Janssen Vaccines & Prevention B.V.*

Non-interventional Post Authorisation Safety Study – Protocol

An Observational Post-Authorisation Safety Study to Assess the Safety of Ad26.COV2.S Using European Healthcare Data through VAC4EU

Observational Study to Assess the Safety of Ad26.COV2.S

Protocol VAC31518COV4003

AMENDMENT [1]

Ad26.COV2.S (JNJ-78436735)

* Janssen Vaccines & Prevention B.V. is a Janssen pharmaceutical company of Johnson & Johnson and is hereafter referred to as the sponsor. The sponsor is identified on the Contact Information page that accompanies the protocol.

Status:	Approved		
Protocol version:	3.0	**Version date:**	9 December 2022
Prepared by:	Janssen Vaccines & Prevention B.V		
EDMS number:	EDMS-RIM-401552		

Compliance: This study will be conducted in compliance with the protocol and applicable regulatory requirements.

PASS INFORMATION

Title:	**An Observational Post-Authorisation Safety Study to Assess the Safety of Ad26.COV2.S Using European Healthcare Data through VAC4EU**
Protocol version:	3.0
Date of last version of the protocol:	05 May 2022*
EU PAS Register No:	EUPAS45362
Active substance (INN common name):	COVID-19 vaccine (Ad26.COV2-S[recombinant])
Pharmaco-therapeutic group (ATC Code):	J07BX03
Medicinal product(s):	JCOVDEN
Product reference:	EMEA/H/C/005737
Procedure number:	EMEA/H/C/005737/MEA/TBD
Name of Marketing Authorisation Holder(s)	Janssen-Cilag International NV
Joint PASS	No
Research question and objectives	This study aims to assess the risk of developing pre-specified adverse events of special interest (AESIs) within disease-specific risk windows following the administration of the JCOVDEN vaccine.
Country(-ies) of study	Italy, Spain, and The Netherlands.
Author	PPD

*Following the submission of the study protocol dated 3 September 2021 to the EMA, no substantial updates have been made to the protocol resulting in a protocol version 2.0 dated 05 May 2022

MARKETING AUTHORISATION HOLDER(S)

Name of Marketing Authorisation* Holder:	Janssen-Cilag International NV
Address:	Turnhoutseweg 30 B-2340 Beerse Belgium
Contact Details:	Corinne Willame, MPH; Cwilla57@its.jnj.com

*Emergency use authorisation granted in the United States on 27 February 2021; Conditional Marketing Authorisation (cMA) granted in the European Union on 11 March 2021; The UK Medicines and Healthcare products Regulatory Agency (MHRA) granted a cMA on 28 May 2021.

Qualified Person Pharmacovigilance:

Name:	PPD
Signature:	
Date:	

RESPONSIBLE PARTIES

Principal Participating Physician:	Not applicable
Coordinating Physician:	Not applicable
Contact person for this protocol:	PPD
E-mail address of contact person:	PPD

Principal Investigator(s) and Contributors to the Protocol

Name	Job title	Affiliation	E-mail address
PPD			

Name	Job title	Affiliation	E-mail address
PPD			

Data Partner Coordinating Investigators

Name	Job title	Affiliation	E-mail address
PPD			

Other Responsible Parties: Scientific Advisory Board members VAC4EU

Name	Job title	Affiliation	E-mail address
PPD			

1. TABLE OF CONTENTS

LIST OF IN-TEXT TABLES AND FIGURES

TABLES

FIGURES

2. LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviations

ACIP	Advisory Committee on Immunization Practices
ADR	Adverse Drug Reaction
ADVANCE	Accelerated Development of VAccine benefit-risk Collaboration in Europe
AE	Adverse Event
AESI	Adverse Event of Special interest
ARS	Agenzia Regionale di Sanità
CDC	Centers for Disease Control and Prevention
CDM	Common Data Model
CI	Confidence Intervals
CLS	Capillary Leak Syndrome
cMA	Conditional Marketing Authorisation
COVID-19	Coronavirus Disease 2019
CPRD	Clinical Practice Research Datalink
CVST	Cerebral Venous Sinus Thrombosis
DAP	Database Access Provider
DRE	Digital Research Environment
ECDC	European Centre for Disease Prevention and Control
EMA	European Medicines Agency
EMR	Electronic Medical Record
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
ESI	Emerging Safety Issue
ETL	Extraction, Transform, and Load
EU	European Union
EUA	Emergency Use Authorisation
EUPI	European Union Product Information
FDA	Food and Drug Administration
GBS	Guillain-Barré Syndrome
GP	General Practitioners
GPP	Good Pharmacoepidemiology Practices
GVP	Good Pharmacovigilance Practices
ICD	International Classification of Diseases
ICPC	International Classification of Primary Care
IRB	Institutional Review Board
IRR	Incidence Rate Ratio
IT	Italy
MHRA	Medicines and Healthcare Products Regulatory Agency
NNH	Number needed to harm
PASS	Post-Authorisation Safety Study
PPV	Positive Predictive Value
PRAC	Pharmacovigilance Risk Assessment Committee
QC	Quality Control
RNA	Ribonucleic Acid
SAP	Statistical Analysis Plan
SARS	Severe Acute Respiratory Syndrome
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus-2
SAS	Statistical Analysis Software
SCRI	Self-Controlled Risk Interval
SIP	Population Information System
SNOMED	Systematized Nomenclature of Medicine
SPEAC	Safety Platform for Emergency vACcines
SQL	Structured Query Language
SmPC	Summary of Product Characteristics
TTS	Thrombosis with Thrombocytopenia Syndrome
UK	United Kingdom

UMCU	University Medical Centre Utrecht
US	United States
VAC4EU	Vaccine Monitoring Collaboration for Europe
VIPIT	Vaccine-induced Prothrombotic Immune Thrombocytopenia
VITT	Vaccine-induced Immune Thrombotic Thrombocytopenia
VRBPAC	Vaccines and Related Biologics Products Advisory Committee
VTE	Venous Thromboembolism
WHO	World Health Organization

Definition of Term(s)

Acute event	Event expected to be recorded within 60 days of vaccination
Non-acute event	Events expected to be recorded >60 days after vaccination
Registry	An organised system that uses observational study methods to collect uniform data (clinical and other) to evaluate specified outcomes for a population defined by a particular disease, condition, or exposure, and that serves one or more predetermined scientific, clinical, or policy purposes
Study	The term “study” indicates the collection of data for research purposes only. The use of this term in no way implies that any interventional treatments or procedures, planned or otherwise, have been provided or performed
Retrospective observational study	A study that has all information collected from source data or a retrospective database. Normally, there is no new collection of information from the individual, although this may be required to address specific questions. Studies/Programs/Related Research Activities with only one visit can be considered prospective or retrospective bearing in mind this definition and the source of information
Post-Authorisation Safety Study (PASS)	Any study relating to an authorized medicinal product conducted with the aim of identifying, characterising, or quantifying a safety hazard, confirming the safety profile of the medicinal product, or of measuring the effectiveness of risk management measures
Washout period	The period of time required to more plausibly identify incident cases
Buffer period	The period of time required between risk and control windows in the SCRI to minimise potential residual risk attributable to the vaccine
Algorithm	A method to ascertain events using a combination of terms and concepts

3. AMENDMENTS AND UPDATES

Protocol Version	Date
Original Protocol	03 September 2021
Non-substantial amendment	05 May 2022
Amendment 1	09 December 2022

Amendments below are listed beginning with the most recent amendment.

Protocol section	Brief description of the change	Reason
Document cover & PASS information section	Protocol version 2.0 date was updated to 05 May 2022	Following the submission of the study protocol dated 3 September 2021 to the EMA, no substantial updates have been made to the protocol resulting in a protocol version 2.0 dated 05 May 2022.
Responsible parties section	PI and investigator lists were updated	The principal investigator for the study is now Fariba Ahmadizar. The Data partner coordinating investigator list has been updated.
Section 2 List of abbreviations and definitions of terms	List of abbreviations updated	To reflect all protocol amendments and updates.
Section 4 Abstract	Updated	To reflect all protocol amendments and updates.
Section 5 Milestones Section 8.1.1 Study design/Feasibility Assessment	Feasibility period, the number of reports, and data extractions have been reduced	Ad26.COV2.S uptake has reached a plateau in the EU, and it is not expected to see an increase in exposure except for data access lag. Therefore, the feasibility period will be shortened by removing one feasibility report, and the Interim reports. The study will produce 2 feasibility reports and a final report that will require only one additional data extraction.
Section 5 Milestones	The timelines of study deliverables have been adjusted	Timelines have been adjusted due to delays in the study setup and launch, which impacted the delivery of the Feasibility report 1. The date of the final report has been adjusted to reflect the complexity of case validations.
Section 6 Background and rationale	Minor editorial changes were made	Spelling, grammar updates, and harmonisation of the use of abbreviations.
Section 7 Research Question and Objectives	Minor editorial changes were made; Primary objective related to the safety evaluation study has been added; Booster dose assessment objective was added; AESI list has been reduced; Asthma Exacerbation was added as an endpoint; Prior COVID-19 was added as a matching variable.	 This was not described in the previous version. Since protocol development, the use of booster doses in recipients of the first dose of Ad26.COV2.S has been approved. 17 AESI were prioritised by the Sponsor for the first feasibility report. Prior COVID-19 infection is an important confounder for most of the studied AESIs.

Protocol section	Brief description of the change	Reason
Section 8.2.1 Study setting, Section 8.4.4 Data sources/ARS, Section 8.8.8 Quality Control/ARS	ARS Toscana (IT) cannot participate in the study until further notice	Due to the ongoing revision of the procedure for data access in the Tuscany Region, ARS must suspend its activities concerning the re-use of data until the revision is complete.
Section 8.1 Study Design	Index date is the time of COVID-19 vaccination (time zero), and the follow-up time is the start of each risk window	This has been clarified in the entire protocol in line with the statistical analysis plan (SAP).
Section 8.1 Study Design	Non-vaccinated individuals were removed from the feasibility assessment	Only vaccinated individuals are included.
Section 8.1.2.2 Cohort design	This section and its corresponding figure (figure 2) were updated according to the SAP	Giving an example of matched cohorts at the index date will clarify the methodology applied.
Section 8.2 Setting and study population	Updated	To reflect all protocol amendments and updates.
Section 8.2.3 Study duration and follow up	The study period start date was changed from 01 December 2020 to the date when vaccination with Ad26.COV2.S began in each country	To have alignment in calendar time for each of the vaccines.
Section 8.2.1 Study setting, Section 8.4.4 Data sources, Section 8.8.8 Quality Control	CPRD (UK), HSD (IT), and Norwegian Health Registers (NO) data sources were discontinued from the study	Capture of the use of the Ad26.COV2.S vaccine is very limited in the UK, in the HSD data source in Italy, and in Norway.
Section 8.2.4.3	Inclusion criteria in cohort design has been adapted according to the SAP	
Section 8.3 Variables	Updated	To reflect all protocol amendments and updates.
Section 8.3.2 Study outcomes	Updated	A coding explanation was added and the outcome list was adjusted to the prioritised AESI list.
Section 8.3.4	COVID-19 Hospitalisation data has been classified as yes/no	Severity may be looked at in a later stage for the final report.
Section 8.7. Data Analysis	Incidence rate difference will not be estimated for the self-controlled risk interval (SCRI) design	Due to methodological reasons.
Section 8.7. Data Analysis	It was clarified that the final main SCRI analysis will be based on a post-vaccination control window	For the final analysis, all individuals are expected to have accrued enough follow-up to use a post-vaccination control window. A post-vaccination control window will also help to address the healthy vaccinee effect.
Section 8.7. Data Analysis	A new sensitivity analysis comparing Ad26.COV2.S with active comparators	This sensitivity analysis will address the different dosing schedules between Ad26.COV2.S and the comparator vaccines used in the analyses.

Protocol section	Brief description of the change	Reason
	with (1) only 1 dose and (2) a 2-dose strategy; Table 5 has been updated accordingly	
Section 8.7. Data Analysis	Analysis of negative control data will be done for the final analysis	This has been added to table 5.
Section 8.7. Data Analysis	Age group categories were adjusted to those specified in the SAP	
Annex 1	Is now the list of stand-alone documents	EnCEPP checklist for study protocols has been completed and shared as a stand-alone document.

4. ABSTRACT

Protocol Title: An Observational Post-Authorisation Safety Study to Assess the Safety of Ad26.COV2.S Using European Healthcare Data through VAC4EU

Sponsor's Responsible Epidemiologist: Corinne Willame, MPH, Associate Director Epidemiology

NOTE: The term "sponsor" used throughout this document refers to the entities listed in the Contact Information page(s), which will be provided separately.

Background and Rationale

SARS-CoV-2 has spread rapidly and globally since its emergence, causing Coronavirus Disease 2019 (COVID-19). The World Health Organization (WHO) declared that the outbreak constituted a public health emergency of international concern on 30 January 2020 and declared the outbreak to be a pandemic on 11 March 2020. Due to the global spread of COVID-19 pandemic, rapid development of a COVID-19 vaccine is a worldwide priority.

Per communication by the European Commission on October 15, 2020. the Commission has entered into agreements with individual vaccine producers on behalf of the Member States, purchasing and/or reserving the right to purchase vaccine doses under Advance Purchase Agreements. Member States and public health authorities should prepare to undertake vaccine effectiveness and safety studies via coordination by the European Medicines Agency (EMA) and the European Centre for Disease Prevention and Control, and specifically to prepare for participation in large-scale EU-wide effectiveness and safety monitoring studies.

Following the primary analysis of the Phase 3 study COV3001, the United States (US) Food and Drug Administration (FDA) granted an emergency use authorisation (EUA) for the Ad26.COV2-S vaccine for use in individuals 18 years of age and older on 27 February 2021. On 11 March 2021, the EMA granted conditional marketing authorisation (cMA) for the Ad26.COV2.S vaccine for use in individuals 18 years of age and older.

However, following a successful efficacy analysis that supports the issuance of a EUA, further evaluation of the Ad26.COV2-S vaccine is still needed, including observational studies that leverage health insurance claims databases and more precise estimation of vaccine effectiveness. Active surveillance of vaccines through additional pharmacovigilance activities such as observational studies should also be considered.

To fulfil the regulatory obligations, the sponsor plans to initiate a Post-Authorisation Safety Study (PASS) aiming to characterise and evaluate the safety profile of Ad26.COV2.S in a large population sample size and to inform the scientific community on AESIs that could be associated with the use of Ad26.COV2.S.

Research Question and Objectives

This study has two chronologically consecutive aims: 1) to conduct a feasibility assessment aiming to inform the safety evaluation study and 2) to assess the risk of developing pre-specified and newly identified AESIs following administration of Ad26.COV2.S vaccine.

Objectives

Feasibility Assessment

The feasibility assessment will include the following objectives (not all objectives may be assessed in all analyses):

- To provide a comprehensive overview of the methods for identification of COVID-19 vaccine exposure including provenance of data and linkage to vaccination registry

- To monitor the number of individuals exposed to any COVID-19 vaccine in each database and to investigate the risk of exposure misidentification. Both missing vaccination instances and misclassification of the type of vaccine will be assessed using the benchmark provided by the ECDC: https://vaccinetracker.ecdc.europa.eu/public/extensions/COVID-19/vaccine-tracker.html#uptake-tab)
- To compare individuals receiving different COVID-19 vaccines concerning demographics, clinical characteristics of recipients, potential confounders, and risk factors for the AESI
- To conduct time-to-onset analyses for the AESIs (including exacerbation of asthma) concerning time since vaccination without prespecified risk windows.
- To describe booster vaccine doses in individuals who received a first dose of Ad26.COV2.S vaccine.

Safety Evaluation Study

Primary objective

The primary objective of this safety evaluation study is to assess the potential association between the occurrence of predefined and newly identified AESIs and vaccination with Ad26.COV2.S compared to individuals unexposed to any COVID-19 vaccine, and to individuals exposed to other types of COVID-19 vaccines (mRNA or adenovirus platforms) within disease-specific risk periods, and among the vaccinated, to a control window within the same individual.

Secondary objectives

The secondary objectives for this study are to assess the potential association between the occurrence of predefined and newly identified AESIs and vaccination with Ad26.COV2.S in the following specific subpopulations (part of the main study population):

- Immunocompromised individuals,
- Pregnant women,
- Individuals who have a prior history of thrombotic events and/or thrombocytopenia,
- Prior SARS-CoV-2 (COVID-19) infection,
- Individuals with a prior history (ever) of the specific event more than a year before start of follow-up.

These results will be compared to COVID-19 vaccine unexposed individuals and, to individuals exposed to other types of COVID-19 vaccines (mRNA or adenovirus-based platforms) or compared to a control window within the same individual.

Endpoints

Feasibility Assessment

The endpoints for the feasibility assessment, including a monitoring phase, are:

- Quality assessment of data that has been extracted from the source data banks (completeness, logic and benchmarking between data sources and against external data)
- The number of doses and uptake of different COVID-19 vaccines by calendar time
- The demographic and morbidity characteristics of individuals receiving different COVID-19 vaccines
- Algorithms for AESI identification (including exacerbation of asthma)
- Incident rates and cumulative incidence (risk) of the selected AESIs and corresponding 95% confidence intervals (CIs), without prespecified risk windows

- Risk factors for AESI.

Safety Evaluation Study

The primary endpoints for the safety evaluation study are to estimate the risk of the selected AESIs listed below among individuals vaccinated with the Ad26.COV2.S vaccine and in corresponding unexposed individuals and individuals exposed to other types of COVID-19 vaccines (split between mRNA platform-based vaccines and adenovirus-based platforms), or during a control window within the same individual.

The contractual agreements with data partners will stipulate the ability to allow newly identified AESIs to be incorporated into the study objectives, and the current protocol that will be used for governance approvals will state that additional AESI may be added. Each data access provider is responsible to inform the appropriate ethics boards of these amendments.

Acute events (events expected to be recorded within 60 days of vaccination).

- The incidence of anaphylaxis within 0-2 days.
- The incidence of generalized convulsion, arrhythmia, acute kidney failure and acute hepatic failure within 1-14 days.
- The incidence of the following events within 1-28 days.
  - thrombotic events (microangiopathy including capillary leak syndrome), disseminated intravascular coagulation, deep vein thrombosis, pulmonary embolism, cerebral vein thrombosis, peripheral thrombosis, thrombosis with thrombocytopenia and ischemic strokes, acute coronary syndrome (acute myocardial infarction, unstable angina, and new onset angina)
  - bleeding events including haemorrhagic stroke (no subarachnoid)
  - composite endpoint: Venous thrombosis (including pulmonary embolism and deep vein thrombosis)
  - composite endpoint: Arterial thrombosis (including acute coronary syndrome and ischemic stroke)
  - composite endpoint: stroke (including haemorrhagic and non-haemorrhagic stroke)
  - heart failure, and stress cardiomyopathy.
- The incidence of immune/inflammatory events: encephalitis (including acute demyelinating encephalomyelopathy and meningoencephalitis), Guillain-Barré Syndrome, Bell's palsy, immune thrombocytopenia, thrombocytopenia, transverse myelitis, and cardiac inflammatory disorders (including myocarditis and pericarditis) within 1-42 days.
- The incidence of sensorineural hearing loss within 1-60 days.
- The incidence of asthma exacerbation within 1-14 days.

Non-acute events (events expected to be recorded >60 days after vaccination)

- The incidence of new onset autoimmune thyroiditis, multiple sclerosis, acute aseptic arthritis, and type 1 diabetes mellitus within 1-365 days.

Newly identified AESI may be incorporated to the list of AESI to be investigated by the current study if a safety signal occurs.

Study Design

The study will comprise a feasibility assessment for a period of 12 months and a safety evaluation study.

This study is a retrospective observational study using electronic healthcare databases of various types in Europe. Eligible individuals will be included in the study from the start of Ad26.COV2.S vaccination in each participating country, and the study will end at the last date of data availability in each database. The AESIs included in this study are considered potential or identified safety risks following the administration of Ad26.COV2.S. The selected AESIs represent a heterogeneous group, including multiple organ systems and acute and chronic conditions.

Feasibility assessment: A feasibility assessment focusing on the availability of Ad26.COV2.S vaccine data (vaccine uptake), characteristics of individuals vaccinated with different COVID-19 vaccines, as well as the measurement of AESIs will be conducted for each participating data source. From the start of the feasibility assessment, participating data sources will extract data twice (one extraction for the first feasibility assessment and one for the second feasibility and the final report). The primary design for the feasibility assessment will be a cohort study including all individuals with at least one day of follow-up after starting Ad26.COV2.S vaccination in each participating country.

Safety evaluation: Evaluation of safety concerns will be conducted using a retrospective observational study in electronic health care databases of various types in Europe. Eligible individuals will be included in the study from the start of Ad26.COV2.S vaccination in each participating country and the study will end at the last date of data availability in each database.

For each AESI, the study design of the safety evaluation study will depend on whether the AESI is an acute or non-acute event and follow the decision framework described in the ACCESS template protocols for evaluation of safety events in electronic health care databases.

The primary study design for the evaluation of acute events (events expected to occur within 2, 14, 28, 42, or 60 days of vaccination) will be a self-controlled risk interval (SCRI) design with pre- and post-vaccination control window, and for non-acute events (events expected to occur within 365 days) will be a cohort design with concurrent unexposed comparators. Additionally, the cohort analysis will also include two active comparator groups: one viral vector vaccine comparator group (i.e., Vaxzervria® [AZD1222] by Oxford/AstraZeneca) and one mRNA vaccine comparator group (i.e., Comirnaty® [tozinameran] by Pfizer/BioNTech and/or Spikevax® [elasomeran] by Moderna), where possible (pending on the feasibility analysis with a maximum follow-up time of one year). For acute events, in a sensitivity analysis, the SCRI design will use a pre-vaccination control window.

Additionally, a cohort design analogous to the non-acute events will be conducted. When it is established that data sources do not capture vaccinations well during the feasibility assessment, we will conduct a sensitivity analysis excluding these data sources from the cohort analysis with non-exposed comparators for the non-acute events, to avoid misclassification of exposure.

Individuals start follow-up at the start of each risk window and end follow-up at the earliest occurrence of latest data availability of the databank, individual exit from the database, completion of the study period, or death. At least one year of enrolment prior to the date of vaccination (index date) will be required to determine whether individuals meet the study criteria and to define baseline characteristics. If more historical data is available, this will be included.

Self-controlled risk interval design (primary study design for acute events)

The SCRI design will compare the risk of the AESI in a post-vaccination risk window to a pre-vaccination (sensitivity analysis) and post-vaccination control window (main analysis) within the same individual.

The SCRI design will include only individuals in the primary analysis who received at least one first-ever dose of the Ad26.COV2.S vaccine during the study period. Vaccinated individuals enter the study at the date of the start of the pre-vaccination control window. The SCRI design will compare the risk of each outcome during the post-vaccination risk window following the COVID-19 dose with a self-matched control interval that may be prior to vaccination, or after the vaccination, risk window to assess the baseline

risk of the outcome. The control window will have the same maximum length as the risk window to minimise time-varying confounding that cannot be measured. In case the follow-up time does not capture the maximum post-vaccination risk or control window due to right censoring, all available follow-up time will be utilized. Follow-up will be censored upon vaccination with another COVID-19 vaccine or end of follow-up. A washout period between the control and risk window will be applied to minimise capturing prevalent events during the risk window. A post-vaccination control window may induce a bias towards the null when an AESI identified during the vaccination risk window, is captured again during the control window due to a re-admission instead of a reoccurrence of the AESI. This situation is more likely to be encountered for acute potentially recurring events such as cardiac events (including acute coronary syndrome, myocarditis, pericarditis, arrhythmia, and stress cardiomyopathy). Moreover, due to lag-times in updating the relevant databanks, there is a higher risk that the post-vaccination control window may not capture all events and underestimate the rate of events. Use of a pre-vaccination control window will deal with these issues. As a sensitivity analysis for acute events, a cohort analysis (following the same approach used in the main analysis for non-acute events) will be conducted. For the final analysis, a pre- and post-vaccination control window will be studied separately. In a sensitivity analysis, vaccination to other non-COVID-19 vaccines will also be an exclusion and censoring event.

Cohort design (primary study design and non-acute events and sensitivity analysis for acute events)

For the primary analysis of non-acute events, a retrospective cohort design will be used to estimate the rate of vaccination with Ad26.COV2.S vaccine, describe the characteristics of these vaccinated individuals and subsequently to estimate the incidence of new-onset autoimmune thyroiditis, multiple sclerosis, acute aseptic arthritis, and type 1 diabetes mellitus after receipt of the vaccine dose and to compare this incidence with that occurring in an unvaccinated matched comparator group, and in two groups of individuals exposed to other types of COVID-19 vaccines (i.e., one viral vector vaccine comparator group vaccinated [i.e., Vaxzevria® {AZD1222}] by Oxford/AstraZeneca) and one mRNA vaccine comparator group (i.e., Comirnaty® [tozinameran] by Pfizer/BioNTech and Spikevax® [elasomeran] by Moderna).

- **Exposed cohort:** individuals will have received at least 1 dose of Ad26.COV2.S vaccine.
- **Concurrent unexposed cohort:** individuals that have not been vaccinated with Ad26.COV2.S or any other COVID-19 vaccines at any time prior to the index date (time zero) matched to the vaccinated individual for important characteristics.
- **Cohort exposed to other COVID-19 vaccines:** individuals will have received at least 1 dose of a viral vector COVID-19 vaccine (i.e., Vaxzevria® [AZD1222] by Oxford/AstraZeneca) or an mRNA COVID-19 vaccine (Comirnaty® [tozinameran] by Pfizer/BioNTech and Spikevax® [elasomeran] by Moderna, respectively).

In this retrospective cohort design, the index date (time zero) in the exposed cohorts (i.e., recipients of the vaccines) will be the day the first dose of the corresponding COVID-19 vaccination was received. This date in the unexposed group will be a day when an individual did not receive a COVID-19 vaccine dose and randomly chosen by calendar matching to the start time of the corresponding Ad26.COV2.S exposed individual (i.e., a random day during the same week that the matching individual in the exposed cohort receives the Ad26.COV2.S vaccine).

Individuals in the Ad26.COV2.S exposed cohort will be individually matched to one individual in the concurrent unexposed cohort and to one individual in each active comparator cohorts on key clinical variables (exact age, sex, prior COVID-19 infection, and presence of one or more risk factors for severe COVID-19 (e.g., cancer, sickle cell disease, obesity, chronic kidney disease, chronic respiratory disease, human immunodeficiency virus infection), and month of vaccination (using calliper if needed)) at the index date. Additional details on the matching process will be specified in the statistical analysis plan (SAP). A single individual may contribute person-time to the exposed and unexposed groups at different time points (details will be described in the SAP). Individuals will be classified into exposure groups that are compatible with their data at the index date. Follow-up under unexposed status is censored if an individual receives a

COVID-19 vaccine. In a sensitivity analysis, vaccination to other non-COVID-19 vaccines will also be an exclusion and censoring event.

Setting and Study Population

For the implementation of the feasibility assessment, electronic health care databases in Southern and Western Europe that have shown interest and are a member of the Vaccine Monitoring Collaboration for Europe (VAC4EU) will be used. The selected data sources and two-letter country codes are as follows:

- SIDIAP (Sistema d'Informació per el Desenvolupament de la Investigació en Atenció Primària) [Information System for Research in Primary Care] (ES)
- VID, the Foundation for the Promotion of Health and Biomedical Research of Valencia Region [FISABIO] (ES)
- PHARMO DATABASE NETWORK (PHARMO Institute for Drug Outcomes Research) (NL)
- ARS Toscana (Agenzia Regionale di Sanità della Toscana) [a research institute of the Tuscany region of Italy] (IT)]. See section 8.2.1 regarding conditional participation of ARS Toscana.

When needed (based on specific safety concerns occurring in specific countries, or because of size) additional data sources may be added for dedicated evaluation studies, if they have adequate data on exposure, outcomes and covariates.

Feasibility Assessment

For the feasibility assessment, the study period will start on the date when vaccination with Ad26.COV2.S began in each participating country or the latest date when the individual is registered in the data source (plus 365 days) or born and will end a maximum of 12 months thereafter. Individuals will be followed until the earliest of the following dates: death, end of data availability, individual exit from the data source, or the completion of the period.

Safety Evaluation Study

- For the SCRI design, follow-up starts at the start of the pre-vaccination control window following an AESI specific washout period. Follow-up ends at the earliest of the following: death, end of data availability, individual withdrawal of the study, end of the post-vaccination risk or control window, or receipt of another COVID-19 vaccine. We will compare the Ad26.COV2.S vaccine with active comparators with 1) only 1-dose strategy and 2) a 2-dose strategy.
- For the cohort design, the study period will start at the time when Ad26.COV2.S vaccination starts in each participating country and will end at the end of follow-up. For the cohort design, follow-up starts at the start of each risk window and ends at the occurrence of each AESI, death, end of data availability, individual exits the database, after one year of follow-up, receipt of a COVID-19 vaccine (unexposed cohort only), receipt of a different COVID-19 vaccine from the one granting access to the study (exposed cohorts only). For the vaccine-exposed comparator groups, a second dose of the same corresponding vaccine will be allowed (more details will be provided in the SAP). In a sensitivity analysis, follow-up in the cohort study will be censored upon receipt of other vaccines (non-COVID-19 vaccine). We will compare the Ad26.COV2.S vaccine with active comparators with 1) only 1-dose strategy and 2) a 2-dose strategy. For the subgroup analysis of subjects without COVID-19 at the start of each risk window, patients will be censored if they develop COVID-19 during follow-up.

Variables

Exposure assessment

Exposure will be based on available prescriptions, dispensing, or administration of the Ad26.COV2.S and other COVID-19 vaccines. Vaccine receipt and date of vaccination will be obtained from all possible sources that capture COVID-19 vaccination, such as pharmacy dispensing records, general practice records, immunisation registers, vaccination records, medical records, or other data banks. During the feasibility assessment, the completeness of information will be assessed and described. Depending on the data source, vaccines may be identified via nationally used product codes (including batch numbers) where possible. The exposure of interest for the safety evaluation study is the receipt of Ad26.COV2.S vaccine.

Study outcomes

AESIs, as listed below and in line with the definitions and code lists that for most of the AESIs have been created for the ACCESS project (https://zenodo.org/communities/vac4eu/) will be identified, with a date of diagnosis, using predefined validated algorithms (where available), based on diagnosis codes (with procedure and/or pharmacy dispensing codes and/or limited to specific medical care settings if applicable to the outcome). The impact of different provenance of data (hospital, general practitioner's diagnoses) and algorithms on the outcome frequency will be assessed and described in the feasibility assessment.

For the first feasibility assessment, the following prioritised AESIs will be studied. The list includes encephalitis, including acute disseminated myelitis (ADEM) and meningoencephalitis; Guillain-Barré syndrome; transverse myelitis; multiple sclerosis, including optic neuritis; immune thrombocytopenia; anaphylaxis; cardiac inflammatory disorders, including myocarditis and pericarditis; stress cardiomyopathy; coronary artery disease, including acute myocardial infarction; deep vein thrombosis; pulmonary embolism; disseminated intravascular coagulation; non-haemorrhagic stroke; haemorrhagic stroke; thrombocytopenia and acute hepatic failure. As per EMA PRAC's request, exacerbation of asthma will also be included in the feasibility assessment. For subsequent feasibility and final report, all AESIs will be analysed.

Evaluation of Safety Outcomes

The sponsor has created a list of AESIs based on current knowledge of the Ad26.COV2.S vaccine. Background incidence rates for most of the data sources and AESIs are available on the VAC4EU dashboard https://vac4eu.org/covid-19-tool/. Definitions and codes are available on Zenodo (https://zenodo.org/communities/vac4eu/).

Covariate definition

Feasibility Assessment

In the feasibility assessment, covariates will be assessed at the time of COVID-19 vaccine administration (index date).

Safety Evaluation Study

For the cohort analysis, covariates for the study populations will be defined at the index date. For the SCRI design, time-varying covariates will be identified both at the start of the risk window and the control window.

Covariates will be used to define and describe the study cohorts and populations of special interest and their baseline characteristics and/or to control for confounding. Additional covariates may be added during the conduct of the study if deemed important. The look-back observation period to define covariates of interest will be all available information before the index date in each data source with a minimum of 12 months, except when an alternative look-back window is specified.

Data Sources

The study will use data from secondary electronic health record databases that are population-based. All data sources will have the ability to provide data on COVID-19 vaccines, outcomes (diagnoses, procedures, laboratory results, and treatments), and important covariates. It is not currently known the extent to which COVID-19 vaccines, product types, and batch numbers will be captured in data sources. To be included in the study, data sources should preferably be updated at a minimum once every 3 months. At the proposal stage, members of VAC4EU were offered the option to participate in the study. Four data sources from 3 countries (Spain, the Netherlands and Italy) will be included in the study. A more detailed description will be included in the VAC4EU FAIR Catalogue.

When establishing the agreements with the data sources to conduct the study, it will be emphasized that the current list of AESIs may be expanded during the course of the study to accommodate newly identified AESIs.

Study Size

The study will be conducted in a source population of approximately 18.1 million individuals, although children will not be vaccinated. It is assumed that up to 10% will be vaccinated with the Ad26.COV2.S vaccine.

Main Statistical Methods

A general description of the planned statistical methods to be used to analyse the data collected in this study is presented in the main body of the document. Additional details will be provided in the SAP.

For the feasibility analysis, the utilisation patterns of Ad26.COV2.S and other COVID-19 vaccines will be characterised and monitored over time. Description of demographics and clinical characteristics will be reported overall and for different groups of vaccine recipients.

The primary analysis of non-acute AESIs will focus on calculating and comparing the incidence rates between individuals exposed to Ad26.COV2.S and:

1. unexposed individuals
2. individuals exposed to another viral vector COVID-19 vaccine (i.e., Vaxzevria® [AZD1222] by Oxford/AstraZeneca); and
3. individuals exposed receiving a mRNA COVID-19 vaccine (cohort).

For acute events, the relative risk between risk window and control window will be estimated (SCRI) among individuals exposed to Ad26.COV2.S.

All analyses will be conducted within each data source and pooled across data sources using a random-effects model.

For the COVID-19 vaccines with a 2-dose schedule, AESI-specific risk windows after each dose of the comparator will be considered for the cohort analyses.

Analyses within the primary study designs (cohort design for non-acute events and SCRI design for acute events) will be stratified by the clinically relevant subgroups below. Stratification will be done also for matching variables when considered relevant.

- Selected comorbidities, including risk factors for severe COVID-19 (by presence or absence of each comorbidity)
- Frailty score (categorized)

- Age (0-17, 18-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80+ years)
- Sex.

Sensitivity analyses will be performed to:

- Assess the risk of AESI within extended disease-specific risk windows for events for which the risk interval is not well known or documented in the vaccine safety literature.
- Assess the impact of data sources that are shown not to capture vaccinations well during the feasibility assessment. In the sensitivity analyses with non-exposed comparators those data sources from the cohort analysis for the acute and non-acute events will be excluded, to avoid misclassification of exposure.
- Address the different dosing schedules between Ad26.COV2.S and the comparator vaccines used in the analyses, comparing the Ad26.COV2.S vaccine with active comparators with (1) a 1-dose strategy and (2) a 2-dose strategy.
- Excluding and censoring upon the administration of other (non-COVID-19) vaccinations in the SCRI and cohort study.
- Conduct a SCRI analysis using a pre-vaccination control window for all acute AESIs.

5. MILESTONES

The initial planned dates for key milestones in this study are outlined below.

Milestone:	**Planned Date:**
Start of data collection[a]	30/04/2022*
Summary report feasibility assessment 1[b]	10/11/2022*
Summary report feasibility assessment 2[b]	Q3 2023
Interim report	No interim report will be generated for this study
Data extraction (querying) for final study report	Q1-Q2 2023
End of data extraction[c]	Q2 2024
Final report of study results to Janssen[d]	Q3 2024

* Actual dates.

[a] Start of data collection is the planned date for starting data extraction for the feasibility assessment.
[b] Summary reports will be generated according to feasibility assessment.
[c] End of data extraction is the planned date the analytical dataset (the minimum set of data required to perform the statistical analyses for the primary objective) is completely available including findings from the case validation process.
[d] Final report will be submitted by Janssen to the appropriate health authority agencies as per health authority agreed timelines.

6. BACKGROUND AND RATIONALE

6.1. Background

SARS-CoV-2 Virology and COVID-19 Disease Burden

The severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) is an enveloped, positive-sense, single-stranded ribonucleic acid (RNA) Beta coronavirus (Coronaviridae Study Group of the International Committee on Taxonomy of Viruses 2020, Wu 2020). It was first identified following reports of a cluster of acute respiratory illness cases in Wuhan, Hubei Province, China in December 2019 (Li Q 2020). Early epidemiological investigations suggested that the majority of early cases were linked to a food market, with patients infected through zoonotic or environmental exposure, followed by the subsequent spread of infection by human-to-human transmission among close contacts (Li Q 2020). However, there is some controversy about the initial origin of the virus (Cyranoski 2020). Genomic sequencing was performed on bronchoalveolar lavage fluid samples collected from patients with viral pneumonia admitted to hospitals in Wuhan, which identified a novel RNA virus from the family Coronaviridae (Lu 2020, Wu 2020). Phylogenetic analysis of the complete viral genome revealed that the virus, SARS-CoV-2, is part of the subgenus Sarbecovirus of the genus Betacoronavirus, and is most closely related (approximately 88% identity) to a group of severe acute respiratory syndrome (SARS)-like coronaviruses previously sampled from bats in China (Lu 2020).

As of 26 August 2021, approximately 212,418,6624 cases and 4,436,327 deaths from Coronavirus Disease 2019 (COVID-19) have been reported worldwide. Approximately 36,307,572 cases including 750,921 deaths have been reported in the EU/European Economic Area (ECDC 2020). In the United States (US) approximately 39,210,669 cases and 640,223 deaths have been reported (Johns Hopkins CSSE 2020). India is the South Asian country that has been affected the most by COVID-19. The country has reported more than 32.8 million cases and the death toll has passed 1 million. In April 2021 alone, India added almost six million new cases partly caused by the new Covid-19 variant, identified as B.1.617 (Yadav 2021). South Africa is the African country that has been affected the most by COVID-19 with 2,777,659 cases and 82,261 deaths reported (Johns Hopkins CSSE 2020).

Ad26.COV2.S Vaccine

Ad26.COV2.S (also known as Ad26COVS1, VAC31518, JNJ-78436735) is a monovalent vaccine composed of a recombinant, replication-incompetent adenovirus type 26 (Ad26) vector, constructed to encode the SARS-CoV-2 Spike (S) protein, stabilized in its prefusion conformation.

The S protein is the major surface protein of coronaviruses. Different animal models have been used for the evaluation of candidate vaccines against SARS-CoV, and the common conclusion that has emerged is that the viral S protein is the only significant target for neutralizing antibodies (Buchholz 2004, Sui 2005, Zhang 2004, Zhou 2004) and the only viral protein that can elicit protective immunity in animal models (Berry 2004, Bisht 2004, Bukreyev 2004, Subbarao 2004).

Ad26.COV2.S encodes a membrane-bound full-length S protein derived from a SARS-CoV-2 clinical isolate (Wuhan 2019, whole genome sequence NC_045512), with 2 amino acid changes in the S1/S2 junction that knock out the furin cleavage site, and 2 proline substitutions in the hinge region.

At the time of the primary analysis of Phase 3 study VAC31518COV3001 (data lock point 22 January 2021) safety analysis of Ad26.COV2.S demonstrated that the vaccine has an acceptable safety and reactogenicity profile when administered as a single dose or as a 2-dose regimen at dose levels up to $1x10^{11}$ viral particles (vp) in adults at 18 years of age or above, including adults at 60 years of age or above. There was a trend towards a decrease in the frequency of solicited local and systemic adverse events (reactogenicity) with a decreasing dose level of Ad26.COV2.S (from $1x10^{11}$ vp to $2.5x10^{10}$ vp). Reactogenicity was demonstrated to be transient and most solicited adverse events (AEs), including pyrexia, generally resolved in 1 to 2 days post-vaccination. No significant safety issues were identified. However, a numerical imbalance was observed for venous thromboembolism (VTE). In general, a lower reactogenicity profile was observed in older adults compared with younger adults (data on file).

On 27 February 2021, the US Food and Drug Administration (FDA) granted an emergency use authorisation (EUA) for the Ad26.COV2.S vaccine for use in individuals 18 years of age and older (FDA 2021). On 11 March 2021, European Medicines Agency (EMA) granted a conditional marketing authorisation (cMA) for Ad26.COV2.S for use in individuals 18 years of age and older (EMA 2021). The UK Medicines and Healthcare products Regulatory Agency (MHRA) granted a cMA)on 28 May 2021 (MHRA 2021).

As of 12 April 2021, more than 6.8 million doses of the Ad26.COV2.S vaccine have been administered in the US. Up to that date, 6 cases of cerebral venous sinus thrombosis (CVST) along with low platelet counts were reported (Cines 2021). In response to these cases, the US FDA and the Centers for Disease Control and Prevention (CDC) recommended a pause in the use of the Ad26.COV2.S vaccine in the US on 13 April 2021, to allow for further investigations on these rare events of thrombosis with thrombocytopenia. Similar cases of thrombosis with thrombocytopenia had also been reported in individuals who received the Vaxzervria/COVID-19 AstraZeneca Vaccine outside the US (Greinacher 2021, Schultz 2021, Scully 2021). This syndrome has been termed “vaccine-induced prothrombotic immune thrombocytopenia (VIPIT)” or “vaccine-induced immune thrombotic thrombocytopenia (VITT)” but is now named “thrombosis with thrombocytopenia syndrome (TTS)”. This rare and new syndrome is characterised by thrombosis (such as CVST or splanchnic thrombosis), thrombocytopenia and confirmed with a positive platelet factor 4 (PF4)-heparin enzyme-linked immunosorbent assay. There is currently no standard case definition for TTS accepted for use by all countries. In this regard, the Brighton Collaboration developed an interim case definition of TTS (Brighton Collaboration 2021). VAC4EU organised a webinar to discuss the operationalisation of this event, for which background rates were generated for EMA (Willame 2021). In addition, the PRAC’s recommendation for the categorisation of TTS will be taken into account.

In response to the pausing in administration of Ad26.COV2.S in the US, 2 emergency Advisory Committee on Immunization Practices (ACIP) meetings were convened to review these reported cases and to consider potential implications on the Ad26.COV2.S vaccination policy. Meanwhile, a series of 12 cases of CVST combined with low platelet counts, were identified through the Vaccine Adverse Event Reporting System in the US (Karron 2021). On 23 April 2021, ACIP concluded that the benefits of resuming Ad26.COV2.S vaccination among individuals aged ≥18 years outweighed the risks and reaffirmed its interim recommendation under FDA's EUA. The sponsor was requested to revise the Ad26.COV2.S vaccine fact sheet to include a warning for rare clotting events among women aged 18 to 49 years.

Similarly, the EMA Pharmacovigilance Risk Assessment Committee (PRAC) assessed cases of thrombosis after vaccination with Ad26.COV2.S vaccine. On 20 April 2021, the PRAC confirmed that the available data show that the Ad26.COV2.S vaccine's known and potential benefits outweigh its known and potential risks in individuals 18 years of age and older. The sponsor was asked to update the Summary of Product Characteristics (SmPC) and Package Leaflet to include important information on the diagnosis and management of TTS. On 6 May 2021, the PRAC recommended to further updating the SmPC and European Union (EU) Risk Management Plan including the addition of TTS as an important identified risk. Thromboembolism should be maintained as an important potential risk. In addition, EMA recommends that PASS should include adverse events of special interest (AESIs) of concern. The Brighton Collaboration Interim Case Definition for TTS should also be considered for case findings.

On 18 June 2021, an Emerging Safety Issue (ESI) notification was sent to EMA regarding an internally identified significant safety issue around cases of Capillary Leak Syndrome (CLS) reported with Ad26.CoV2.S.

- Following the AESI notification to EMA a Type 2 variation was submitted on 25 June 2021 to update the European Union Product Information (EUPI) with a contraindication for individuals who have previously experienced episodes of CLS. In addition, an updated direct healthcare professional communication to inform healthcare professionals of the contraindication for individuals who have previously experienced episodes of CLS was submitted to EMA on 25 June 2021. Upon EMA request wording related to CLS was also added to sections 4.4 and 4.8 of the EUPI. This variation was approved on 09 July 2021.
- In the US Janssen notified the FDA of the internally identified significant safety issue around cases of CLS reported with Ad26.CoV2.S. also on 18 June 2021 and submitted as per guidance obtained from the FDA a Clinical Overview which included an executive summary for CLS and a cumulative review that includes supportive data, narratives and Council for International Organizations of Medical Sciences on 25 June 2021.

On 7 July 2021, based on an FDA review of cases of Guillain-Barré syndrome (GBS) reported in recipients of the Janssen COVID-19 Vaccine, FDA required updates to the product label to include a warning statement and associated information regarding GBS following vaccination. Janssen agreed to consider GBS a significant safety issue. In addition, FDA provided comments regarding their review of capillary leak syndrome (CLS). FDA did not conclude that in individuals with a history of CLS, the risk of receiving the Janssen COVID-19 vaccine clearly outweighs any possible

benefit and the agency believes it is appropriate to include CLS in Section 6.2 Post Authorisation Experience of the Fact Sheet for Healthcare Providers Administering Vaccine (Vaccination Providers). On 8 July 2021, Janssen submitted updated factsheets to the EUA accepting the FDA positions on GBS and CLS. Approval was received on 12 July 2021.

On 12 July 2021, a type II variation was submitted to EMA to update the EUPI to include GBS as an Adverse Drug Reaction (ADR) in section 4.8 and to include related wording in Section 4.4 of the EUPI.

6.2. Overall Rationale for the Study

SARS-CoV-2 has spread rapidly and globally since its emergence, causing COVID-19. The World Health Organization (WHO) declared that the outbreak constituted a public health emergency of international concern on 30 January 2020 and declared the outbreak to be a pandemic on 11 March 2020. Due to the global spread of COVID-19 pandemic, rapid development of a COVID-19 vaccine is a worldwide priority.

Following the primary analysis of the Phase 3 study COV3001, the US FDA granted an EUA for the Ad26.COV2-S vaccine for use in individuals 18 years of age and older on 27 February 2021. However, following a successful efficacy analysis that supports issuance of a EUA, further evaluation of the Ad26.COV2-S vaccine is still needed, including observational studies that leverage health insurance claims databases and more precise estimation of vaccine effectiveness (Center for Disease Control and Prevention 2020). Active surveillance of vaccines through additional pharmacovigilance activities such as observational studies should also be considered.

Per communication by the European Commission on October 15, 2020 (European Commission 2020), the Commission has entered into agreements with individual vaccine producers on behalf of the Member States, purchasing and/or reserving the right to purchase vaccine doses under Advance Purchase Agreements (European Union 2020). Member States and public health authorities should prepare to undertake studies of vaccine effectiveness and safety via coordination by the EMA and the European Centre for Disease Prevention and Control (ECDC), and specifically to prepare for participation in large-scale EU-wide effectiveness and safety monitoring studies.

To fulfil its regulatory obligations, the sponsor plans to initiate a post-authorisation safety study (PASS) aiming to characterise and evaluate the safety profile of Ad26.COV2.S in a large population sample size and to inform the scientific community on AESIs that could be associated with the use of Ad26.COV2.S.

7. RESEARCH QUESTION AND OBJECTIVES

Research Question

This study has 2 chronologically consecutive aims: 1) to conduct a feasibility assessment aiming to inform the safety evaluation study and 2) to assess the risk of developing pre-specified and newly identified AESIs following administration of Ad26.COV2.S.

Objectives

Feasibility Assessment

The feasibility assessment will include the following objectives (not all objectives may be assessed in all analyses):

- To provide a comprehensive overview of the methods for identification of COVID-19 vaccine exposure including provenance of data and linkage to vaccination registry
- To monitor the number of individuals exposed to any COVID-19 vaccines (Ad26.COV2.S, Vaxzervria vaccine and mRNA COVID-19 vaccines) and to investigate the risk of misidentification of the COVID-19 vaccine exposure. Both missing vaccination instances and misclassification of the type of vaccine will be assessed with the benchmark provided by the ECDC: https://vaccinetracker.ecdc.europa.eu/public/extensions/COVID-19/vaccine-tracker.html#uptake-tab)
- To compare individuals receiving different COVID-19 vaccines with respect to demographics and clinical characteristics of recipients, potential confounders and risk factors for the AESI
- To conduct time-to-onset analyses for the sponsor prioritised AESIs (including exacerbation of asthma) with respect to time since vaccination and without prespecified risk windows
- To describe booster vaccine doses in individuals who received a first dose of Ad26.COV2.S.

The joint evaluation of the objectives mentioned above will inform on the overall feasibility of the study.

Safety Evaluation Study

Primary objective

The primary objective of this safety evaluation study is to assess the potential association between the occurrence of predefined and newly identified AESIs and vaccination with Ad26.COV2.S compared to individuals unexposed to any COVID-19 vaccine, and to individuals exposed to other types of COVID-19 vaccines (mRNA or adenovirus platforms) within disease-specific risk periods, and among the vaccinated, to a control window within the same individual.

Secondary objectives

The secondary objectives of this study are to assess the potential association between the occurrence of predefined and newly identified AESIs and vaccination with Ad26.COV2.S in the following specific subpopulations (part of the main study population):

- Immunocompromised individuals,
- Pregnant women,
- Individuals who have a prior history of thrombotic events and/or thrombocytopenia,
- Prior COVID-19 infection,
- Individuals with a prior history (ever) of the specific event more than a year before start of follow-up.

These results will be compared to COVID-19 vaccine unexposed individuals and, to individuals exposed to other types of COVID-19 vaccines (mRNA or adenovirus-based platforms) or compared to a control window within the same individual.

Endpoints

Feasibility Assessment

The endpoints for the feasibility assessment, including a monitoring phase, are:

- Quality assessment of data that has been extracted from the source data banks (completeness, logic and benchmarking between data sources and against external data)
- The number of doses and uptake of different COVID-19 vaccines by calendar time
- The demographic and morbidity characteristics of individuals receiving the different COVID-19 vaccines
- Algorithms for AESI identification (including exacerbation of asthma)
- Incident rates and cumulative incidence (risk) of he selected AESIs and corresponding 95% confidence intervals (CIs), without prespecified risk windows
- Risk factors for AESI.

Safety Evaluation Study

The primary endpoints for the safety evaluation study are to estimate the risk of the selected AESIs listed below among individuals vaccinated with the Ad26.COV2.S vaccine and in corresponding unexposed individuals and individuals exposed to other types of COVID-19 vaccines (split between mRNA platform-based vaccines and adenovirus-based platforms), or during a control window within the same individual.

The contractual agreements with data partners will stipulate the ability to allow newly identified AESIs to be incorporated into the study objectives, and the current protocol that will be used for governance approvals will state that additional AESI may be added. Each data access provider is responsible to inform the appropriate ethics boards of these amendments.

Acute events (events expected to be recorded within 60 days of vaccination).

- The incidence of anaphylaxis within 0-2 days.

- The incidence of generalized convulsion, arrhythmia, acute kidney failure and acute hepatic failure within 1-14 days.
- The incidence of the following events within 1-28 days
  - thrombotic events (microangiopathy including capillary leak syndrome), disseminated intravascular coagulation, deep vein thrombosis, pulmonary embolism, cerebral vein thrombosis, peripheral thrombosis, thrombosis with thrombocytopenia and ischemic strokes, acute coronary syndrome (acute myocardial infarction, unstable angina, and new onset angina),
  - bleeding events including haemorrhagic stroke (no subarachnoid),
  - composite endpoint: Venous thrombosis (including pulmonary embolism and deep vein thrombosis)
  - composite endpoint: Arterial thrombosis (including acute coronary syndrome and ischemic stroke),
  - composite endpoint: stroke (including haemorrhagic and non-hemorrhagic stroke)
  - heart failure, and stress cardiomyopathy.
- The incidence of immune/inflammatory events: encephalitis (including acute demyelinating encephalomyelopathy and meningoencephalitis), Guillain-Barré Syndrome, Bell's palsy, immune thrombocytopenia, thrombocytopenia, transverse myelitis and cardiac inflammatory disorders (including myocarditis and pericarditis) within 1-42 days.
- The incidence of sensorineural hearing loss within 1-60 days.
- The incidence of asthma exacerbation within 1-14 days.

Non-acute events (events expected to be recorded >60 days after vaccination)

- The incidence of new-onset autoimmune thyroiditis, multiple sclerosis, acute aseptic arthritis, and type 1 diabetes mellitus within 1-365 days.

Newly identified AESI may be incorporated into the list of AESI to be investigated by the current study if a safety signal occurs.

8. RESEARCH METHODS

8.1. Study Design

The study will comprise a feasibility assessment, during a period of 12 months, and a safety evaluation study.

8.1.1. Feasibility Assessment

A feasibility assessment focusing on the availability of Ad26.COV2.S vaccine data (vaccine uptake), characteristics of individuals vaccinated with different COVID-19 vaccines and the measurement of AESIs will be conducted for each participating data source. From the start of the feasibility assessment, participating data sources will extract data twice (one extraction for the first feasibility assessment and one for the second feasibility and the final report). Summary reports (feasibility assessment reports and the final report) will be produced at each extraction. The primary design for the feasibility assessment will be a cohort study including all individuals with at least one day of follow-up after the start of Ad26.COV2.S vaccination in each participating country.

8.1.2. Overview of Study Design for Safety Evaluation Study

Evaluation of safety concerns will be conducted using a retrospective observational study in electronic health care databases of various types in Europe. Eligible individuals will be included in the study from the start of Ad26.COV2.S vaccination in each participating country, and the study will end at the last date of data availability in each database.

For each AESI, the study design of the safety evaluation study will depend on whether the AESI is an acute or non-acute event and follow the decision framework described in the ACCESS template protocols for evaluation of safety events in electronic health care databases (Dodd 2020).

The primary study design for the evaluation of acute events (events expected to occur within 2, 14, 28, 42, or 60 days of vaccination) will be a self-controlled risk interval (SCRI) design with pre- and post-vaccination control window and for non-acute events (events expected to occur within 365 days) will be a cohort design with concurrent unexposed comparators. Additionally, the cohort analysis will also include two active comparator groups: one viral vector vaccine comparator group (i.e., Vaxzevria® [AZD1222] by Oxford/AstraZeneca) and one mRNA vaccine comparator group (i.e., Comirnaty® [tozinameran] by Pfizer/BioNTech and/or Spikevax® [elasomeran] by Moderna), where possible (pending on the feasibility analysis with a maximum follow-up time of one year).

In sensitivity analyses, an equivalent cohort design will also be used to assess acute events. When it is established that data sources do not capture vaccinations well during the feasibility assessment, we will conduct a sensitivity analysis excluding these data sources from the cohort analysis with non-exposed comparators for the non-acute events, to avoid misclassification of exposure.

Another sensitivity analysis will compare the Ad26.COV2.S vaccine with active comparators (1) a 1-dose strategy and (2) a 2-dose strategy.

Individuals start follow-up at the start of each risk window and end follow-up at the earliest occurrence of latest data availability of the databank, individual exit from the database, completion of the study period, or death. At least one year of enrolment prior to the index date (date of vaccination) will be required to determine whether individuals meet the study criteria and to define baseline characteristics. If more historical data is available, this will be included.

8.1.2.1. Self-controlled Risk Interval Design (Primary Study Design for Acute Events)

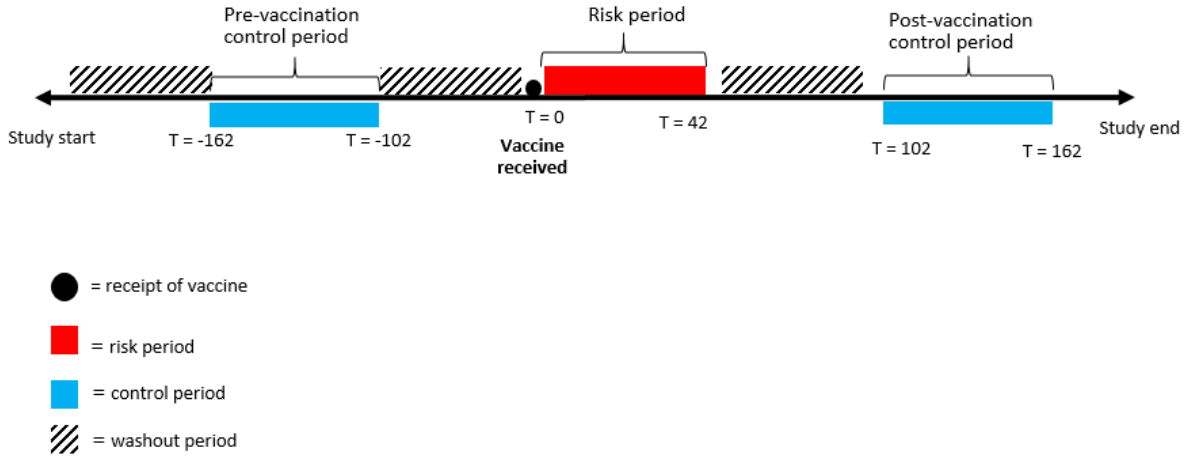
The SCRI design will compare the risk of the AESI in a post-vaccination risk window to a pre-vaccination control window (sensitivity analysis) and post-vaccination control window (main analysis) within the same individual.

The SCRI design will include only individuals in the primary analysis who received at least one first-ever dose of the Ad26.COV2.S vaccine during the study period. Vaccinated individuals enter the study at the date of the start of the pre-vaccination control window. The SCRI design will compare the risk of each outcome during the post-vaccination risk window following the COVID-19 dose with a self-matched control interval that may be prior to vaccination, or after the vaccination risk window, to assess the baseline risk of the outcome. The control window will have the same maximum length as the risk window to minimise time-varying confounding that cannot be measured. In case the follow-up time does not capture the maximum post-vaccination risk or control window due to right censoring, all available follow-up time will be utilized. Follow-up will be censored upon vaccination with another COVID-19 vaccine or end of follow-up. A washout period between the control and risk window will be applied to minimise capturing of prevalent events during the risk window. A post-vaccination control window may induce a bias towards the null when an AESI that was identified during the vaccination risk window, is captured again during the control window due to a re-admission instead of reoccurrence of the AESI. This situation is more likely to be encountered for acute potentially recurring events such as cardiac events (including acute coronary syndrome, myocarditis, pericarditis, arrhythmia, and stress cardiomyopathy). Moreover, due to lag-times in updating the relevant databanks, there is a higher risk that the post-vaccination control window may not capture all events and underestimate the rate of events. Use of a pre-vaccination control window will deal with these issues. In a sensitivity analyses, we will use a pre-vaccination control window and a post-vaccination control window will be used for the main analysis. In a sensitivity analysis, vaccination to other non-COVID-19 vaccines will also be an exclusion and censoring event.

We will also compare the Ad26.COV2.S vaccine with active comparators with 1) only 1-dose strategy and 2) a 2-dose strategy.

Figure 1 shows the SCRI study design diagrammatically using an example with a risk window and control window of 42 days.

Figure 1: Self-Controlled Risk Interval Design



T is time and T=0 is the index date.

8.1.2.2. Cohort Design (Primary Study Design for Non-acute Events and Sensitivity Analysis Design for Acute Events)

For the primary analysis of non-acute events, a retrospective cohort design will be used to estimate the rate of vaccination with Ad26.COV2.S vaccine, describe the characteristics of these vaccinated individuals and subsequently to estimate the incidence of new-onset autoimmune thyroiditis, multiple sclerosis, acute aseptic arthritis, and type 1 diabetes mellitus after receipt of the vaccine dose and to compare this incidence with that occurring in an unvaccinated matched comparator group, and in two groups of individuals exposed to other types of COVID-19 vaccines (i.e., one viral vector vaccine comparator group vaccinated [i.e., Vaxzevria® {AZD1222}] by Oxford/AstraZeneca) and one mRNA vaccine comparator group (i.e., Comirnaty® [tozinameran] by Pfizer/BioNTech and Spikevax® [elasomeran] by Moderna).

As a sensitivity analysis for acute events, a cohort analysis (following the same approach used in the main analysis for non-acute events) will be conducted.

- **Exposed cohort:** individuals will have received at least 1 dose of Ad26.COV2.S vaccine.
- **Concurrent unexposed cohort:** individuals that have not been vaccinated with Ad26.COV2.S vaccine or any other COVID-19 vaccine at any time prior to the index date (time zero) matched to the vaccinated individual for important characteristics as described in Section 8.7.2.1.
- **Cohort exposed to other COVID-19 vaccines:** individuals will have received at least 1 dose of a viral vector COVID-19 vaccine (i.e., Vaxzevria® [AZD1222] by Oxford/AstraZeneca) or an mRNA COVID-19 vaccine (Comirnaty® [tozinameran] by Pfizer/BioNTech and Spikevax® [elasomeran] by Moderna, respectively).

A sensitivity analysis will compare the Ad26.COV2.S vaccine with the active comparators with 1) a 1-dose strategy and 2) a 2-dose strategy.

In this retrospective cohort design, the start of follow-up will be defined as the time when study outcomes start to be counted. The index date in the exposed cohorts (i.e., recipients of the vaccines)

will be the day the first dose of the corresponding COVID-19 vaccination was received. This date in the unexposed group will be a day when an individual did not receive a COVID-19 vaccine dose and randomly chosen by calendar matching to the start date of the corresponding Ad26.COV2.S exposed individual (i.e., a random day during the same week that the matching individual in the exposed cohort receives the Ad26.COV2.S vaccine).

Individuals in the Ad26.COV2.S exposed cohort will be individually matched to one individual in the concurrent unexposed cohort and to one individual in each active comparator cohorts on key clinical variables (exact age, sex, previous COVID-19 infection and presence of one or more risk factors for severe COVID-19 [e.g., cancer, sickle cell disease, obesity, chronic kidney disease, chronic respiratory disease, human immunodeficiency virus infection], and month of vaccination (using calliper if needed)) at the index date.

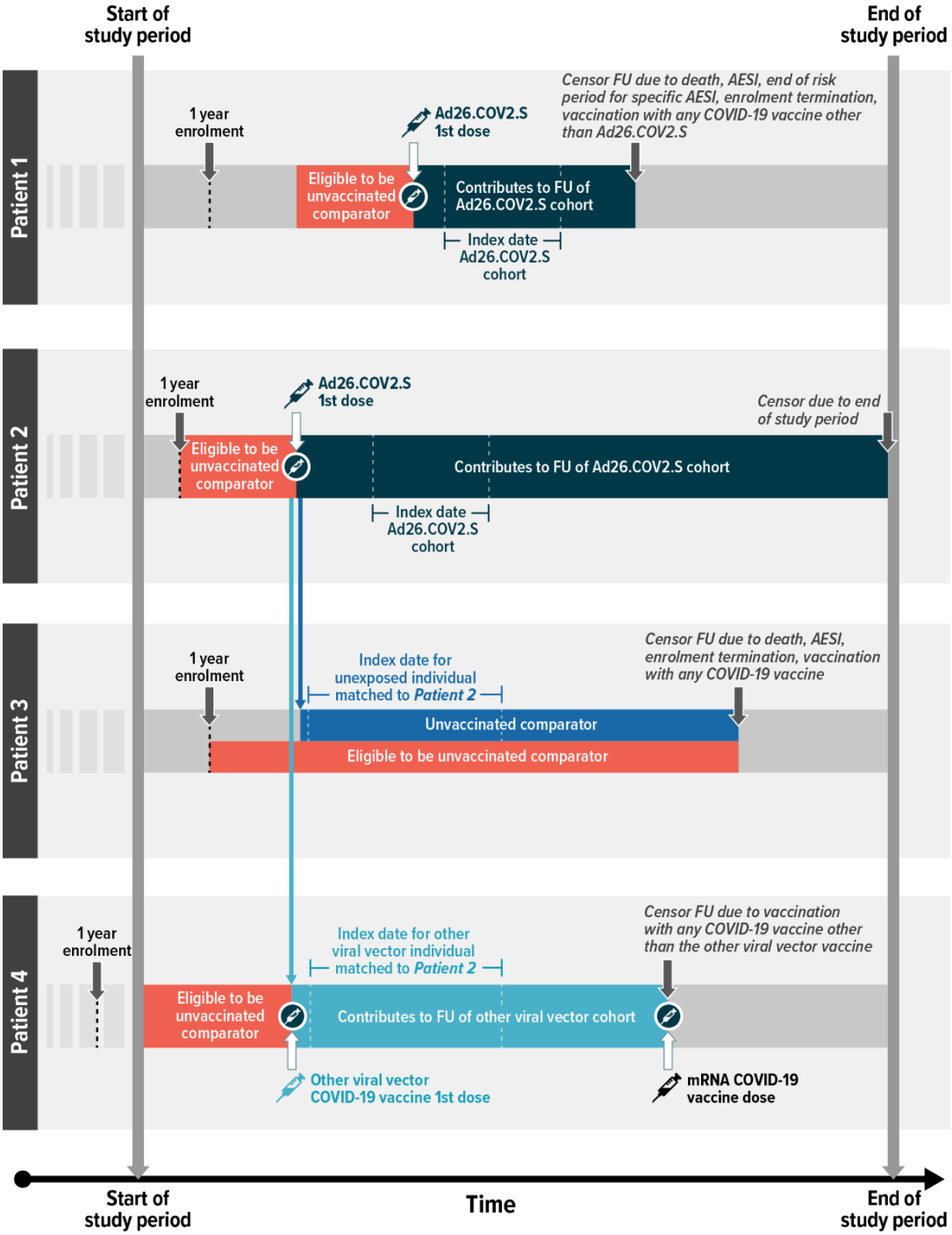
Three sets of matched cohorts of individuals who meet all inclusion and exclusion criteria will be considered as follows:

- Recipients of a first Ad26.COV2.S dose matched to recipients of the first dose of a viral vector COVID-19 vaccine (AstraZeneca vaccine)
- Recipients of a first Ad26.COV2.S dose matched to recipients of the first dose of an mRNA COVID-19 vaccine (Pfizer-BioNTech or Moderna)
- Recipients of a first Ad26.COV2.S dose matched to unvaccinated individuals at the index date (individuals in this cohort may later become eligible to enter the exposed cohorts if vaccinated with a COVID-19 vaccine).

For the matched exposed cohorts, the index date (time zero) will be the date of the corresponding COVID-19 vaccine. For the unvaccinated individuals, the index date will be assigned as the index date of the matched exposed individual (see Figure 2). Additional details regarding matching will be included in the SAP.

Follow-up for unexposed persons is censored when the individual receives a COVID-19 vaccine. In a sensitivity analysis, vaccination with other non-COVID-19 vaccines will also be an exclusion and censoring event.

Figure 2: Schematic Overview of the Cohort Study Design



AESI = adverse event of special interest; COVID-19 = coronavirus disease 2019; FU = follow-up;

SARS-CoV-2 = severe acute respiratory syndrome coronavirus-2.

Note: In the figure 1 year enrolment refers to the inclusion criterion to have at least 12 months of data available before the index date.

8.1.3. Rationale for Study Design: Safety Evaluation Study

The AESIs included in this study are considered as potential or identified safety risks following administration of Ad26.COV2.S. The selected AESIs represent a heterogeneous group including multiple organ systems, and acute and chronic conditions (Section 8.3.3).

The SCRI design has been selected as the primary design for acute events (i.e., AESIs expected to occur within a disease risk window of 60 days) because it only includes vaccinated individuals, which avoids misclassification of unexposed status due to incomplete capture of vaccinations in some of the electronic health care databases. Because it uses each individual as its own control, the SCRI design avoids potential confounding by factors that do not vary with time. To account for the possibility that increased risks of AESIs might extend beyond the control interval, the control and risk windows are separated by a buffer period to minimise potential of capturing prevalent events in the risk window (length to be defined in SAP). Although most AESIs are age-dependent, confounding by age is not anticipated because the follow-up period is short compared with the age effect on the incidence of AESIs. However, because the SCRI design generally has less statistical power and for some AESIs may be subject to confounding by other time-variant confounders such as seasonality (a proxy for respiratory infections), a cohort design will be used in sensitivity analyses of acute events, with matching on calendar time and other key variables.

Because the SCRI design is not appropriate for non-acute events, the cohort design will be used to assess these types of events (i.e., AESIs expected to occur within a disease risk window of more than 60 days). Some of the participating data sources may not have complete information on COVID-19 vaccination (e.g., PHARMO) and therefore risk misclassification of the non-exposed group (these may have been vaccinated). In a sensitivity analysis, these will be excluded to investigate the impact of misclassification of exposure in the non-exposed group on the pooled analysis of non-acute events.

Disease-specific risk windows have been defined for all AESIs; however, these risk windows are based on currently available evidence or best estimates (e.g., from recent post-marketing use, experience with other marketed vaccines or knowledge of COVID-19 or case reports) and may not be strictly applicable to the Ad26.COV2.S vaccine. For this reason, sensitivity analyses with extended risk windows will be considered in the SCRI study for outcomes that do not have a well-defined risk interval (to be defined in SAP). For these events, a descriptive time-to-onset analysis will be conducted during a period up to one year after the Ad26.COV2.S vaccine to inform the risk interval for the sensitivity analysis.

8.2. Setting and Study Population

8.2.1. Study Setting

For the implementation of the feasibility assessment, electronic health care databases in Southern and Western Europe that have shown interest and are members of the Vaccine Monitoring Collaboration for Europe (VAC4EU) will be used. The selected data sources and two-letter country codes are as follows:

- SIDIAP (Sistema d'Informació per el Desenvolupament de la Investigació en Atenció Primària) [Information System for Research in Primary Care] (ES)
- VID (La Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunitat Valenciana) [The Foundation for the Promotion of Health and Biomedical Research of Valencia Region] (ES).
- PHARMO DATABASE NETWORK (PHARMO Institute for Drug Outcomes Research) (NL)
- ARS Toscana (Agenzia Regionale di Sanità della Toscana) [a research institute of the Tuscany region of Italy] (IT)]. Since the approval of protocol version 2 by the EMA PRAC, a revision of the procedure for data access in the Tuscany Region has started. Hence, the study team at ARS Toscana has suspended its activities concerning the re-use of data until the revision of the procedures is complete. Until then ARS Toscana cannot contribute data to the study.

Further information on the data sources used in this study can be found in Section 8.4. When needed (based on specific safety concerns occurring in specific countries, or because of size) additional data sources may be added for dedicated evaluation studies, if they have adequate data on exposure, outcomes and covariates.

8.2.2. Source Population

The source population will be made up of all individuals registered in each of the participating health care data sources.

8.2.3. Study Duration and Follow-Up

Feasibility Assessment

For the feasibility assessment, the study period will start on the date when each participating country begins with the Ad26.COV2.S vaccine or the latest date when the individual is registered in the data source (plus 365 days) or born and will end a maximum of 12 months thereafter. Individuals will be followed until the earliest of the following dates: death, end of data availability, individual exit from the data source, or the completion of the period.

Safety Evaluation Study

- For the SCRI design, the study period will start at the date of the start of the pre-vaccination control window (Table 1). Follow-up ends at the earliest of the following: death, end of data availability, individual withdrawal of the study, end of the post-vaccination risk or control window, or receipt of another COVID-19 vaccine. In a sensitivity analysis, follow-up will be censored upon receipt of other vaccines (non-COVID-19 vaccine). We will also compare the Ad26.COV2.S vaccine with active comparators with 1) only 1-dose strategy and 2) a 2-dose strategy. Figure 1 shows the different periods considered for the SCRI design.
- For the cohort design, the study period will start at the date when each participating country will begin with the Ad26.COV2.S vaccine and will end at the end of data availability at each data source. For the cohort design, the follow-up will start at the beginning of the vaccination risk window, as detailed in Table 1 and ends at the occurrence of each AESI, death, end of data availability, individual exits the database, after one year of follow-up, receipt of a

COVID-19 vaccine (unexposed cohort only), receipt of a different COVID-19 vaccine from the one granting access to the study (exposed cohorts only). For the vaccine-exposed comparator groups, a second dose of the same corresponding vaccine will be allowed (more details will be provided in the SAP). In a sensitivity analysis, follow-up in the cohort study will be censored upon receipt of other vaccines (non-COVID-19 vaccine). Another sensitivity analysis will also compare the Ad26.COV2.S with active comparators with (1) a 1-dose strategy and (2) a 2-dose strategy. For the subgroup analysis of subjects without COVID-19 at the start of risk window, patients will be censored if they develop COVID-19 during follow-up.

8.2.4. Inclusion Criteria

8.2.4.1. Feasibility Assessment

For the feasibility assessment, the individual will be included if there is at least one day of follow-up and the individual has at least 12 months of data in the data source at the index date.

8.2.4.2. SCRI Design

For analyses of outcomes assessed with the SCRI design, the following criteria must be met. Note that the study population for each outcome-specific analysis will thus be different.

- Received a dose of Ad26.COV2.S vaccine during the study period.
- Have experienced the corresponding AESI during the control or risk window (as defined in Table 1)
- Have at least one day of follow-up in both the control and the risk window.
- Have at least 12 months of data/registration in the data sources prior to start of pre-vaccination control window.

8.2.4.3. Cohort Design

Individuals must meet all the following inclusion criteria to be eligible for inclusion in the cohort study:

- Have received their first-ever dose of the corresponding COVID-19 vaccine within the study period
- Had at least 12 months of data available before the index date or were born in the previous 12 months and have been enrolled since birth

8.2.5. Exclusion Criteria

For the feasibility assessment, there will be no exclusion criteria.

Individuals will be excluded from the safety evaluation study if:

- They have been vaccinated with any other COVID-19 vaccines before starting each risk window.

- Record of the specific AESI during the specific corresponding washout periods for the SCRI and cohort analyses. Individuals with such acute diagnoses before the AESI-specific washout periods will not be excluded. For clarification, when one AESI will be investigated, any history or prevalence of other AESIs will not be excluded. Table 1 shows the specific washout periods for each AESI.

8.3. Variables

8.3.1. Exposure Assessment

Exposure will be based on available prescriptions, dispensing, or administration of the Ad26.COV2.S and other COVID-19 vaccines. Vaccine receipt and date of vaccination will be obtained from all possible sources that capture COVID-19 vaccination, such as pharmacy dispensing records, general practice records, immunisation registers, vaccination records, medical records, or other data banks. During the feasibility assessment, the completeness of information will be assessed and described. Depending on the data source, vaccines may be identified via nationally used product codes (including batch numbers) where possible. The exposure of interest for the safety evaluation study is the receipt of Ad26.COV2.S vaccine.

- SIDIAP (ES): SIDIAP will have available information on the administration of COVID-19 vaccines to individuals linked to a unique and anonymous identifier. The information will be originated from the electronic medical records. For each patient, SIDIAP will have the date and centre of administration, health professional administering the vaccine, dose, brand, reasons for vaccination (e.g., risk group), and other information related to vaccination.
- VID (ES): For the inhabitants of the Valencia region, FISABIO will capture all COVID-19 vaccines through the immunisation register. The Vaccine Information System includes administration dates, brand and batch from each dose.
- PHARMO DATABASE NETWORK (NL): Data on vaccination will be obtained from PHARMO's general practitioners' GP database. Information on vaccines includes brand, batch, and date of administration/recording. Several COVID-19 vaccines have been administered outside traditional medical care settings and linkage with a GP database is not yet implemented, this may change over time.
- ARS Toscana (IT): ARS will identify vaccines from the regional immunisation register using the nationally used product code, including batch number. See section 8.2.1 regarding conditional participation of ARS Toscana.

The vaccination strategies for the different exposure groups will be defined as follows:

- Individuals who receive the first dose of the Ad26.COV2.S vaccine will be classified as exposed to Ad26.COV2.S vaccine. Individuals will be censored if and when they receive a non-Ad26.COV2.S COVID-19 vaccine during follow-up. Individuals who receive a second dose of Ad26.COV2.S vaccine will also be censored at the date of that dose as this is not the current dosing schedule. When the dosing schedule changes for Ad26.COV2.S vaccine, the study protocol may be amended.

- The vaccination strategy for the matched unexposed group will be defined as not receiving a COVID-19 vaccine of any brand during the follow-up period. Individuals will be censored when they receive a dose of any COVID-19 vaccine during the observation period.
- The vaccination strategy for the matched groups receiving other COVID-19 vaccines will be defined as 1) receiving a viral vector COVID-19 vaccine (i.e., Vaxzevria® [AZD1222] by Oxford/AstraZeneca) or 2) receiving a mRNA COVID-19 vaccine (i.e., Comirnaty® [tozinameran] by Pfizer/BioNTech or Spikevax® [elasomeran] by Moderna). Individuals will be censored when they receive a dose of a COVID-19 vaccine of a different type during the observation period (e.g., a patient with the first dose of AZD1222 receives a second dose with a mRNA COVID-19 vaccine).

For the SCRI design, person-time in the risk interval will be considered "exposed" while person-time in the control interval will be considered "unexposed." Risk intervals are specific to the outcome of interest.

8.3.2. Study Outcomes

AESIs, as listed below and in line with the definitions and code lists that for most of the AESIs have been created for the ACCESS project (https://zenodo.org/communities/vac4eu/), will be identified, with a date of diagnosis, using predefined validated algorithms (where available), based on diagnosis codes (with procedure and/or pharmacy dispensing codes and/or limited to specific medical care settings if applicable to the outcome). Only codes tagged as "narrow" (i.e., specific) will be used. Whenever no "narrow" codes are available in any applicable coding system, "possible" codes will be used instead to identify the corresponding AESI. It is possible that some AESIs will not have "narrow" codes in some coding systems. Therefore, some AESIs may not be assessable in all data sources. The impact of different provenance of data (hospital, general practitioner's diagnoses) and algorithms on the outcome frequency will be assessed and described in the feasibility assessment.

For the first feasibility assessment, the following rare and prioritised AESIs will be studied. The list includes encephalitis, including acute disseminated myelitis (ADEM) and meningoencephalitis; Guillain-Barré syndrome; transverse myelitis; multiple sclerosis, including optic neuritis; immune thrombocytopenia; anaphylaxis; cardiac inflammatory disorders, including myocarditis and pericarditis; stress cardiomyopathy; coronary artery disease, including acute myocardial infarction; deep vein thrombosis; pulmonary embolism; disseminated intravascular coagulation; non-haemorrhagic stroke; haemorrhagic stroke; thrombocytopenia and acute hepatic failure. As per EMA PRAC request, exacerbation of asthma will also be included in the feasibility assessment. For subsequent feasibility and final report, all AESIs will be analysed.

8.3.3. Evaluation of Safety Outcomes

The sponsor has created a list of AESIs based on current knowledge of the Ad26.COV2.S vaccine. Table 1 lists and explains the AESIs to be ascertained in the study with specific risk windows. Background incidence rates for most of the data sources and AESIs are available on the VAC4EU dashboard https://vac4eu.org/covid-19-tool/. Definitions and codes are available on Zenodo (https://zenodo.org/communities/vac4eu/).

Table 1: AESI-Specific Risk Windows and Washout Periods for Patient Inclusion in the Study

Body System	AESI	Disease-specific washout period before the index date	Disease-specific risk window following the index date
Nervous and central nervous system	Encephalitis, including ADEM and meningoencephalitis	365 days	1-42 days
	Guillain-Barré Syndrome	365 days	1-42 days
	Transverse myelitis	365 days	1-42 days
	Bell's palsy	365 days	1-42 days
	Multiple sclerosis, including optic neuritis	365 days	1-365 days
	Sensorineural Hearing loss	365 days	1-60 days
	Generalised convulsion	365 days (w/o epilepsy) 90 days (w/o non-epileptic seizures)	1-14 days
Immune system	Autoimmune thyroiditis	365 days	1-365 days
	Immune thrombocytopenia	365 days	1-42 days
	Type 1 diabetes mellitus	Any time prior*	1-365 days
	Acute aseptic arthritis	365 days	1-365 days
	Anaphylaxis	90 days	0-2 days
Cardiac system	Cardiac inflammatory disorders, including myocarditis and pericarditis	365 days	1-28 days
	Microangiopathy	365 days	1-28 days
	Heart failure	365 days	1-28 days
	Stress cardiomyopathy	365 days	1-28 days
	Coronary artery disease, including acute myocardial infarction	365 days	1-28 days
	Arrhythmia	365 days	1-14 days
Blood and lymphatic system disorders	Deep vein thrombosis	365 days	1-28 days
	Pulmonary embolism	365 days	1-28 days
	Disseminated intravascular coagulation	365 days	1-28 days
	Non-haemorrhagic stroke	365 days	1-28 days
	Haemorrhagic Stroke	365 days	1-28 days
	thrombocytopenia	365 days	1-28 days
	peripheral thrombosis	365 days	1-28 days
	Sinus thrombosis	365 days	1-28 days
	Thrombotic Thrombocytopenia syndrome	365 days	1-28 days
	Composite: venous thrombosis	365 days	1-28 days
	Composite: arterial thrombosis	365 days	1-28 days
	Composite: strokes	365 days	1-28 days
Renal system	Acute kidney failure	365 days	1-14 days
Hepatic system	Acute hepatic failure	365 days	1-14 days
Respiratory system	Exacerbation of asthma	365 days	1-14 days

Abbreviations: ADEM = Acute disseminated encephalomyelitis, AESI = adverse event of special interest,
All non-acute events are considered to have a washout period of 365 days. * For diabetes mellitus which is a chronic condition we will use any time prior for washout.

Exacerbation of asthma ascertainment will require the definition of an asthma cohort and identifying exacerbations of asthma within that cohort.

The asthma cohort will be defined by having:

- a code for asthma and;
- at least 2 prescriptions for an inhaled corticosteroid within 5 years before the index date.

Among this asthma cohort, an exacerbation of asthma was defined by having:

- At least 1 prescription of a systemic glucocorticosteroid during the defined risk window or;

- At least 1 hospitalisation for asthma during the defined risk window.

The date of this event will be the date of the first prescription of a systemic glucocorticosteroid or the admission date for hospitalisation for asthma, whichever comes first.

8.3.3.1. Case Ascertainment and Validation Process

The incidence of AESIs within disease-specific risk windows following the administration of the Ad26.COV2.S vaccine will be ascertained using diagnosis codes that were developed and reported to EMA for the ACCESS study (ACCESS 2021). Additional and more refined electronic algorithms (e.g., by requiring multiple codes, or having confirmatory medications/procedures, laboratory values or laboratory assessment) will be created and described in the SAP and assessed in the feasibility assessment. It will also be described how to assess the case index date, whether this will be based on a diagnosis or the first retrievable symptom.

For the safety evaluation study, case validation to confirm the disease diagnosis and date of onset may be conducted on AESIs that are judged likely to be misclassified based on clinical expert opinion, prior validation studies (if available), and data access providers. Manual review of medical records or chronological listings of diagnosis, procedure, and pharmacy dispensing codes in patient profiles by clinicians may be conducted, depending on the AESI. Where available, Brighton Collaboration definitions or any other clinical definitions from published literature or learned societies will be used as case definitions if medical record review is implemented for the AESI. For outcomes selected for validation, a sampling strategy will be implemented based on the rarity of events and identified cases in the study population, to identify cases that will undergo case validation. Following case validation, if the positive predictive value (PPV) of a given AESI is above a pre-specified threshold (e.g., ≥80%), all identified cases in the electronic data will be included in the final analysis. If the PPV is lower than the pre-specified threshold (e.g., <80%), then the validation results will be used to inform or adjust relative risk estimates (e.g., in quantitative bias analysis) in electronic data. The PPV threshold for determining whether all cases identified in electronic data are to be included in the final analysis will be specific to each AESI and will depend on the background rate of each AESI. The sampling strategy, the selection of which AESIs are to be validated and the rationale for selection, details on the methods for validation, and the plan for integrating the validation results into the final analysis will be described in a data validation plan.

SIDIAP (ES): In SIDIAP, the validation process is part of data quality control. On the other hand, specific validation studies on selected endpoints will be based on the review of the electronic medical record information by the own primary care clinicians exploring the overall coded and text information and existing reports.

VID (ES): In VID, validation will be based on the review of the electronic medical record information by members of the FISABIO research group who will be blinded to COVID-19 vaccination status.

PHARMO DATABASE NETWORK (NL): For validation, information on selected endpoints from patient medical records will be abstracted by local medical professionals or PHARMO

employees, provided that medical record review is approved by ethics committees and other local and/or national governing bodies.

ARS (IT): Review of medical records is not possible for ARS. See section 8.2.1 regarding conditional participation of ARS Toscana

8.3.4. Covariate Definition

Feasibility Assessment

In the feasibility assessment, covariates will be assessed at the time of COVID-19 vaccine administration or time-matched non-vaccinated individuals.

Safety Evaluation Study

Covariate assessment for the description of the patient population at baseline (index date) will use all available time of enrolment in the corresponding data source (minimum required enrolment period is 1 year).

Covariates will be assessed at index date (for the cohort design) to define patient populations of special interest or priority vaccination groups, to define subgroups of interest for sensitivity analyses, or to control for confounding.

We will consider the following time-varying covariates and corresponding period of evaluation for the SCRI measured at start of the risk window and the control window. To capture use of medications of interest during the month prior to start of the study periods (i.e., risk control and control risk), we will consider the drug supply covering the month prior to the start of the two periods of interest.

- Analgesics (month prior)
- Antibiotics (month prior)
- Antiviral medications (month prior)
- Systemic Corticosteroids (month prior)
- Immunosuppressant therapy (month prior)
- Non-steroidal anti-inflammatory drugs (month prior)
- Heparin (month prior)
- Novel oral anticoagulants (month prior)
- Warfarin (month prior)
- Aspirin (month prior)
- Influenza vaccine (month prior)
- Any pneumococcal vaccine (month prior)
- Any herpes zoster vaccine (month prior)

- Human papillomavirus vaccine (month prior)
- Meningitis vaccine (month prior)
- Pregnancy (at start of each window of interest)
- COVID-19 infection (month prior)

For the cohort design, covariate status for all time-varying factors will be measured at index date. All covariates will be assessed in specific periods. For the secondary analysis in dedicated subgroups or patients with a past history, any time prior as the look-back period for the assessment of the covariates will be used, reference point is index date. For vaccinations other than COVID-19, follow-up time will be censored in a sensitivity analysis.

Population characteristics will be identified based on diagnoses, medicines, laboratory data, survey observation or medical observations, and observation period information.

Demographic characteristics

- Age
- Sex
- Pregnancy

Pregnancy status and pregnancy trimester at index date.

- Frailty

The eFI frailty score as described by Clegg et al (Clegg 2016) with a count of the number of different conditions will be created. Details will be available in the SAP.

Comorbidities (any time prior)

- Cancer (other than non-melanoma skin cancer)
- Chronic kidney disease (exclusion for acute kidney injury)
- Acute coronary syndrome (exclusion for cardiac/cardiovascular events)
- Chronic respiratory disease (chronic obstructive pulmonary disease, asthma)
- Obesity
- Down's syndrome
- Type 1 (excluded for type 1 when AESI) or type 2 diabetes
- Prior thrombosis (exclusion for thrombosis as outcome – see Table 1)
- Morbidity index: number of different ATC codes (Level 5) dispensed in the year prior to cohort entry
- History of anaphylaxis
- History of any type of allergic reaction
- Immunocompromised conditions (will be used to define subgroups for secondary analyses)

  - Immunodeficiencies
  - Systemic Immunosuppressant medication use
  - Human immunodeficiency virus and other immunosuppressing conditions
- Parkinson's disease
- Stroke/brain injury
- Dementia
- Difficulty walking
- Chronic use of home oxygen (>3 months of continuous use)
- Palliative care
- Heart failure (exclusion for heart failure as an outcome).

Covid-19 History

- Prior recorded COVID-19 infection as yes/no for the feasibility studies and by severity for the final study (non-hospitalised, hospitalised)

Comedication (dispensed/prescribed/used in period or episode covering this period)

- Analgesics (month prior)
- Antibiotics (month prior)
- Antiviral medications (month prior)
- Systemic Corticosteroids (month prior)
- Non-steroidal anti-inflammatory drugs (month prior)
- Psychotropics (year prior)
- HMG-CoA Reductase Inhibitors (statins) (year prior)
- Novel oral anticoagulants (month prior/time-varying)
- Warfarin (month prior)
- Aspirin (month prior)
- Oral contraceptives (year prior)
- Hormone replacement therapy (year prior)
- Heparin (month prior)
- Immunosuppressants
- Use of antihypertensive drugs.

Vaccinations

- Administration of any other COVID-19 vaccine from the reference date until the end of the follow-up period (censoring criterion)

- Influenza vaccine (month prior/time-varying)
- Any pneumococcal vaccine (month prior/time-varying)
- Any zoster vaccine (month prior/time-varying)
- Human papillomavirus vaccine (month prior/time-varying)
- Meningitis vaccine (month prior/time-varying).

Health Care Utilisation in the Year Before and in the 4 Weeks Prior to Index Date

- Number of individuals with at least one hospitalisation (hospitalisation data sources)
- Number of individuals with at least one GP-visit (Primary care data sources)
- Number of different dispensed/prescribed drugs ATC Level 5 dispensed (all data sources)

The AESI may have different sets of risk factors, and outcome-specific analyses may contain different covariate sets. Potential covariates may include the following information, as available in each data source. For the SCRI design, covariate status will be measured at start of risk window and control window for time-varying factors.

For the secondary analyses in subgroups, the following categories will be used:

- Immunocompromised individuals,
- Pregnant women,
- Prior COVID-19 infection,
- Individuals who have a prior history of thrombotic events or thrombocytopenia,
- Individuals with a prior history (ever) of that event (i.e., more than a year before).

8.4. Data Sources

The study will use data from secondary electronic health record databases that are population-based. All data sources will have the ability to provide data on COVID-19 vaccines, outcomes (diagnoses, procedures, laboratory results, and treatments), and important covariates. It is not currently known the extent to which COVID-19 vaccines, product types, and batch numbers will be captured in data sources. To be included in the study, data sources should preferably be updated at a minimum once every 3 months. At the proposal stage, members of VAC4EU (https://vac4eu.org/) were offered the option to participate in the study. Four data sources from 3 countries (Italy, Spain, and The Netherlands) countries will be included in the study. A more detailed description will be included in the VAC4EU FAIR Catalogue that Molgenis based (https://www.molgenis.org/).

When establishing the agreements with the data sources to conduct the study, it will be emphasized that the current list of AESIs may be expanded during the course of the study to accommodate newly identified AESIs.

8.4.1. SIDIAP (ES) (7 million Active Individuals)

The Information System for Research in Primary Care (Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària' [SIDIAP]; www.sidiap.org) was created in 2010 by the Catalan Health Institute and the IDIAP JGol Institute. It includes information collected since 01 January 2006 during routine visits at 278 primary care centres pertaining to the Catalan Health Institute in Catalonia (North-East Spain) with 3,414 participating GPs. SIDIAP has pseudo-anonymised records for 7 million people (80% of the Catalan population) and is highly representative of the Catalan population. The SIDIAP data comprise the clinical and referral events registered by primary care health professionals (e.g., GPs, paediatricians, and nurses) and administrative staff in electronic medical records, comprehensive demographic information, community pharmacy invoicing data, specialist referrals, and primary care laboratory test results. The SIDIAP data can also be linked to other data sources, such as the hospital discharge database, on a project-by-project basis. Health professionals gather this information using ICD-10-CM codes, ATC codes, and structured forms designed for the collection of variables relevant to primary care clinical management, such as country of origin, sex, age, height, weight, body mass index, tobacco and alcohol use, blood pressure measurements, and blood and urine test results. Regarding vaccinations, SIDIAP includes all routine childhood and adult immunisations, including the antigen and the number of administered doses. Encoding personal and clinic identifiers ensures the confidentiality of the information in the SIDIAP database. Currently, with the COVID-19 pandemic, there is the possibility to have shorter-term updates in order to monitor the evolution of the pandemic. Recent reports have shown the SIDIAP data to be useful for epidemiological research. SIDIAP is listed under the ENCePP resources database.

8.4.2. VID ES (5 million Active Population)

The region of Valencia, with 5 million inhabitants, is part of the Spanish National Health System, a universal public healthcare system. Information will be obtained from the population-based electronic data banks of the Valencia Health Service and the regional Government of Valencia: (1) The Population Information System (SIP) provides an identification number for each individual under Valencian Health Service coverage and registers some demographic characteristics, and dates and causes of Valencia Health Service discharge, including death. (2) The minimum basic dataset at hospital discharge is a synopsis of clinical and administrative information on all hospital discharges, including diagnoses and procedures (all electronic health systems in the Valencian Health Service use the ICD- 9-CM and ICD-10-CM). (3) The Emergency Department module including emergency department dates of visit and discharge and reason for discharge. (4) The electronic medical record (EMR) for ambulatory care, is available in all primary healthcare centres and other ambulatory settings. It has all the information on patients regarding diagnoses, their personal and family medical history, laboratory results, lifestyle, etc (5) The pharmaceutical module (prescription information system), part of EMR, includes information about both physician prescriptions and dispensations from pharmacy claims. (6) The Corporate Resource Catalogue provides information about the geographical and functional organisation of Valencian Health Service, its health centres, health services provided and professionals in healthcare. Specific public health registries are available and linkable at an individual level (such as the perinatal registry and the congenital anomalies registry, from which pregnancy outcomes

can be obtained) All the information in these systems can be linked at an individual level through the SIP number.

8.4.3. PHARMO (NL) (2.5 million Active Individuals)

The PHARMO Database Network, which is maintained by the PHARMO Institute for Drug Outcomes Research, is a population-based network of electronic health record databases that combines anonymous data from different primary and secondary health care settings in the Netherlands. These different data banks—including data from general practices, in- and outpatient pharmacies, clinical laboratories, hospitals, the cancer register, the pathology register, and the perinatal register—are linked on a patient level through validated algorithms. To ensure data privacy in the PHARMO Database Network, the collection, processing, linkage, and anonymisation of the data are performed by Stichting Informatievoorziening voor Zorg en Onderzoek, which is an independent, ISO/IEC 27001 certified foundation that acts as a trusted third party between the data sources and the PHARMO Institute. The longitudinal nature of the PHARMO Database Network enables the follow-up of more than 9 million individuals of a well-defined population in the Netherlands for an average of 12 years. Currently, the PHARMO Database Network covers over 6 million active individuals out of 17 million inhabitants of the Netherlands. Data collection period, catchment area, and overlap between data banks differ. Therefore, the final cohort size for any study will depend on the data banks that are required. All electronic patient records in the PHARMO Database Network include information on age, sex, socioeconomic status, and mortality. Other available information depends on the data banks. A detailed description of the different data banks is given in subsequent sections. The PHARMO Institute is always seeking new opportunities to link with additional databanks, currently, it is exploring linkage with the COVID-19 immunisation register that is collected by RIVM.

The General Practitioner databank comprises data from electronic patient records registered by GPs. The records include information on diagnoses and symptoms, laboratory test results, referrals to specialists, and health care product/drug prescriptions. The prescription records include information on the type of product, prescription date, strength, dosage regimen, quantity, and route of administration. Drug prescriptions are coded according to the WHO ATC coding system. Diagnoses and symptoms are coded according to the International Classification of Primary Care (ICPC) [www.nhg.org], which can be mapped to the International Classification of Diseases (ICD) codes but can also be entered as free text. General practitioner data cover a catchment area representing 3.2 million residents (~20% of the Dutch population).

The Outpatient Pharmacy Databank comprises GP- or specialist-prescribed health care products dispensed by the outpatient pharmacy. The dispensing records include information on the type of product, date, strength, dosage regimen, quantity, route of administration, prescriber speciality, and costs. Drug dispensing is coded according to the WHO ATC classification system. Outpatient pharmacy data cover a catchment area representing 4.2 million residents (~25% of the Dutch population). The PHARMO Database Network is listed under the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) resources database. PHARMO GP databank captures vaccinations supplied by the GP (influenza, zoster, COVID-19).

As of 08 May 2021, ~65,000 doses of Ad26.COV2.S vaccine have been administered in the Netherlands. Assuming PHARMO covers ~20% of the Dutch population, this should be 16,250 doses in PHARMO's catchment area. However, different healthcare providers administer the COVID-19 vaccines (i.e., GPs, the public health service and healthcare institutions) in the Netherlands. The vaccination data are recorded in a central register (if people have given permission beforehand), the COVID-19 vaccination Information and Monitoring System. PHARMO is currently exploring the possibilities to link with this register. Until then, the GP Database will be the basis for the vaccination data. The GP Database contains vaccinations administered by GPs and by the public health service, as GPs receive an automated notification when a patient has a positive corona test or has been vaccinated via the public health service (as long as people have given their consent).

It is difficult to estimate how many persons will be vaccinated with Ad26.COV2.S vaccine during the study period in the Netherlands. The Dutch government wants everyone from the age of 18 to have had at least one dose of COVID-19 vaccine by the beginning of July 2021. This vaccination schedule depends on many factors (e.g., approval and effectiveness, delivery and distribution of vaccines to injection sites, such as hospitals and GPs, new developments and advice from, for instance, the Health Council of the Netherlands (i.e., de Gezondheidsraad). Also, the choice of the type of vaccination depends on different factors, such as type of work (care workers), home living (yes or no), year of birth and existing comorbidities. PHARMO acknowledges that the data source they have access to includes data on vaccine delivery and registration and undertakes to cooperate in addressing the Study objectives by contributing to providing reports based on such data.

8.4.4. ARS Toscana Database (IT) (3.6 million Active Individuals)

Disclaimer: ARS Toscana's participation in the study is subject to the revision of the procedures for data access in the Tuscany Region. See section 8.2.1 regarding conditional participation of ARS Toscana.

The Italian National Healthcare System is organised at the regional level: the national government sets standards of assistance and tax-based funding for each region, which regional governments are responsible for providing to all their inhabitants. Tuscany is an Italian region, with approximately 3.6 million inhabitants. The Agenzia Regionale di Sanità della Toscana (ARS Toscana) is a research institute of the Tuscany region. The ARS Toscana database comprises all information collected by the Tuscany region to account for the health care delivered to its inhabitants. Moreover, ARS Toscana collects data from regional initiatives. All data banks in the ARS Toscana data source can be linked at the individual level through a pseudo-anonymous identifier. Two data banks collect dispensing of reimbursed medicines from, respectively, community pharmacies and hospital pharmacies. In the latter data bank, dispensing for outpatient and ambulatory use are complete, and dispensing for inpatient use are partial. Other data banks include hospital discharges, emergency care admissions, records of exemptions from co-payment, diagnostic tests and procedures, causes of death, the mental health services register, the birth register, the spontaneous abortion register, and the induced terminations register. A pathology register is available, mostly recorded in free text, but with morphology and topographic SNOMED codes. A COVID-19 registry including all positive cases with clinical follow-up is also available.

Mother-child linkage is possible through the birth register. Vaccination data are available for children since 2016 and for adults since 2019. All the data banks can be linked at the individual level through a pseudonymous identifier. Data banks are updated approximately every 2 months. Some of them are updated at the date of transmission (e.g., vaccines, COVID-19 registry, access to emergency room), while others (e.g., medicines dispensing and hospital discharge records) have a delay of approximately 4 months.

8.5. Study Size

The study will be conducted in a source population of approximately 18.1 million individuals, although children will not be vaccinated. It is assumed that up to 10% will be vaccinated with the Ad26.COV2.S vaccine.

8.5.1. Cohort

Subject to the feasibility assessment, a minimum target sample size of 100,000 individuals in the exposed cohort will be included in the study. Table 2 shows the number of exposed subjects needed for various upper limits of the 95% CI of 2, 3, 5, and 10, assuming a ratio of 1:1 to unexposed or exposed and a power of 80% for different background incidence rates. For example, 80,000 individuals in the exposed cohort and 80,000 individuals in the unexposed cohort will allow the detection of an upper limit of the 95% CI of the relative risk equal to or greater than 3 with 80% statistical power for diseases with a background incidence rate of ≥100 per 100,000 person-years. We assume similar numbers of exposed comparators for those comparisons.

Table 2: Number of Exposed Subjects* needed to have an Upper Limit of the 95% CI of 2, 3, 5, and 10 Assuming a Ratio of 1:1 to Unexposed and a Power of 80% for Different Background Incidence Rates

	Upper limit of 95% CI for IRR			
Background IR per 100,000 PY	**2**	**3**	**5**	**10**
1	19,875,972	7,912,075	3,686,641	1,801,142
10	1,987,598	791,208	368,665	180,115
50	397,520	158,242	73,733	36,023
100	198,760	79,121	36,867	18,012

Rothman 2018.
CI = confidence interval; IR = incidence rate; IRR = incidence rate ratio; PY = per year.
*Assuming each individual contributes a 60-day risk window.

8.5.2. Self-Controlled Risk Interval design

Table 3 shows the statistical power that can be obtained for a range of relative risks and a range of sample sizes. For example, a sample size of 100 cases either in the risk or control window will allow the detection of a relative risk equal to or greater than 2 with 93% statistical power.

Table 3: Detectable Relative Risk and Statistical Power for SCRI Design

Relative Risk	Sample Size*	Power
1.5	20	14.2%
2	20	32.0%
2.5	20	49.5%
3	20	63.8%
1.5	50	29.2%
2	50	66.7%

Table 3: Detectable Relative Risk and Statistical Power for SCRI Design

Relative Risk	Sample Size*	Power
2.5	50	88.1%
3	50	96.3%
1.5	100	51.9%
2	100	92.6%
2.5	100	99.4%
3	100	100.0%
1.5	150	69.2%
2	150	98.7%
2.5	150	100.0%
3	150	100.0%
1.5	200	81.2%
2	200	99.8%
2.5	200	100.0%
3	200	100.0%

*Sample size = number of events in risk and control window.

Table 4 shows the number of vaccinated individuals assuming the same time in the control and risk window to obtain an 80% statistical power for a range of relative risks. For example, a sample size of 69 vaccinated individuals with an AE of interest will allow the detection of a relative risk equal to or greater than 2 with 80% statistical power.

Table 4: Detectable Relative Risk and Sample Size for 80% Statistical Power for SCRI Design

Relative Risk	Individuals with AE of Interest	Power
1.5	195	80.0%
2	69	80.5%
2.5	41	80.9%
3	29	80.4%

8.6. Data Management

This study is conducted in a distributed manner using a common protocol, common data model (CDM), and common analytics programs (Figure 3) The data pipeline has been developing from the EU-ADR project and was further improved in the IMI-ConcePTION project (https://www.imi-conception.eu/) and used to generate background rates in the ACCESS project. This process maximises the involvement of the data providers in the study by utilising their knowledge of the characteristics and the process underlying the data collection which makes analysis more efficient.

8.6.1. Data Extraction

Each database access provider (DAP) creates extraction, transform, and load (ETL) specifications using the standard ConcePTION ETL design template (accessible via this link: https://docs.google.com/document/d/1SWi31tnNJL7u5jJLbBHmoZa7AvfcVaqX7jiXgL9uAWg/edit) and upload it to the VAC4EU FAIR Catalogue.

The most recent version of the ConcePTION CDM (2.2) will be used for this analysis.

Following completion of this template and review with study statisticians and principal investigators, each DAP extracts the relevant study data locally using their software (e.g., Stata, Statistical Analysis Software [SAS], R, Oracle). These data are loaded into the CDM structure in CSV format. These data remain local (Figure 3).

8.6.2. Description of Data Transformation and Analysis Pipeline

This study uses data that is already collected for analysis and available in electronic health care data sources in 4 EU countries and follows the following principles.

First, to harmonise the structure of the data sets held by each partner, a shared syntactic foundation is utilised, the CDM will be used that was developed in the IMI-ConcePTION project. In this CDM, data are represented in a common structure, but the content of the data remains in their original format. The ETL design is made available on paper and later on in the VAC4EU FAIR Catalogue. The validity of the ETL will be assessed using Level 1 (completeness) and Level 2 (logical consistency) R-scripts that have been developed as part of the IMI-ConcePTION project.

Second, to reconcile differences across terminologies a shared semantic foundation is built for the definition of events under study by collecting relevant concepts in a structured fashion using a standardised event definition template. The Codemapper tool (https://vac4eu.org/codemapper/) was used to create diagnosis code lists based upon completed event definition templates for each event and comorbid risk condition in the ACCESS-BGR protocol (Zenodo.org).

Based on the relevant diagnostic medical codes, as well as other relevant concepts (e.g., medications), algorithms were constructed to operationalise the identification and measurement of each event. These algorithms may differ per data source, as the components that go into the study variable may differ. Wherever possible the event definition sheet specifies prior validation of algorithms and codes for benchmarking. Scripts for semantic harmonisation are created centrally and provided in R and distributed to data access providers for local deployment. This will result in a set of study variables which are both semantically and syntactically harmonised. The quality of the semantic harmonisation will be assessed using Level 3 checks and compared against published rates and between databases (benchmarking). Limited to outcomes, components will be analysed in each data source, to assess the unique contribution of each data bank, therefore providing evidence on the sensitivity and accuracy of the algorithms. Results from validation activities will be used to inform the calculation of the validity of the composite algorithms.

Third, following conversion to harmonised study variable sets, additional R- or SAS-scripts for the calculation of analytical datasets are distributed to data access providers for local deployment. The output datasets produced by these scripts are then uploaded to the Digital Research Environment (DRE) for pooled analysis of incidence and visualisation (Figure 3). The DRE is made available through UMCU/VAC4EU (https://www.andrea-consortium.org/). The DRE is a cloud-based, globally available research environment where data are stored and organised securely and where researchers can collaborate (https://www.andrea-consortium.org/azure-dre/).

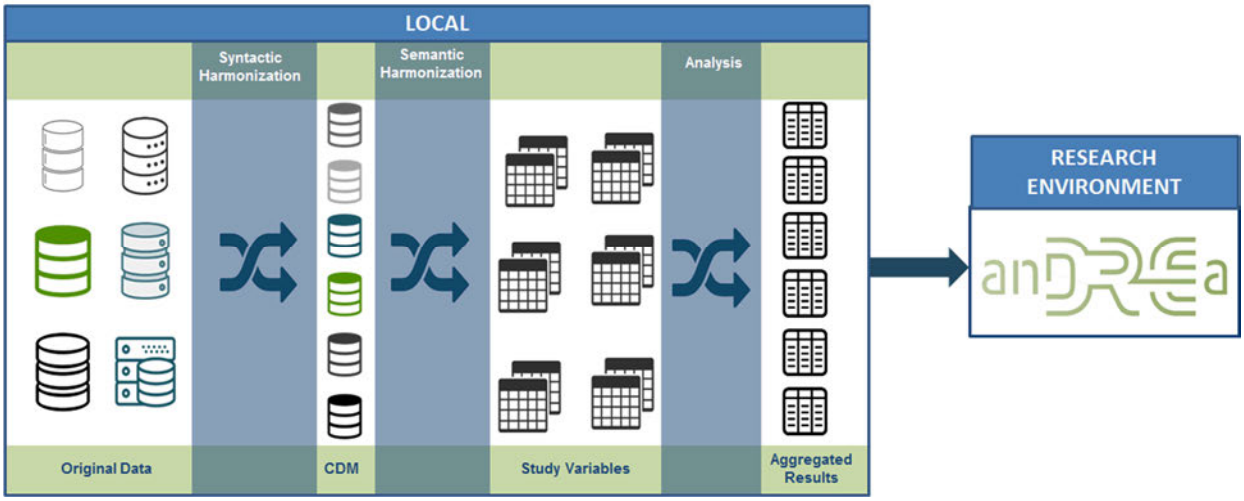
All final statistical computations are performed on the DRE using R and SAS. Data access providers have access to the project workspace for verification of the results.

8.6.3. Data Access

Within the DRE, each project-specific area consists of a separate secure folder called a "workspace." Each workspace is completely secure, and researchers are in full control of their data. Each workspace has its own list of users, which can be managed by its administrators. The DRE architecture allows researchers to use a solution within the boundaries of data management rules and regulations. Although General Data Protection Regulation and Good (Clinical) Research Practice still apply to researchers, the DRE offers tools to more easily control and monitor which activities take place within projects. All researchers who need access to the DRE are granted access to study-specific secure workspaces.

Access to this workspace is only possible with double authentication using an identification code and password together with the user's mobile phone for authentication. Upload of files is possible for all researchers with access to the workspace within the DRE. Downloading files is only possible after requesting and receiving permission from a workspace member with an "owner" role.

Figure 3: Data Transformation and Flow



8.6.4. Data Processing

Due to the nature of the study, a repeated data processing procedure is envisioned, based on the pipeline described in the previous section. This allows for optimising the data processing timelines and archiving procedures (see Section 8.6.6).

In the study start, a baseline data extraction will be made by each of the DAPs. This creates a baseline instance of the data source. This is ETL'ed into the ConcePTION CDM and forms the baseline instance of the CDM. The data pipeline will be run for the first time on the baseline instance of the CDM of each DAP, and produce a baseline set of analytic datasets that will be centrally analysed for the baseline assessment.

Periodically (every 4 months), DAPs are requested to perform the extraction of the new data, which will be called the x^{th} supplementary instance of their data source. The supplementary instance of the data source will be ETL'ed to the ConcePTION CDM, to form a supplementary instance of the CDM, which will be analysed, this instance comprises all the prior data.

8.6.5. Case Report Forms/Data Collection Tools

This study will use secondary data collected in electronic health record databases. For the purpose of validating selected study endpoints, special forms may need to be developed and securely saved in environments assuring data protection and patient confidentiality according to the requirements of each country and DAP.

8.6.6. Record Retention

DAPs are responsible locally to archive each data source instance that is used for the study. The meta-data table in the CDM allows for storing details on the data source instance. The DAP has the obligation to archive the data source instances, the ETL scripts, the R-scripts that were used and the results that were uploaded to the DRE, locally.

Aggregated results from DAPs, ETL design documents, and a repository of study scripts will be stored in the DRE for inspection by the study sponsor for at least five years. The final study aggregated results sets and statistical programs to pool and visualise will be archived and stored on the DRE for five years.

Documents that individually and collectively permit evaluation of the conduct of a study and the quality of the data produced will be retained for a period of 5 years in accordance with Good Pharmacoepidemiology Practices (GPP) guidelines. Study records or documents may also include the analysis files, syntaxes (usually stored at the site of the database), ETL specifications, and output of data quality checks.

After 5 years all materials from the DRE will be retained for at least 15 years on a UMCU secure drive. The final study protocol and possible amendments, the final statistical report, statistical programs and output files will be archived on the UMCU secure drive according to Julius Clinical standard operating procedures.

8.7. Data Analysis

Detailed methodology for summary and statistical analyses of data collected in this study will be documented in an SAP for the feasibility assessment and the respective SCRI and cohort safety analyses. All analyses will be conducted using R version R-4.0.3 or higher (Foundation for Statistical Computing, Vienna, Austria; https://www.R-project.org) or SAS version 9.3 software or higher (Cary, North Carolina, USA; SAS Institute, Inc.). The SAP will contain more detail about the analysis and data pooling and 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.

A general description of the planned statistical methods to be used to analyse the data collected in this study is presented in the following subsections. Additional details will be provided in the SAP. A summary of planned analyses is provided in Table 5.

Table 5: Summary of Planned Analyses

Analysis	Acute AESIs Study design	Non-acute Study design for Non-acute AESIs	Risk window Definition of Risk window	Notes on the main purpose of analysis/difference from the primary analysis
Feasibility	Cohort	Cohort	none	Baseline characteristics, vaccine uptake, assessing the quality of data and of algorithms for outcomes and covariates, and the number of outcomes and time-to-event.
Primary (Final)	SCRI: post-vaccination control window	Cohort: exposed versus concurrent unexposed comparator and versus other COVID-19 vaccine exposed	Disease-specific risk window	SCRI: Using a pre and post-vaccination control window to assess the effect of the order of windows; Calculating and comparing the incidence rates of each AESI; Analysis of negative control data
Secondary (Final)	Subpopulation SCRI	Subpopulation cohort	Disease-specific risk window	Sub-population analysis
Sensitivity (Final)	Cohort: exposed versus concurrent unvaccinated comparator and/or versus other COVID-19–vaccine exposed	Cohort: exposed versus concurrent unexposed comparator and versus other COVID-19 vaccine exposed	Extended risk windows for acute events	Acute events: Compare different analytical approaches and incorporate the use of extended risk windows; Non-acute: not applicable
Sensitivity (Final)	SCRI: Comparing Ad26.COV2.S with active comparators with (1) only 1 dose and (2) a 2-dose strategy	Cohort: Comparing Ad26.COV2.S with active comparators with (1) only 1 dose and (2) a 2-dose strategy	Disease-specific risk window	Address the different dosing schedules between Ad26.COV2.S and the comparator vaccines used in the analyses
Sensitivity (Final)	SCRI, pre-vaccination control window	Cohort analyses with non-exposed restricted to data sources with good ascertainment of COVID-19 vaccine exposure	Disease-specific risk window	SCRI: Using a pre- and post-vaccination control window to assess the effect of the order of windows; Cohort: Restricted analysis to avoid misclassification of unvaccinated status; Risk window
Sensitivity (Final)	SCRI, excluding and censoring upon taking other (non-COVID-19) vaccinations	Cohort, excluding and censoring upon taking other (non-COVID-19) vaccinations	Disease-specific risk window	To avoid the confounding effects of non-COVID-19 vaccines
Stratified analysis (Final)	SCRI: pre-vaccination and post-vaccination control window	Cohort: exposed versus concurrent unexposed comparator and versus other COVID-19 vaccine exposed	Disease-specific risk window	Subgroup analysis according to the baseline factors such age and sex

AESI= adverse event of special interest; SCRI= self-controlled risk interval
Note: Acute AESIs are events expected to occur within 60 days of vaccination; non-acute events are events expected to be recorded >60 days after vaccination.

8.7.1. Main Summary Measures

For the feasibility analysis the utilisation patterns of Ad26.COV2.S and other COVID-19 vaccines will be characterised and monitored over time. Description of demographics and clinical characteristics will be reported overall and for different groups of vaccine recipients.

The primary analysis will focus on calculating and comparing the incidence rates of each non-acute AESI between individuals exposed to Ad26.COV2.S and (1) unexposed individuals; (2) individuals exposed to another viral vector COVID-19 vaccine (i.e., Vaxzevria® [AZD1222] by Oxford/AstraZeneca); and 3) individuals exposed receiving a mRNA COVID-19 vaccine (cohort).

For acute events, the relative risk between risk window and control window will be estimated (using an SCRI design) among individuals exposed to Ad26.COV2.S. All analyses will be conducted within each data source and pooled across data sources using a random-effects model.

For the COVID-19 vaccines with a 2-dose schedule, AESI-specific risk windows after each dose of the comparator will be considered for the analyses. Further details on how to deal with different doses will be included in the SAP.

8.7.1.1. Feasibility Analysis

Demographic and Baseline Characteristics

The distributions of baseline characteristics at start of COVID-19 vaccination for each COVID-19 vaccine exposure group will be calculated to describe differences between the groups. For continuous variables, means, standard deviations, medians, and other quartiles will be estimated. For categorical variables, counts and proportions will be estimated. The missingness of variables will also be described.

Vaccine uptake

For every data source, the number of administered doses per COVID-19 vaccine brand within the primary series (dose 1 and dose 2) will be calculated by calendar time (in months) over the follow-up period.

For individuals with a given year of birth vaccination uptake for dose 1 over time will be calculated. The coverage at month i for birth year j is calculated by dividing the number of vaccinated individuals n_ij by the total number of individuals under follow-up at month i (N_ij), expressed as a percentage. Estimates of vaccination uptake will be compared with data provided by ECDC.

Algorithms for outcome definitions

For each outcome, the number of cases by diagnosis code will be counted, and the impact on the case count will be assessed when confirmatory medications /procedures or laboratory tests are used. All such counts will be stratified by meaning (GP record, hospitalisation primary diagnosis, hospitalisation secondary diagnosis).

Time-to-onset

Descriptive analyses of the time-to-onset of event (up to one year after vaccination) may inform the selection of risk intervals for the sensitivity analyses.

8.7.1.2. Cohort Design

Demographic and Baseline Characteristics

The distributions of baseline characteristics at index date by exposure group will be calculated to describe the study cohort and illustrate differences between the groups. For continuous variables, means, standard deviations, medians, and other quartiles will be estimated. For categorical variables, counts and proportions will be estimated. The missingness of variables will also be described. Comparative analyses will be included in the final report. Further details will be described in the SAP. To describe the relative imbalance of characteristics between exposed and unexposed groups, absolute standardised differences will be calculated for each baseline characteristic. Multilevel categorical variables will calculate an overall standardised difference across all levels. The larger the absolute standardised difference values, the greater the imbalance between baseline characteristics.

Measures of Occurrence and Association

Incidence rates for each AESI will be calculated by dividing the number of cases by the follow-up person-time. Poisson regression models will estimate incidence rate ratios and 95% CIs. Robust variance estimation will account for an individual's possibility to be included in the exposed and unexposed cohorts. The Poisson model for each AESI will include the number of events in each cohort as the dependent variable, the exposure status as the binary independent variable, and the natural logarithm of the person-time as the offset variable. The attributable risk will be computed as the difference between the incidence rate in the exposed cohort and the incidence rate in the unexposed cohort.

Adjustment for Baseline Imbalances

Individuals following in the exposed and non-exposed groups may have different characteristics that may determine their risk of AESI. To account for such potential confounding, key variables will be matched, and it will be explored whether other covariates are associated with exposure and also whether they are independent risk factors for the outcome (Poisson regression) in the non-exposed. If many factors are imbalanced between the matched cohorts, a propensity score will be created; if only a few are imbalanced, there will be an adjustment for all factors that are both associated with exposure and outcome. Also, if the number of events is low, the use of propensity score methods for adjustment will be considered. More details will be provided in the SAP.

8.7.1.3. Self-controlled Risk Interval Design

Descriptive Statistics

The number of cases of each AESI will be reported in the control and risk windows, overall and by important covariates.

Measures of Association

Conditional Poisson regression will be used to estimate incidence rate ratios and 95% confidence intervals (CIs). The SCRI inherently adjusts for both measured and unmeasured time-constant factors such as sex and chronic health conditions with onset before the start of follow-up. Time-varying confounders (e.g., medications and other vaccinations) may be included as covariates in regression models. Subgroup analyses will be conducted by groups defined by demographic and clinical characteristics and other covariates of interest.

8.7.2. Main Statistical Methods

8.7.2.1. Matching

For the cohort designs, matching will be used to ensure comparability on observed covariates between groups to be compared. There will be a match on the following key variables at index date:

- Age (year of birth)
- Sex
- Pregnancy status
- Immunocompromised state
- Prior COVID-19 infection
- Calendar time (with calliper approach) Prior recorded COVID-19 infection (yes/no)
- Presence of one or more risk factors for severe COVID-19 (e.g., cancer, sickle cell disease, obesity, chronic kidney disease, chronic respiratory disease, human immunodeficiency virus infection).

For the SCRI design, an individual is matched with itself, and therefore analyses will be controlled for all stable variables.

8.7.3. Secondary Analyses

Analyses within the primary study designs (cohort design for non-acute events and SCRI design for acute events) will be conducted for sub-populations:

- Immunocompromised individuals,
- Pregnant women,
- Individuals who have a prior history of thrombotic events and/or thrombocytopenia,
- Prior COVID-19 infection (yes/no),
- Individuals with a prior history (ever) of the specific event more than a year before start of follow-up.

8.7.4. Stratified Analyses

Analyses within the primary study designs (cohort design for non-acute events and SCRI design for acute events) will be stratified by the clinically relevant subgroups below. Stratification will be done also for matching variables when considered relevant.

- Selected comorbidities, including risk factors for severe COVID-19 (by presence or absence of each comorbidity)
- Frailty score (categorised)
- Age (0-17, 18-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80+ years)
- Sex.

8.7.5. Missing Values

Further details of the analysis, including handling of missing data, will be described in the SAP.

8.7.6. Sensitivity Analyses

Sensitivity analyses will be performed to:

- Assess the risk of AESI within extended disease-specific risk windows for events for which the risk interval is not well known or documented in the vaccine safety literature.
- Assess the impact of data sources that are shown not to capture vaccinations well during the feasibility assessment. In the sensitivity analyses with non-exposed comparators those data sources from the cohort analysis for the acute and non-acute events will be excluded, to avoid misclassification of exposure.
- Excluding and censoring upon the administration of other (non-COVID-19) vaccinations in the SCRI and cohort study (see Section 8.1.2.2).
- Comparing Ad26.COV2.S with active comparators with (1) a 1-dose strategy and (2) a 2-dose strategy in the cohort study.
- Conduct a SCRI analysis using a pre-vaccination control window for all acute AESIs.

8.8. Quality Control

Rigorous quality control (QC) will be applied to all deliverables. Data transformation into the CDM will be conducted by each subcontracted research partner in its associated database, with processes as described in the following corresponding sections. Standard operating procedures or internal process guidance at each research centre will be used to guide the conduct of the study. These procedures include rules for secure and confidential data storage, backup, and recovery; methods to maintain and archive project documents; QC procedures for programming; standards for writing analysis plans; and requirements for scientific review by senior staff.

8.8.1. Coordinating Centres

At UMCU, as the scientific coordinating centre responsible for central data management and analysis, all documents undergo QC review and senior scientific review. Data management and

statistical analysis will follow standard operating procedures. All statistical analysis programs will be double-coded.

At Teamit, as the project coordinating centre, all key study documents will undergo QC review and editorial review. Senior reviewers with expertise in the appropriate subject matter area will provide advice on the design of research study approaches and the conduct of the study and will review results, reports, and other key study documents.

8.8.2. VAC4EU

All organisations working on this project will work under the VAC4EU umbrella, which means that VAC4EU will provide access to tools that facilitate the conduct of collaborative distributed studies. As VAC4EU members, the institutions agree to be responsible for implementing and maintaining a standard quality management system with written development procedures and functional area standard operating procedures to ensure that studies are conducted, and data are generated, documented, and reported in compliance with the protocol, accepted standards of Good Clinical and Pharmacoepidemiology Practice and the Accelerated Development of VAccine benefit-risk Collaboration in Europe (ADVANCE) code of conduct, and all applicable federal, state, and local laws, rules and regulations relating to the conduct of the study.

The VAC4EU quality framework is based on the principles of the IMI-ADVANCE code of conduct (Kurz 2017). The quality framework provides requirements and indicators. These requirements are overarching, and each organisation needs to comply with them if they wish to participate in a VAC4EU study.

As information technology tools to support distributed studies, VAC4EU provides partners with:

- A secure VAC4EU SharePoint
- A secure double authentication DRE with study-specific workspaces
- A task management system, with authentication.

VAC4EU also provides members with:

- Work instructions for the SAP
- Access to a series of standard documents and templates for ETL
- Quality verification tools for checking the data: Level 1-3 checks
- A series of verified functions and R-scripts to transform data from the CDM into study variables and evidence.

8.8.3. UMCU (NL)

University Medical Centre Utrecht Julius Centre's department of Data Science and Biostatistics will function as scientific, data engineering and statistical coordinator and adheres to high standards throughout the research process based on robust methods, QC procedures, transparency, and scientific independence. UMCU is an ENCePP centre that conducts studies in accordance with

the ENCePP Guide on Methodological Standards in Pharmacoepidemiology and the ENCePP Code of Conduct. The Julius Centre has standard operating procedures, work instructions, and quality checks conducted for each study. These procedures and documents include internal quality assessments, rules for secure and confidential data storage, methods to maintain and archive project documents, rules and procedures for execution and QC of SAS/R programming, standards for writing protocols, statistical analysis plans, reports and data management.

UMCU personnel is trained in the ConcePTION CDM, coordination of data quality checks 1-3 and able to code and develop R-scripts for the ConcePTION CDM. UMCU has qualified statisticians and data engineers to support the coding of the study-specific R-scripts and trained epidemiologists for scientific coordination.

8.8.4. RTI Health Solutions (RTI-HS)

Standard operating procedures, internal process guidance, or routine practice at each research centre will be used to guide the conduct of the study. These procedures may include, among others, internal quality audits, rules for secure and confidential data storage, methods to maintain and archive project documents, quality control procedures for programming, standards for writing analysis plans, and requirements for senior scientific review.

All programming written by one study analyst will be reviewed independently by a different analyst, with oversight by a senior statistician, if possible. All key study documents, such as the analysis plan, abstraction forms, and study reports, will undergo quality control review, senior scientific review, and editorial review.

A quality assurance audit of this study may be conducted by the sponsor, the sponsor's designees, or a regulatory agency. Note that individual patient-level data are available at the centres only. Selected data fields are not available to be viewed by pharmaceutical companies.

For work conducted at RTI-HS, an independent Office of Quality Assurance will perform internal audits and assessments that involve various aspects of the project, including but not limited to education and training documentation, data entry and data transfer procedures and documentation, and institutional review board (IRB) documentation. Such audits will be conducted by the Office of Quality Assurance according to established criteria in standard operating procedures and other applicable procedures and according to country-specific laws governing audits.

8.8.5. IDIAP Jordi Gol (ES)

IDIAP Jordi Gol will act as a DAP and expert for the SIDIAP data. Data quality processes are implemented at each phase of the data flow cycle. Quality control checks are performed at the extraction and uploading steps of the original SIDIAP databanks prior to release for research. To assess data completeness, the elements present are described by geographical areas, registering physician, time and the distribution function of values. Correctness is assessed by validity checks on outliers, out-of-range values, formatting errors and logical data incompatibilities. Completeness and correctness measures are used to inform decisions on the required transformations to improve data quality (e.g., harmonisation, normalisation, and clean-up) and data fitness for the purpose of

specific research projects. SIDIAP personnel is trained in the conversion of SIDIAP data to the ConcePTION CDM and participating in studies using this CDM. IDIAP Jordi Gol participated in the ACCESS study with background rates of most of the AESI.

8.8.6. FISABIO (ES)

FISABIO will act as a DAP and expert for the Valencian databanks (Valencia Health System Integrated Database, VID).

After Ethical Review Board approval, raw data will be extracted in-text file format and will undergo a data quality check. Data will be stored on secure servers at FISABIO in accordance with Spanish and data protection requirements and ensuring that no identifiable data will be stored longer than required.

All the procedures that will be implemented for data collection, storage, protection, retention and destruction will comply with national and EU legislation. The research team will stay up to date with the detailed provisions of the EU GDPR, which came into force in May 2018, and which will supersede national legislation within the 28 EU Member States.

FISABIO personnel is trained in the conversion of Valencian data to the ConcePTION CDM and participating in studies using this CDM. FISABIO participated in the ACCESS study with background rates of most of the AESI.

8.8.7. PHARMO DATABASE NETWORK (NL)

PHARMO DATABASE NETWORK will act as a DAP and expert on PHARMO data and adheres to high standards throughout the research process based on robust methodologies, transparency, and scientific independence. PHARMO conducts studies in accordance with the ENCePP Guide on Methodological Standards in Pharmacoepidemiology and the ENCePP Code of Conduct. PHARMO DATABASE NETWORK is ISO 9001:2015 certified. Standard operating procedures, work instructions, and checklists are used to guide the conduct of a study. These procedures and documents include internal quality audits, rules for secure and confidential data storage, methods to maintain and archive project documents, rules and procedures for execution and QC of SAS programming, standards for writing protocols and reports, and requirements for senior scientific review of key study documents. PHARMO personnel is trained in the conversion of PHARMO data to the ConcePTION CDM and participating in studies using this CDM. PHARMO DATABASE NETWORK participated in the ACCESS study with background rates of most of the AESI.

8.8.8. ARS Toscana (IT)

Disclaimer: ARS Toscana's participation in the study is subject to the revision of the procedures for data access in the Tuscany Region. See section 8.2.1 regarding conditional participation of ARS Toscana.

ARS will act as a DAP and expert on ARS data. One or two researchers will review the study documents. ARS Toscana receives data on a bimonthly basis from the Tuscany region (where it

undergoes a first QC); the ARS Toscana statistical office appends it to an Oracle database and checks it using a dashboard to identify any inconsistencies with historical data. The Pharmacoepi Unit has standardised parametric procedures in Structured Query Language (SQL) and Stata to extract data from the Oracle database. Parametric procedures are also available to convert the data into various CDMs. ARS personnel is trained in the conversion of ARS data to the ConcePTION CDM and has developed many of the training materials. ARS participated in the ACCESS study with background rates of most of the AESI.

8.9. Limitations of the Research Methods

This study is subject to limitations related to both the study design and the use of secondary healthcare data.

Design

A list of AESI was defined based on Safety Platform for Emergency vACcines (SPEAC) work and additional safety concerns that were raised by the FDA/CDC in the US and EMA in the EU around cases of TTS. The SPEAC list includes a large range of events considered as potential risks that may occur after the administration of a COVID-19 vaccine. The listed AESIs are potentially related to vaccine platforms, natural COVID-19 development, or raised with other marketed vaccines. This study protocol will include all events that are considered as identified risks or risks potentially related to the Ad26.COV2.S vaccine, some of which events were also added as they were observed during the Phase 3 study VAC31518COV3001 and in the US or Europe during the post-authorisation use. Over the 12-month period between vaccine launch and study implementation, the list of AESIs may be amended based on available evidence on the safety profile of the Ad26.COV2.S vaccine.

Given the broad list of potential events that could be associated with the Ad26.COV2.S vaccine and the large heterogeneity in the type of events, acute versus long latency diseases, different study design approaches will be used depending on the type of event. For acute events, the primary study design will be the SCRI design (Baker 2015), which is preferred due to its inherent control for time-constant factors; additionally, because it is restricted to vaccinated individuals, it avoids bias due to misclassification of unexposed individuals due to vaccination outside of traditional medical care settings, which may not be captured the data bank.

Data

Data-related limitations of this study are the reliance on the secondary use of data that has been collected for other purposes. Outcomes and their dates of occurrence will be validated, but the extent of validation may be limited because of the use of medical records. Furthermore, it is unknown the extent to which vaccine brands and batch numbers/lot numbers will be available in the secondary data sources.

Misclassification of COVID-19 vaccination

The COVID-19 vaccines have been rolled out quickly and in special circumstances because of cold-chain requirements and specific risk group channelling. In many healthcare systems,

dedicated systems have been set up to track COVID-19 vaccine administration, in others it is fragmented between primary care and other places. Linkable COVID-19 vaccine registers are in place for Norway, ARS, SIDIAP and FISABIO. For PHARMO information on COVID-19 vaccination is not collected comprehensively in the primary care records, especially not for those vaccines that were administered early and targeted at the frail population. For the SCRI design and acute events, this is not of concern for internal validity, as it will include only individuals vaccinated with the Ad26.COV2.S vaccine. For the non-acute design with non-exposed comparators, misclassification may occur in the non-exposed cohort, which might have been vaccinated with another COVID-19 vaccine. Matching on age will reduce this risk as most of the roll-out strategies were age-based, but still, healthcare workers or other professional groups, which were eligible for early vaccination may be erroneously classified as non-exposed. This would lead to an attenuation of the risk of AESI if it would be associated with the AESI. We will conduct a sensitivity analysis that will exclude non-exposed comparators in databases with misclassification for non-COVID-19 vaccinated

Misclassification of risk window

Experience with other vaccines indicates that safety risks may occur within specific periods post-vaccination, for example, the risk of Guillain-Barré syndrome within 6 weeks after the swine flu vaccine. However, for most of the AESIs included in this study, the risk windows following Ad26.COV2.S vaccines are not clearly established; theoretical risk windows based on previous experience, mechanistic evidence and biological plausibility are considered. For this reason, a subset of the listed AESIs (to be defined in the SAP) will also be assessed within a larger risk window in sensitivity analyses.

For non-acute events, the primary study design will be the cohort design with concurrent matched unexposed comparators. The advantage is that misclassification errors will be reduced because of misspecification of the risk window that may happen in the SCRI, but at the same time, person-time at risk because of the larger time windows may be diluted, which would dilute a potential increased risk.

Confounding

COVID-19 is a complex disease for which people at higher risk of developing severe COVID-19 have been identified (Centers for Disease Control and Prevention 2020). The proposed matched cohort design (primary design for non-acute events) will allow adjustment for observed potential confounders at the time of the reference date, and the SCRI design (primary design for acute events) inherently adjusts for all time-constant factors. In this observational database study conducted using heterogeneous electronic health record databases, confounders that can influence vaccine behaviours such as beliefs, educational level, or socioeconomic status cannot be considered because they are either not recorded or incompletely recorded in the data. Even though many potential confounders will be included in this study, there may be residual bias due to unmeasured, unreported, or misclassified confounders in the cohort. As mentioned before, the SCRI design will be able to adjust for the time-constant characteristics (e.g., socioeconomic status, beliefs), but this will be limited to the acute AESIs. For most acute events, infections and vaccinations are time-varying risk factors, which can be captured by most of the data sources.

Unmeasured confounding can be individually addressed using quantitative bias analysis methods. For each potential confounder of interest that is not available in the study data sources, three bias parameters will be estimated using available literature or knowledge from the study data source: (1) the expected association between the unmeasured confounder and the incidence rate ratio (IRR) (IRRCD), (2) the prevalence of the confounder among those exposed (p1), and (3) the prevalence of the confounder among those not exposed (p0). Based on these three bias parameters, if IRRObs is the observed IRR without adjustment for the unmeasured confounder, then an IRR adjusted for the unmeasured confounder (IRRAdj, UC) will be computed. Specific analytic details will be included in the SAP.

Outcome misclassification

Algorithms including medical encounters with diagnosis codes and/or medications have been and will be developed to identify the AESI that may occur in both study cohorts. Most of the data sources have been developed already for ACCESS. Case validation to confirm the disease diagnosis and date of onset may be conducted on AESIs that are likely to be misclassified in electronic data. The validation will be based on manual medical record review by a medically trained person depending on the AESI.

9. PROTECTION OF HUMAN SUBJECTS

This is a non-interventional study using secondary data collection and does not pose any risks for individuals. Each data source research partner will apply for an independent ethics committee review according to local regulations.

Data protection and privacy regulations will be observed in collecting, forwarding, processing, and storing data from study participants.

9.1. Individual Information

This study mainly involves data that exist in anonymised/pseudonymised structured format and contains no patient personal information. No patient names or any identifiable information will be shared and if based on validation, it will be replaced by a single, specific, numerical code. No identifiable level data will be transferred to the sponsor or any organisation. Data remain local and only aggregated data that is the result of the analytical script implementing the study design will be transferred to the DRE for postprocessing and pooling. When there are privacy concerns around anonymised SCRI data sharing, postprocessing can be done locally and only coefficients can be shared.

9.2. Patient Consent

As this study does not involve data subject to privacy laws according to applicable legal requirements, obtaining informed consent from individuals is not required.

9.3. Institutional Review Board /Independent Ethics Committee

Each DAP will be following the local country and institutional requirements to apply for access and analysis of the data for this study. At the coordinating centre, UMCU will ask for approval for exemption from review by the UMCU International IRB. All correspondence with the IRB or independent ethics committee and applicable documentation will be retained as part of the study materials.

When establishing the agreements with the data sources to conduct the study, it will be emphasised that the current list of AESIs may be expanded during the course of the study to accommodate newly identified AESIs.

9.4. Ethical Conduct of the Study

This study will adhere to the Guidelines for GPP and has been designed in line with the ENCePP Guide on Methodological Standards in Pharmacoepidemiology. The ENCePP Checklist for Study Protocols will be completed (see ANNEX 1). The study is a post-authorisation study of vaccine safety and will comply with the definition of the non-interventional (observational) study referred to in the International Conference on Harmonisation tripartite guideline Pharmacovigilance Planning E21 and provided in the EMA Guideline on Good Pharmacovigilance Practices (GVP) Module VIII: Post-Authorisation Safety Studies and with the 2012 EU pharmacovigilance legislation, adopted 19 June 2012. The study will be registered in the EU PAS Register before data collection commences. The research team and study sponsor should adhere to the general

principles of transparency and independence in the ENCePP Code of Conduct and the ADVANCE Code of Conduct.

The study will be conducted in accordance with legal and regulatory requirements, as well as with scientific purpose, value, and rigour and follow generally accepted research practices described in Guidelines for GPP issued by the International Society for Pharmacoepidemiology, and Good Epidemiological Practice guidelines issued by the International Epidemiological Association. An independent scientific advisory board will be installed, comprising experts in vaccine safety studies.

10. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

For studies in which the research team uses only data from automated healthcare databases, according to the International Society for Pharmacoepidemiology Guidelines for GPP.

"Aggregate analysis of database studies can identify an unexpected increase in risk associated with a particular exposure. Such studies may be reportable as study reports, but typically do not require reporting of individual cases. Moreover, access to automated databases does not confer a special obligation to assess and/or report any individual events contained in the databases. Formal studies conducted using these databases should adhere to these guidelines."

For non-interventional study designs that are based on secondary use of data, such as studies based on medical record reviews or electronic health records, systematic reviews, or meta-analyses, reporting of adverse events/adverse drug reactions is not required. Reports of adverse events/adverse drug reactions should only be summarised in the study report, where applicable.

According to the EMA Guideline on GVP, Module VI – Management and Reporting of Adverse Reactions to Medicinal Products,

"All adverse events/reactions collected as part of [non-interventional post-authorisation studies with a design based on secondary use of data], the submission of suspected adverse reactions in the form of [individual case safety reports] is not required. All adverse events/reactions collected for the study should be recorded and summarised in the interim safety analysis and in the final study report."

Module VIII – Post-Authorisation Safety Studies echoes this approach. Legislation in the EU further states that for certain study designs such as retrospective cohort studies, particularly those involving electronic health records, it may not be feasible to make a causality assessment at the individual case level.

11. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

Results of analysis and interpretation will be delivered in the form of reports. At the end of the study follow-up period, the final report will be produced including the analysis and interpretation of each outcome.

Study results will be published following guidelines, including those for authorship, established by the International Committee of Medical Journal Editors. When reporting the results of this study, the appropriate Strengthening the Reporting of Observational Studies in Epidemiology checklist will be followed. Independent publication rights will be granted to the research team in line with Section VIII.B.5., Publication of study results, of the EMA Guideline on GVP Module VIII: Post-Authorisation Safety Studies.

Communication via appropriate scientific venues will be made according to the ENCePP Code of Conduct.

12. REFERENCES

ACCESS 2021 Available at: https://drive.google.com/drive/folders/1Y_3cuGRN1g-jBv2ec1fC0aYcpxEjtrY9?usp=sharing.

Baker MA, Lieu TA, Li L, Hua W, Qiang Y, Kawai AT, Fireman BH, Martin DB, Nguyen MD.Baker MA, et al. A vaccine study design selection framework for the postlicensure rapid immunization safety monitoring program Am J Epidemiol. 2015 Apr 15;181(8):608-18.

Berry JD, Jones S, Drebot MA, et al. Development and characterization of neutralising monoclonal antibody to the SARS-coronavirus. J Virol Methods. 2004;120(1):87-96.

Bisht H, Roberts A, Vogel L, et al. Severe acute respiratory syndrome coronavirus spike protein expressed by attenuated vaccinia virus protectively immunizes mice. Proc Natl Acad Sci USA. 2004;101(17):6641-6646.

Brighton Collaboration. Interim case definition of Thrombosis with Thrombocytopenia Syndrome (TTS) https://brightoncollaboration.us/thrombosis-with-thrombocytopenia-syndrome-interim-case-definition/ Accessed on 8 May 2021.

Buchholz UJ, Bukreyev A, Yang L, et al. Contributions of the structural proteins of severe acute respiratory syndrome coronavirus to protective immunity. Proc Natl Acad Sci USA. 2004;101(26):9804-9809.

Bukreyev A, Lamirande EW, Buchholz UJ, et al. Mucosal immunisation of African green monkeys (Cercopithecus aethiops) with an attenuated parainfluenza virus expressing the SARS coronavirus spike protein for the prevention of SARS. Lancet. 2004;363(9427):2122-2127.

Centers for Disease Control and Prevention. Coronavirus Disease 2019 (COVID-19) Symptoms and Complications. https://www.cdc.gov/coronavirus/2019-ncov/about/symptoms.html. Accessed 15 February 2020.

Cines DB and Bussel JB. SARS-CoV-2 Vaccine-Induced Immune Thrombotic Thrombocytopenia. N Engl J Med. 2021;384(23):2254-2256.

Clegg A, Bates C, Young J, et al. Development and validation of an electronic frailty index using routine primary care electronic health record data. Comparative Study Age Ageing. 2016;45(3):353-360.

Coronaviridae Study Group of the International Committee on Taxonomy of Viruses. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nat Microbiol. 2020;5(4):536-544.

Cyranoski D. The biggest mystery: what it will take to trace the coronavirus source. Nature. 2020.

European Centre for Disease Prevention and Control (2020). COVID-19 Situation update for the EU/EEA and the UK, as of 22 September 2020. Available at: https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-casesa. Accessed: 1 September 2021.

European Commission. Preparedness for COVID-19 vaccination strategies and vaccine deployment https://ec.europa.eu/health/sites/health/files/vaccination/docs/2020_strategies_deployment_en.pdf Accessed 10 November 2020.

European Medicines Agency (2021). EMA recommends COVID-19 Vaccine Janssen for authorisation in the EU. 11 March 2021 https://www.ema.europa.eu/en/news/ema-recommends-covid-19-vaccine-janssen-authorisation-eu (accessed 8 May 2021).

Dodd C, Willame C et al. Rapid safety assessment of COVID-19 vaccines in electronic healthcare databases: a protocol template from the ACCESS project. 2020.
Available at: http://www.encepp.eu/encepp/viewResource.htm?id=39362. Accessed: 10 May 2021.

European Union. Council Regulation (EU) 2016/369 of 15 March 2016 on the provision of emergency support within the Union: https://eurlex.europa.eu/eli/reg/2016/369/oj Accessed 10 November 2020.

Greinacher A, Thiele T, Warkentin TE, Weisser K, Kyrle PA, Eichinger S. Thrombotic Thrombocytopenia after ChAdOx1 nCov-19 Vaccination. N Engl J Med. 2021;384(22):2092-2101.

Janssen Vaccines & Prevention B.V. Data on file.

Johns Hopkins CSSE. Coronavirus COVID-19 Global Cases. https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6. Accessed 1 September 2021.

Karron RA, Key NS, Sharfstein JM. Assessing a Rare and Serious Adverse Event Following Administration of the Ad26.COV2.S Vaccine. JAMA. 2021;325(24):2445-2447.

Kurz X, Bauchau V, Mahy P, Glismann S, van der Aa LM, Simondon F; ADVANCE consortium. The ADVANCE Code of Conduct for collaborative vaccine studies. Vaccine. 2017;35(15):1844-1855.

Li Q, Guan X, Wu P, et al Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. N Engl J Med. 2020;382:119-1207.

Lu R, Zhao X, Li J, et al Genomic characterization and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet. 2020;395(10224):565-574.

Medicines and Healthcare products Regulatory Agency (MHRA). Regulatory approval of COVID-19 Vaccine Janssen. Dated 23 June 2021. https://www.gov.uk/government/publications/regulatory-approval-of-covid-19-vaccine-janssen. Accessed 03 September 2021.

Rothman KJ and Greenland S. Planning Study Size Based on Precision Rather Than Power. Epidemiology. 2018;29(5):599-603.

Schultz NH, Sørvoll IH, Michelsen AE, et al. Thrombosis and Thrombocytopenia after ChAdOx1 nCoV-19 Vaccination. N Engl J Med. 2021;384(22):2124-2130.

Scully M, Singh D, Lown R, et al. Pathologic Antibodies to Platelet Factor 4 after ChAdOx1 nCoV-19 Vaccination. N Engl J Med. 2021;384(23):2202-2211.

Subbarao K, McAuliffe J, Vogel L, et al Prior infection and passive transfer of neutralizing antibody prevent replication of severe acute respiratory syndrome coronavirus in the respiratory tract of mice. J Virol. 2004;78(7):3572-3577.

Sui J, Li W, Roberts A, et al Evaluation of human monoclonal antibody 80R for immunoprophylaxis of severe acute respiratory syndrome by an animal study, epitope mapping, and analysis of spike variants. J Virol. 2005;79(10):5900-5906.

US Food and Drug Administration (2021). FDA Issues Emergency Use Authorization for Third COVID-19 Vaccine 27 February 2021. https://www.fda.gov/news-events/press-announcements/fda-issues-emergency-use-authorization-third-covid-19-vaccine.

VAC4EU FAIR Catalogue that Molgenis. https://www.molgenis.org/communities.html. Accessed 03 September 2021.

Willame C et al. (2021). Background rates of Adverse Events of Special Interest for monitoring COVID-19 vaccines, an ACCESS study. D3-Draft Final Report, April 30, 2021. Available at: http://www.encepp.eu/documents/ACCESS_Report_BGR_20210430_v.1.2.pdf

Wu F (2020), Zhao S, Yu B, et al A new coronavirus associated with human respiratory disease in China. Nature. 2020;579(7798):265-269.

Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome, NC_045512.2

Yadav PD, Sapkal GN, Abraham P, et al. Neutralization of variant under investigation B.1.617 with sera of BBV152 vaccinees. Clin Infect Dis. 2021.

Zhang H (2004), Wang G, Li J, et al Identification of an antigenic determinant on the S2 domain of the severe acute respiratory syndrome coronavirus spike glycoprotein capable of inducing neutralizing antibodies. J Virol. 2004;78(13):6938-6945.

Zhou T (2004), Wang H, Luo D, et al An exposed domain in the severe acute respiratory syndrome coronavirus spike protein induces neutralizing antibodies. J Virol. 2004;78(13):7217-7226.

ANNEX 1: ENCEPP CHECKLIST FOR STUDY PROTOCOLS (REVISION 4)

Study title: An Observational Post-Authorisation Safety Study to Assess the Safety of Ad26.COV2.S Using European Healthcare Data through VAC4EU

EU PAS Register® number: EUPAS45362
Study reference number (if applicable): VAC31518COV4003

Section 1: Milestones	**Yes**	**No**	**N/A**	**Section Number**
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection[†]	☒	☐	☐	5
1.1.2 End of data collection[‡]	☒	☐	☐	5
1.1.3 Progress report(s)	☒	☐	☐	5
1.1.4 Interim report(s)	☐	☒	☐	5
1.1.5 Registration in the EU PAS Register®	☒	☐	☐	5
1.1.6 Final report of study results.	☒	☐	☐	5

Section 2: Research question	**Yes**	**No**	**N/A**	**Section Number**
2.1 Does the formulation of the research question and objectives clearly explain:	☒	☐	☐	7
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)	☒	☐	☐	7
2.1.2 The objective(s) of the study?	☒	☐	☐	7
2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	☐	☐	☐	8.2.2
2.1.4 Which hypothesis(-es) is (are) to be tested?	☐	☐	☒	
2.1.5 If applicable, that there is no *a priori* hypothesis?	☐	☐	☒	

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)	☒	☐	☐	8.1
3.2 Does the protocol specify whether the study is based on primary, secondary or combined data collection?	☒	☐	☐	9

[†] 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.

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

Section 3: Study design		**Yes**	**No**	**N/A**	**Section Number**
3.3	Does the protocol specify measures of occurrence? (e.g., rate, risk, prevalence)	☒	☐	☐	8.7
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))	☐	☐	☐	8.7
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)	☒	☐	☐	8.3.3

Section 4: Source and study populations		**Yes**	**No**	**N/A**	**Section Number**
4.1	Is the source population described?	☒	☐	☐	8.2
4.2	Is the planned study population defined in terms of:				
	4.2.1 Study time period	☒	☐	☐	8.2.1
	4.2.2 Age and sex	☒	☐	☐	8.2.2
	4.2.3 Country of origin	☒	☐	☐	8.2.1
	4.2.4 Disease/indication	☒	☐	☐	8.1
	4.2.5 Duration of follow-up	☒	☐	☐	8.2.3
4.3	Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	☒	☐	☐	8.2.4

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)	☒	☐	☐	8.3.1
5.2	Does the protocol address the validity of the exposure measurement? (e.g. precision, accuracy, use of validation sub-study)	☐	☒	☐	
5.3	Is exposure categorised according to time windows?	☐	☒	☐	
5.4	Is intensity of exposure addressed? (e.g. dose, duration)	☐	☒	☐	
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?	☐	☒	☐	

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?	☒	☐	☐	8.3.2
6.2 Does the protocol describe how the outcomes are defined and measured?	☒	☐	☐	8.3.2
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)	☒	☐	☐	8.3.3.1
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)	☐	☒	☐	

Section 7: Bias	**Yes**	**No**	**N/A**	**Section Number**
7.1 Does the protocol address ways to measure confounding? (e.g. confounding by indication)	☒	☐	☐	8.9
7.2 Does the protocol address selection bias? (e.g. healthy user/adherer bias)	☒	☐	☐	8.9
7.3 Does the protocol address information bias? (e.g. misclassification of exposure and outcomes, time-related bias)	☒	☐	☐	8.9

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)	☒	☐	☐	8.3.4

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:				
9.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview)	☒	☐	☐	8.4
9.1.2 Outcomes? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics)	☒	☐	☐	8.4
9.1.3 Covariates and other characteristics?	☒	☐	☐	8.4
9.2 Does the protocol describe the information available from the data source(s) on:				8.4
9.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☒	☐	☐	8.4

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.2.2 Outcomes? (e.g. date of occurrence, multiple event, severity measures related to event)	☒	☐	☐	8.4
9.2.3 Covariates and other characteristics? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	☒	☐	☐	8.4
9.3 Is a coding system described for:				
9.3.1 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	☒	☐	☐	8.4
9.3.2 Outcomes? (e.g. International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))	☒	☐	☐	8.4
9.3.3 Covariates and other characteristics?	☒	☐	☐	8.4
9.4 Is a linkage method between data sources described? (e.g. based on a unique identifier or other)	☒	☐	☐	8.4

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Section Number**
10.1 Are the statistical methods and the reason for their choice described?	☒	☐	☐	8.7
10.2 Is study size and/or statistical precision estimated?	☒	☐	☐	8.5
10.3 Are descriptive analyses included?	☒	☐	☐	8.7.7.1
10.4 Are stratified analyses included?	☒	☐	☐	8.7.4
10.5 Does the plan describe methods for analytic control of confounding?	☒	☐	☐	8.1.2
10.6 Does the plan describe methods for analytic control of outcome misclassification?	☒	☐	☐	8.9
10.7 Does the plan describe methods for handling missing data?	☒	☐	☐	8.7.5
10.8 Are relevant sensitivity analyses described?	☒	☐	☐	8.7.6

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)	☒	☐	☐	8.6
11.2 Are methods of quality assurance described?	☒	☐	☐	8.8
11.3 Is there a system in place for independent review of study results?	☒	☐	☐	9.4

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?	☒	☐	☐	8.9

Section 12: Limitations	**Yes**	**No**	**N/A**	**Section Number**
12.1.2 Information bias?	☒	☐	☐	8.9
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).	☒	☐	☐	8.9
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)	☒	☐	☐	8.9

Section 13: Ethical/data protection issues	**Yes**	**No**	**N/A**	**Section Number**
13.1 Have requirements of Ethics Committee/ Institutional Review Board been described?	☒	☐	☐	9
13.2 Has any outcome of an ethical review procedure been addressed?	☐	☐	☒	
13.3 Have data protection requirements been described?	☒	☐	☐	8.4/8.6

Section 14: Amendments and deviations	**Yes**	**No**	**N/A**	**Section Number**
14.1 Does the protocol include a section to document amendments and deviations?	☒	☐	☐	3

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)?	☒	☐	☐	11
15.2 Are plans described for disseminating study results externally, including publication?	☒	☐	☐	11

SPONSOR AND MAIN AUTHOR'S SIGNATURE

I have read this protocol and agree that it contains all the necessary details for carrying out this study. I will conduct the study as outlined herein and will complete the study within the time designated.

I will provide copies of the protocol and all pertinent information to all individuals responsible to me who assist in the conduct of this study. I will discuss this material with them to ensure that they are fully informed regarding the conduct of the study and the obligations of confidentiality.

Sponsor's Responsible Contact Person:

Name (typed or printed): PPD

Institution and Address: Janssen Vaccines & Prevention B.V.

Signature: _____ Date: _____
(Day Month Year)

Principal Investigator (Main author):

Name (typed or printed): PPD

Institution: UMCU, The Netherlands

Signature: _____ Date: _____
(Day Month Year)

Note: If the address or telephone number of the participating physician changes during the course of the study, written notification will be provided to the sponsor; a protocol amendment will not be required.