposure, outcomes, and covariates. A nested case-control design can provide valid and precise estimates of associations and is a cost-effective alternative for full-cohort analysis.

Limitations:

The main limitation is related to the case definition, at this moment there is no clear agreement on an acute thrombosis (AT) with thrombocytopenia (TP) case definition. Another limitation is related to the diagnostic codes, the existence of two coding vocabularies, as well as multiple codes for the same medical concept within these vocabularies as well as the data lags, there are 7 days lag in recording into GP records. It must be noted that most general practices started lab-links (electronic in 2003-2004 in preparation for the start of the pay-for-performance scheme, the [redacted] which was initiated on 31st March 2004. Additionally, electronic transfer of records between practices (GP2GP) started in 2007; with large scale reliable transfer from 2011 (there has been some data loss as practices moved to the now most popular GP computerised medical record systems suppliers.

10. STUDY CONDUCT

10.1 Study Flow Chart and Plan

See section 2.1 for the study flow chart

10.2 Procedures

Not applicable for secondary data studies

10.3 Quality Control

The Principal Investigator (PI) is responsible for ensuring protocol compliance in accordance with AZ standards of quality. The PI may implement activities that could include but are not limited to:

- ensure appropriate storage of programming codes, code book, variables' definition etc
- double programming of the data analysis to ensure high quality of data analyses and avoidance of coding errors
- confirm that the research team is complying with the protocol

The Principal Investigator will ensure that appropriate training relevant to the Observational Study is given to investigational staff, and that any new information relevant to the performance of this Observational Study is forwarded to the staff involved.

10.4 Reporting guidelines and conventions

Results will be reported according to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) and REporting of studies Conducted using Observational Routinely collected Data (RECORD) (via the COVID-19 extension) guidelines.

10.5 Protection of Human Subjects

The study protocol is to be approved by the Integrated research application system (IRAS) system, by applying the filters for commercially funded and research database.

10.6 Subject Informed Consent (Primary Data Collection Only)

Not applicable

10.7 Confidentiality of Study/Subject Data (Primary Data Collection Only)

Not applicable

10.8 Collection and Reporting of Adverse Events/Adverse Drug Reactions

Not applicable since this is a secondary data study.

REFERENCES

1. (ECDC), E.C.f.D.P.a.C. *COVID-19 Vaccine Tracker*. 2021 15 Apr 2021]; Available from: https://qap.ecdc.europa.eu/public/extensions/COVID-19/vaccine-tracker.html#distribution-tab.
2. Zhou, P., et al., *A pneumonia outbreak associated with a new coronavirus of probable bat origin*. Nature, 2020. **579**(7798): p. 270-273.
3. Organization, W.H., *Coronavirus disease (COVID-19) weekly epidemiological update - 4 April 2021 – Overview*. . 2021, WHO.
4. Lu, R., et al., *Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding*. Lancet, 2020. **395**(10224): p. 565-574.
5. Chan, B.T., et al., *Meta-Analysis of Risk of Vaccine-Induced Immune Thrombotic Thrombocytopenia Following ChAdOx1-S Recombinant Vaccine*. medRxiv, 2021: p. 2021.05.04.21256613.
6. Haematology, B.S.f., *Guidance produced from the Expert Haematology Panel (EHP) focussed on syndrome of Thrombosis and Thrombocytopenia occurring after coronavirus Vaccination*. 2021.
7. Agency, M.H.p.R. *Research and analysis Coronavirus vaccine - weekly summary of Yellow Card reporting*. 2021 4 April 2021]; Available from: https://www.gov.uk/government/publications/coronavirus-covid-19-vaccine-adverse-reactions/coronavirus-vaccine-summary-of-yellow-card-reporting.
8. Agency, E.M. *European database of suspected adverse drug reaction reports*. 2021 April 9, 2021].
9. Collaboration, B., *CASE DEFINITION OF THROMBOSIS AND THROMBOEMBOLISM*. 2021.
10. NHS. *Who can get the coronavirus (COVID-19) vaccine*. 2021; Available from: https://www.nhs.uk/conditions/coronavirus-covid-19/coronavirus-vaccination/who-can-get-the-vaccine/.
11. *COVID-19 - SARS-CoV-2*. 2021.
12. Engbers, M.J., A. van Hylckama Vlieg, and F.R. Rosendaal, *Venous thrombosis in the elderly: incidence, risk factors and risk groups*. J Thromb Haemost, 2010. **8**(10): p. 2105-12.
13. Rothman, K.J. and S. Greenland, *Planning study size based on precision rather than power*. Epidemiology, 2018. **29**(5): p. 599-603.
14. Team, R.C. *R: A language and environment for statistical computing*. . 2020; Available from: https://www.R-project.org/.
15. Farrington, C.P., H.J. Whitaker, and M.N. Hocine, *Case series analysis for censored, perturbed, or curtailed post-event exposures*. Biostatistics, 2009. **10**(1): p. 3-16.

[redacted]

Appendix A List of stand-alone documents

None

[redacted]

Appendix B ENCePP checklist for study protocols

Study title:
An assessment of a relationship between the exposure to COVID-19 vaccines and risk of thrombotic thrombocytopenia syndrome

EU PAS Register® number:
Study reference number (if applicable):

Section 1: Milestones	**Yes**	**No**	**N/A**	**Section Number**
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection[1]	☒	☐	☐	9.1
1.1.2 End of data collection[2]	☒	☐	☐	9.1
1.1.3 Progress report(s)	☐	☒	☐	
1.1.4 Interim report(s)	☐	☒	☐	
1.1.5 Registration in the EU PAS Register®	☐	☒	☐	
1.1.6 Final report of study results.	☐	☒	☐	

Comments:

Section 2: Research question	**Yes**	**No**	**N/A**	**Section Number**
2.1 Does the formulation of the research question and objectives clearly explain:	☒	☐	☐	8
2.1.1 Why the study is conducted? (e.g. to address an important public health concern, a risk identified in the risk management plan, an emerging safety issue)	☒	☐	☐	8
2.1.2 The objective(s) of the study?	☒	☐	☐	8
2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	9.1

[1] Date from which information on the first study is first recorded in the study dataset or, in the case of secondary use of data, the date from which data extraction starts.

[2] Date from which the analytical dataset is completely available.

[redacted]

Section 2: Research question	**Yes**	**No**	**N/A**	**Section Number**
2.1.4 Which hypothesis(-es) is (are) to be tested?	☐	☒	☐	
2.1.5 If applicable, that there is no *a priori* hypothesis?	☐	☐	☒	

Comments:

Section 3: Study design	**Yes**	**No**	**N/A**	**Section Number**
3.1 Is the study design described? (e.g. cohort, case-control, cross-sectional, other design)	☒	☐	☐	9.1
3.2 Does the protocol specify whether the study is based on primary, secondary or combined data collection?	☒	☐	☐	9.1
3.3 Does the protocol specify measures of occurrence? (e.g., rate, risk, prevalence)	☒	☐	☐	9.6
3.4 Does the protocol specify measure(s) of association? (e.g. risk, odds ratio, excess risk, rate ratio, hazard ratio, risk/rate difference, number needed to harm (NNH))	☒	☐	☐	9.6
3.5 Does the protocol describe the approach for the collection and reporting of adverse events/adverse reactions? (e.g. adverse events that will not be collected in case of primary data collection)	☐	☒	☐	

Comments:

Section 4: Source and study populations	**Yes**	**No**	**N/A**	**Section Number**
4.1	☒	☐	☐	
	☒	☐	☐	
	☒	☐	☐	
	☒	☐	☐	
	☒	☐	☐	
	☒	☐	☐	
	☒	☐	☐	9

[redacted]

Comments:

Section 5: Exposure definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
5.1 Does the protocol describe how the study exposure is defined and measured? (e.g. operational details for defining and categorising exposure, measurement of dose and duration of drug exposure)	☒	☐	☐	9.2.2
5.2 Does the protocol address the validity of the exposure measurement? (e.g. precision, accuracy, use of validation sub-study)	☒	☐	☐	9.8 and 9.9
5.3 Is exposure categorised according to time windows?	☒	☐	☐	9.2.2
5.4 Is intensity of exposure addressed? (e.g. dose, duration)	☒	☐	☐	9.2.2
5.5 Is exposure categorised based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	☐	☐	☒	
5.6 Is (are) (an) appropriate comparator(s) identified?	☐	☐	☒	

Comments:

Section 6: Outcome definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
6.1 Does the protocol specify the primary and secondary (if applicable) outcome(s) to be investigated?	☒	☐	☐	9.2.2
6.2 Does the protocol describe how the outcomes are defined and measured?	☒	☐	☐	9.2.2
6.3 Does the protocol address the validity of outcome measurement? (e.g. precision, accuracy, sensitivity, specificity, positive predictive value, use of validation sub-study)	☒	☐	☐	9.2.2
6.4 Does the protocol describe specific outcomes relevant for Health Technology Assessment? (e.g. HRQoL, QALYs, DALYS, health care services utilisation, burden of disease or treatment, compliance, disease management)	☐	☐	☒	

Comments:

[redacted]

Section 7: Bias	**Yes**	**No**	**N/A**	**Section Number**
7.1 Does the protocol address ways to measure confounding? (e.g. confounding by indication)	☐	☐	☒	
7.2 Does the protocol address selection bias? (e.g. healthy user/adherer bias)	☐	☐	☒	
7.3 Does the protocol address information bias? (e.g. misclassification of exposure and outcomes, time-related bias)	☒	☐	☐	9.12

Comments:

To reduce the potential for ascertainment bias, we will carry out a sensitivity analysis with a censoring date prior to the first media reports of possible TTS events associated with AZ vaccine (i.e., March 07, 2021).

Section 8: Effect measure modification	**Yes**	**No**	**N/A**	**Section Number**
8.1 Does the protocol address effect modifiers? (e.g. collection of data on known effect modifiers, sub-group analyses, anticipated direction of effect)	☒	☐	☐	9.7

Comments:

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.1 Does the protocol describe the data source(s) used in the study for the ascertainment of:	X			9.2
9.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview)	☒	☐	☐	9.2
9.1.2 Outcomes? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics)	☒	☐	☐	9.2
9.1.3 Covariates and other characteristics?	☒	☐	☐	9.1
9.2 Does the protocol describe the information available from the data source(s) on:				
9.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☒	☐	☐	9.2
9.2.2 Outcomes? (e.g. date of occurrence, multiple event, severity measures related to event)	☒	☐	☐	9.2

[redacted]

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.2.3 Covariates and other characteristics? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	☒	☐	☐	9.1
9.3 Is a coding system described for:				
9.3.1 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	☐	☒	☐	
9.3.2 Outcomes? (e.g. International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))	☐	☒	☐	
9.3.3 Covariates and other characteristics?	☐	☒	☐	
9.4 Is a linkage method between data sources described? (e.g. based on a unique identifier or other)	☒	☐	☐	9.3

Comments:

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Section Number**
10.1 Are the statistical methods and the reason for their choice described?	☒	☐	☐	9.6
10.2 Is study size and/or statistical precision estimated?	☒	☐	☐	9.6
10.3 Are descriptive analyses included?	☒	☐	☐	9.6
10.4 Are stratified analyses included?	☒	☐	☐	9.6
10.5 Does the plan describe methods for analytic control of confounding?	☒	☐	☐	9.6
10.6 Does the plan describe methods for analytic control of outcome misclassification?	☐	☒	☐	
10.7 Does the plan describe methods for handling missing data?	☒	☐	☐	9.11
10.8 Are relevant sensitivity analyses described?	☒	☐	☐	9

Comments:

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Section Number**
11.1 Does the protocol provide information on data storage? (e.g. software and IT environment, database maintenance and anti-fraud protection, archiving)	☐	☒	☐	
11.2 Are methods of quality assurance described?	☐	☒	☐	

[redacted]

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Section Number**
11.3 Is there a system in place for independent review of study results?	☐	☒	☐	

Comments:

Section 12: Limitations	**Yes**	**No**	**N/A**	**Section Number**
12.1 Does the protocol discuss the impact on the study results of:				
12.1.1 Selection bias?	☐	☐	☐	
12.1.2 Information bias?	☐	☐	☐	
12.1.3 Residual/unmeasured confounding? (e.g. anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods).	☐	☐	☐	
12.2 Does the protocol discuss study feasibility? (e.g. study size, anticipated exposure uptake, duration of follow-up in a cohort study, patient recruitment, precision of the estimates)	☐	☐	☐	

Comments:

The study strengths and limitations are described I section 9.13

Section 13: Ethical/data protection issues	**Yes**	**No**	**N/A**	**Section Number**
13.1 Have requirements of Ethics Committee/ Institutional Review Board been described?	☒	☐	☐	10.4
13.2 Has any outcome of an ethical review procedure been addressed?	☒	☐	☐	10.4
13.3 Have data protection requirements been described?	☒	☐	☐	10.4

Comments:

Section 14: Amendments and deviations	**Yes**	**No**	**N/A**	**Section Number**
14.1 Does the protocol include a section to document amendments and deviations?	☒	☐	☐	5

Comments:

[redacted]

Section 15: Plans for communication of study results	**Yes**	**No**	**N/A**	**Section Number**
15.1 Are plans described for communicating study results (e.g. to regulatory authorities)?	☐	☒	☐	
15.2 Are plans described for disseminating study results externally, including publication?	☐	☒	☐	

Comments:

Name of the study scientific lead of the protocol: _____

Date: dd/Month/year

Signature: _____