NON-INTERVENTIONAL (NI) STUDY PROTOCOL

PASS information

Title	Post-Authorisation Safety Study of Comirnaty Original/Omicron BA.1 and Comirnaty Original/Omicron BA.4-5 in Europe
Protocol number	C4591052
Protocol version identifier	V3.0
Date	08 January 2024
EU Post Authorization Study (PAS) register number	Study to be registered prior to start of data collection.
Active substance	Riltozinameran (ATC code J07BX03) is a single-stranded, 5'-capped messenger RNA (mRNA) produced using cell-free *in vitro* transcription from the corresponding DNA templates, encoding the viral spike (S) protein of SARS-CoV-2 (Omicron BA.1). Famtozinameran (ATC code J07BX03) is a single-stranded, 5'-capped messenger RNA (mRNA) produced using cell-free *in vitro* transcription from the corresponding DNA templates, encoding the viral spike (S) protein of SARS-CoV-2 (Omicron BA.4-5).
Medicinal product	Comirnaty Original/Omicron BA.1 and Comirnaty Original/Omicron BA.4-5
Marketing Authorization Holder(s) (MAH)	BioNTech Manufacturing GmbH
Joint PASS	No

Research question and objectives	The research question is: Is there an increased risk of pre-specified adverse events of special interest (AESIs) after vaccination with Comirnaty Original/Omicron BA.1 (bivalent BA.1) or Comirnaty Original/Omicron BA.4-5 (bivalent BA.4-5) compared with no vaccination against COVID-19 among individuals with comparable vaccination history? *Primary objective* To determine whether there is an increased risk of pre-specified AESIs following the administration of bivalent BA.1 or bivalent BA.4-5 compared with not receiving any COVID-19 vaccine during follow-up.
Country(-ies) of study	Italy (IT), The Netherlands (NL), Norway (NO), Spain (ES), United Kingdom (UK)
Authors	Xabier García de Albéniz, Rachel Weinrib, Alejandro Arana, RTI Health Solutions, Barcelona, Spain Daniel Weibel, University Medical Center Utrecht, Utrecht, The Netherlands

Marketing Authorization Holder(s)

Marketing Authorization Holder(s)	BioNTech Manufacturing GmbH An der Goldgrube 12 55131 Mainz Germany
MAH contact person	Constanze Blume

1. TABLE OF CONTENTS

2. LIST OF ABBREVIATIONS

Abbreviation	Definition
ACCESS	vACcine Covid-19 monitoring readinESS
ADVANCE	Accelerated Development of VAccine benefit-risk Collaboration in Europe
AESI	adverse event of special interest
CDC	Centers for Disease Control and Prevention
CDM	common data model
CI	confidence interval
COVID-19	coronavirus disease 2019
CPRD	Clinical Practice Research Datalink (UK)
CRF	case report form
CSV	comma-separated values
DAP	database access provider
DCT	data collection tools
DRE	Digital Research Environment (NL)
DSRU	Drug Safety Research Unit (UK)
EHR	electronic health record
EMA	European Medicines Agency
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
EpiChron	EpiChron Research Group on Chronic Diseases at the Aragon Health Sciences Institute (Spain)
ES	Spain
ETL	extraction, transformation, and loading
EU PAS Register	European Union electronic register of post-authorisation studies
EU	European Union
FAIR	findable, accessible, interoperable, and re-usable
GP	general practitioner
GPP	Good Pharmacoepidemiology Practices

Abbreviation	Definition
GVP	Good Pharmacovigilance Practices
HES	Hospital Episode Statistics
IACS	Instituto Aragonés de Ciencias de la Salud
IR	incidence rate
IT	Italy
KUHR	Norway Control and Payment of Health Reimbursement
MBRN	Medical Birth Registry of Norway
mRNA	messenger ribonucleic acid
NHR	Norwegian health registers
NHS	National Health Service
NIS AEM	non-interventional study adverse event monitoring
NL	The Netherlands
NO	Norway
NorPD	Norwegian Prescription Database
ONS	Office for National Statistics
PASS	post-authorisation safety study
PHARMO	PHARMO Institute for Drug Outcomes Research or PHARMO Database Network (Netherlands)
QC	quality control
RTI-HS	RTI Health Solutions
SAB	scientific advisory committee
SAP	statistical analysis plan
SARS-CoV-2	severe acute respiratory syndrome coronavirus 2
SCRI	self-controlled risk interval
SIDIAP	Sistema d'Informació per el Desenvolupament de la Investigació en Atenció Primària [Information System for the Improvement of Research in Primary Care] (Spain)
SQL	Structured Query Language

Abbreviation	Definition
SSB	Statistics Norway
SYSVAK	Nordic national, electronic immunisation register
TSD	Tjeneste for Sensitive Data
TTS	thrombotic thrombocytopenia syndrome
UK	United Kingdom
UMCU	University Medical Center Utrecht
VAC4EU	Vaccine monitoring Collaboration for Europe
WHO	World Health Organization
YRR	Your Reporting Responsibilities

3. RESPONSIBLE PARTIES

Principal Investigator(s) of the Protocol

Name, degree(s)	Job Title	Affiliation	Address
Cynthia De Luise, PhD, MPH	Senior Director, Safety Surveillance Research	Pfizer, Inc.	66 Hudson Blvd East New York, NY 10001-2192, USA
Daniel Weibel, PhD	Assistant Professor	University Medical Center Utrecht	Internal mail no Str 6.131 | PO Box 85500 | 3508 GA Utrecht, The Netherlands
Alejandro Arana, MD, MPH	Senior Director, Epidemiology	RTI Health Solutions	Av. Diagonal, 605, 9-1, 08028 Barcelona, Spain
Miriam Sturkenboom, PhD	Professor	University Medical Center Utrecht	Internal mail no Str 6.131 | PO Box 85500 | 3508 GA Utrecht, The Netherlands
Constanza Andaur, DDS, PhD	Assistant Professor	University Medical Center Utrecht	Internal mail no Str 6.131 | PO Box 85500 | 3508 GA Utrecht, The Netherlands
Amirreza Dehghan Tarazjani, MD	Research Assistant	University Medical Center Utrecht	Internal mail no Str 6.131 | PO Box 85500 | 3508 GA Utrecht, The Netherlands
Xabier Garcia de Albeniz, MD, PhD	Director, Epidemiology	RTI Health Solutions	Av. Diagonal, 605, 9-1, 08028 Barcelona, Spain
Jaume Aguado, PhD	Director, Biostatistics	RTI Health Solutions	Av. Diagonal, 605, 9-1, 08028 Barcelona, Spain
Rachel Weinrib, MPH	Research Epidemiologist	RTI Health Solutions	Av. Diagonal, 605, 9-1, 08028 Barcelona, Spain
Julia Vila Guilera, PhD	Research Epidemiologist	RTI Health Solutions	Av. Diagonal, 605, 9-1, 08028 Barcelona, Spain

Country Coordinating Investigators

Name, degree(s)	Job Title	Affiliation	Address
Rosa Gini, PhD	Head Pharmacoepidemiology Unit	Agenzia Regionale di Sanità della Toscana	Via Pietro Dazzi 1, 50141 Firenze, Italy
Ron Herings, PhD	Director	PHARMO Institute for Drug Outcomes Research	Van Deventerlaan 30-40 3528 AE Utrecht The Netherlands
Karin Swart, PhD	Senior Research Manager		
Jetty Overbeek, PhD	Senior Research Manager		
Jesse van den Berg, PhD	Researcher		
Carlo Giaquinto, MD	President	PENTA foundation	Corso Stati Uniti 4 35127 Padova Italy

Name, degree(s)	Job Title	Affiliation	Address
Saad Shakir, MB ChB, LRCP&S, FRCP, FFPM, FISPE, MRCGP	Director	Drug Safety Research Unit (DSRU)	Drug Safety Research Unit Bursledon Hall Blundell Lane Southampton, Hampshire SO31 1AA United Kingdom
Debabrata Roy, PhD	Head of Epidemiology and Research		
Samantha Lane, MSc, PgDip, BSc	Senior Research Fellow		
Catherine Fry, PhD	Research Fellow		
Kathryn Morton, BSc, BM, MRCGP	Clinical Research Fellow		
Alexandra Prados Torres, MD, MPH, PhD	National Health Service (NHS) Senior Researcher	Instituto Aragonés de Ciencias de la Salud (IACS)	Hospital Universitario Miguel Servet Paseo Isabel la Católica 1-3 50009, Zaragoza, Spain
Antonio Gimeno Miguel, MPH, PhD	Researcher		
Beatriz Poblador-Plou, MPH, PhD	Researcher		
Felipe Villalobos, MD, PhD	Researcher	IDIAP-Jordi Gol	Gran Via Corts Catalanes 587 08007, Barcelona, Spain
Mertixell Pallejà, MSc	Statistician		
Carlo Alberto Bissacco, MSc	Statistician		
Elena Segundo, MSc	Data Scientist		
Angela Lupattelli, PhD	Professor	University of Oslo	PO 1068 Blindern 0316 Oslo, Norway

Note: Country coordinating investigators have reviewed and contributed to this protocol.

Other Responsible Parties Contributing to the Protocol

Name, degree(s)	Job Title	Affiliation	Address
Natasha Yefimenko, PhD	Program Manager	Teamit Institute	c/ Sant Antoni María Claret, 167, 08025 Barcelona, Spain
Elena Palà, PhD	Study Manager		
Margaret Haugh, PhD	Medical Writer	MediCom Consult	39 rue Clément Michut 69100 Villeurbanne France

4. ABSTRACT

Title: Post-Authorisation Safety Study of Comirnaty Original/Omicron BA.1 and Comirnaty Original/Omicron BA.4-5 in Europe

Version 3.0, 08 January 2024

Authors: Xabier García de Albéniz, Rachel Weinrib, Alejandro Arana, RTI Health Solutions, Barcelona, Spain; Daniel Weibel, University Medical Center Utrecht, Utrecht, The Netherlands

Rationale and background: In November 2021, World Health Organization (WHO) designated the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variant B.1.1.529 a variant of concern and named it Omicron. The Omicron variant, and its descendant sub-lineages, such as BA.4-5, have since spread to become the dominant variant of the virus circulating worldwide. The Omicron variant has greatly increased transmissibility and immune-evasion capacity compared with prior circulating variants. In addition, there is evidence that the original mRNA vaccines provide lower protection against the Omicron variant.

Bivalent formulations (Comirnaty Original/Omicron BA.1 (bivalent BA.1) and Comirnaty Original/Omicron BA.4-5 (bivalent BA.4-5)) of Pfizer-BioNTech's Coronavirus Disease 2019 (COVID-19) mRNA-based vaccine were authorised as booster doses in the European Union (EU) on 01 September 2022 and 12 September 2022, respectively. On 06 December 2022, the Emergency Task Force of the European Medicines Agency (EMA) concluded that "bivalent BA.4-5 may be used as a primary vaccination series in adult and children" who have not been previously vaccinated against SARS-CoV-2".[1] Pfizer and Vaccine monitoring Collaboration for Europe (VAC4EU) are conducting this study to monitor the safety of these bivalent formulations in European countries.

Research question and objectives: Is there an increased risk of pre-specified adverse events of special interest (AESIs) after vaccination with bivalent BA.1 or bivalent BA.4-5 compared with no vaccination against COVID-19 among individuals with comparable vaccination histories?

Primary study objective
To determine whether there is an increased risk of pre-specified AESIs following the administration bivalent BA.1 or bivalent BA.4-5 compared with not receiving any COVID-19 vaccine during follow-up.

Study design: A retrospective cohort design will be used to estimate the incidence of AESIs after receiving a Pfizer-BioNTech COVID-19 bivalent vaccine, and these incidences will be compared with those in a comparator group that did not receive any COVID-19 vaccine during follow-up. Exposed individuals will be matched to unexposed individuals using relevant individual characteristics listed below. For selected AESIs a self-controlled risk interval (SCRI) study design will also be used, when appropriate.

Population: The source population will comprise all individuals registered in each of the health care data sources who are eligible to receive bivalent BA.1 or bivalent BA.4-5. The study period will start on the date of availability of the bivalent BA.1, which was the first bivalent vaccine to receive authorisation in the EU (on 01 September 2022), in each participating country and will end on 31 August 2024 or the date of the latest data availability. BA.4-5 received authorisation in the EU on 12 September 2022.

Individuals will be evaluated for eligibility and time zero will be determined as the date of exposure (vaccination with bivalent BA.1 or bivalent BA.4-5. Matching will occur at time zero and follow-up will begin at time zero. Individuals who have received at least one dose of bivalent BA.1 or bivalent BA.4-5 will be included in the exposed cohort. Individuals who have not received a dose of any COVID-19 vaccine at time zero will be included in the unexposed cohort.

Variables: Exposure will be based on recorded prescription, dispensing, or administration of the bivalent BA.1 and bivalent BA.4-5 vaccines. Outcomes will be based on the AESIs proposed by the European Medicines Agency (EMA)-sponsored ACCESS project (vACcine COVID-19 monitoring readinESS) and also, for consistency, those included in the ongoing post-authorisation safety study of the Pfizer BioNTech COVID-19 monovalent vaccine (EUPAS41623). Clinical validation, using individual health records, will be carried out for specific AESIs. This manual review will be conducted by clinicians blinded to COVID-19 vaccine exposure. Confirmation of the certainty of an event diagnosis will be classified using available Brighton Collaboration case definitions and, those from other organisations, if Brighton Collaboration case definitions are not available.

The following and other variables will be considered as potential confounders in the matched cohort design and, those listed below will be used as matching variables:

- age,
- sex,
- place of residence (as available),
- prior COVID-19 vaccination history,
- history of hospitalization for COVID-19
- pregnancy status and duration of pregnancy
- immunocompromised status,
- COVID-19 disease risk criteria, as defined by the US Centers for Disease Control and Prevention,
- socioeconomic status/education level (as available).

The list of potential confounders may be updated, based on emerging scientific evidence.

Data sources: The study will be performed within the following selected data sources: Pedianet (IT), PHARMO Institute for Drug Outcomes Research (PHARMO) (NL), the Norwegian health registers (NHR) (NO), EpiChron Research Group on Chronic Diseases at the Aragon Health Sciences Institute (EpiChron) (ES), Sistema d'Informació per el Desenvolupament de la Investigació en Atenció Primària (SIDIAP) [Information System for the Improvement of Research in Primary Care] (ES), and Clinical Practice Research Datalink (CPRD) (UK).

Study size: The sample size will be determined by the uptake of bivalent BA.1 and bivalent BA.4-5 during the study period.

Data analysis: Data from the matched cohort design will be analysed as follows:

- Conditional exchangeability: The pairs will be matched using several variables considered as potential confounders to ensure conditional exchangeability. Additional standard epidemiological methods, based on propensity scores, will be used to improve adjustment for confounding, if necessary.
- The effect estimates will be reported as risk ratios and risk differences (and their corresponding 95% CIs) for those exposed to a Pfizer-BioNTech COVID-19 bivalent vaccine compared with those not exposed to any COVID-19 vaccine during follow-up.
- Appropriate data analysis models will be used to estimate the incidence rate ratios of AESIs in the risk and the control windows in the SCRI study.

Milestones:

Milestone	Planned Date[a]
Registration in EU PAS Register	After regulatory endorsement and prior to the start of data collection
Start of data collection[b]	31 March 2024
End of data collection	30 September 2025
Interim report 1[c]	14 May 2025
Final study report[c,d]	30 April 2026

EU PAS Register; European Union Electronic Register of Post-Authorisation Studies.
a Planned dates are dependent on protocol endorsement date, uptake of Comirnaty Original/Omicron BA.1 and Comirnaty Original/Omicron BA.4-5, approvals for data extraction and contracts with data access providers (DAPs).
b Start of data collection: the date on which information about the first individual is recorded in the study data set or, in the case of secondary use of data, the date from which data extraction starts [IR Art 37(1)]. Simple counts in a database to support the development of the study protocol, for example, to inform the sample size and statistical precision of the study, are not part of this definition.[2]
c Pending timelines and implementation of booster campaigns. Once actual timelines are known, milestones will be revised as necessary.
d Will include full follow-up for all AESIs including pregnancy outcomes.

5. AMENDMENTS AND UPDATES

The following amendments have been made to the protocol:

Amendment number	Date	Section of protocol changed	Summary of amendment/update	Reason
2	08 January 2024	5. Amendments and Updates	Changed "spell" to "time" in the reason section of the update to section 9.1.1.3 Matching process	Clarification as requested by the United States Food and Drug Administration
1	18 October 2023	4. Abstract and 6. Milestones	Start of data collection moved to 31 March 2024. Footnote c removed.	Previous start of data collection was prior to protocol endorsement Footnote removed for clarity and consistency.
1	18 October 2023	4. Abstract, 6. Milestones, and 9.7. Data Analysis	Interim report 1 moved to 14 May 2025	Timeline has been extended to account for current progress of the study and to avoid overlap with the production of study 1021 (safety of the original monovalent PFE vaccine) and study 1038 (Natural history of myocarditis) final report. The 3 studies share resources and avoiding overlap will impact positively in the quality of the report.
1	18 October 2023	9.1.1 Matched cohort design	Matching to be done on Pfizer-BioNTech vaccination	For clarification about the matched control cohort
1	18 October 2023	9.1.1.3 Matching process	Rolling cohort design was removed	An alternative approach using time periods for matching will be used
1	18 October 2023	9.1.1.4 Follow-up and 9.7.1.1 Exposure assignment and follow-up	End of study period moved earlier, to 31 August 2024, with follow-up of pregnancy outcomes through 30 June 2025 Censoring of matched pairs due to SARS-CoV-2 infection removed from main analysis	The administration of the bivalent vaccine ended by August 2023, therefore, all vaccinees will have 1 year of follow up by 31 August 2024. To align the main analysis with the ongoing PASS of the Pfizer BioNTech monovalent vaccine, C4591021
1	18 October 2023	9.2.1.2 Self-controlled risk interval design	Inclusion criteria for SCRI study modified	Eligibility criteria of 12 months continuous enrolment in data source was missing

Amendment number	Date	Section of protocol changed	Summary of amendment/update	Reason
1	18 October 2023	9.2.3 Study period	End of study period moved earlier, to 31 August 2024, with follow-up of pregnancy outcomes through 30 June 2025 Figure 2 updated. Last bivalent administered time point moved to August 2023, Follow-up of AESIs end period moved to August 2024.	The administration of the bivalent vaccine ended by August 2023, therefore, all vaccinees will have one year of follow up by 31 August 2024.
1	18 October 2023	9.3.2 Outcomes (Table 1)	Added myositis and hypermenorrhoea. Retained glomerulonephritis and cerebral venous sinus thrombosis. Removed tinnitus, Stevens-Johnson syndrome or toxic epidermal necrolysis. Changed VAED to severe COVID-19 Corrected “Thrombosis” (instead of “Thrombotic”) Changed the risk intervals from “1-X” to “X” and updated footnote (*).	Request from EMA to align the AESIs with those in the ongoing PASS of the Pfizer BioNTech monovalent vaccine, C4591021. Risk intervals amended to show the length of the intervals (instead of the range from day 1 to X), and clarified in the footnote * that the risk intervals start at time 0, corresponding to the day of vaccination.
1	18 October 2023	9.3.3 Covariates	Covariates list modified	Request from EMA to align with ongoing PASS of the Pfizer BioNTech monovalent vaccine, C4591021
1	18 October 2023	9.5 Study size	Table 2 title updated, and table formatting modified	Title updated for clarity. First 2 columns of Table 2 have been reformatted to make the table less busy but the meaning is unchanged.
1	18 October 2023	9.7.1 Cohort design and 9.7.1.11 Sensitivity analysis	Censoring of matched pairs due to SARS-CoV-2-infection removed from main analysis and added as a sensitivity analysis. Study end date updated to 31 August 2024.	To align the main and sensitivity analyses with the ongoing PASS of the Pfizer BioNTech monovalent vaccine, C4591021 The administration of the bivalent vaccine ended by August 2023, therefore, all vaccinees will have one year of follow up by 31 August 2024.

Amendment number	Date	Section of protocol changed	Summary of amendment/update	Reason
1	18 October 2023	9.7.1.7 Estimation of the total and direct effects of vaccination	Section on estimation of total and direct effects of vaccination added	To explain the differences between direct and total effect and justify choice of measuring total effect in main analyses and direct effect in sensitivity analyses
1	18 October 2023	9.7.1.8 Censoring	Section moved below 9.7.1.7 and text expanded	Text added to specify the censoring criteria used to estimate the total effect in the main analyses.
1	18 October 2023	General	Administrative, formatting, and typographical corrections have been made	Updated to provide clarity and to be consistent with remainder of protocol

6. MILESTONES

Milestone	Planned Date[a]
Registration in EU PAS Register	After regulatory endorsement and prior to the start of data collection
Start of data collection[b]	31 March 2024
End of data collection	30 September 2025
Interim report 1[c]	14 May 2025
Final study report[c,d]	30 April 2026

EU PAS Register: European Union Electronic Register of Post-Authorisation Studies.
a Schedule is dependent on protocol endorsement date, uptake of Comirnaty Original/Omicron BA.1 and Comirnaty Original/Omicron BA.4-5, approvals for data extraction and contracts with data access providers (DAPs).
b Start of data collection: the date on which information on the first individual is recorded in the study data set or, in the case of secondary use of data, the date from which data extraction starts [IR Art 37(1)]. Simple counts in a database to support the development of the study protocol, for example, to inform the sample size and statistical precision of the study, are not part of this definition.[2]
c Pending timelines and implementation of booster campaigns. Once actual timelines are known, milestones will be revised as necessary.
d Will include full follow-up for all AESIs including pregnancy outcomes.

7. RATIONALE AND BACKGROUND

In November 2021, World Health Organization (WHO) designated the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variant B.1.1.529 as a variant of concern and named it Omicron.[3] The Omicron variant, and its descendant sub-lineages, eg, BA.4-5, have since spread to become the dominant circulating variant worldwide.[4] The Omicron variant has greatly increased transmissibility and immune-evasion capacity compared with prior circulating variants.[5] In addition, there is evidence that the original mRNA vaccines provide lower protection against the Omicron variant.[6-8]

Pfizer-BioNTech has developed bivalent formulations (Comirnaty Original/Omicron BA.1 (bivalent BA.1) and Comirnaty Original/Omicron BA.4-5 (bivalent BA.4-5)) of the Coronavirus Disease 2019 (COVID-19) mRNA-based vaccine. On 01 September 2022 and 12 September 2022, bivalent BA.1 and bivalent BA.4-5, respectively, were authorised for use in the European Union as a booster following a primary vaccination series, at least 3 months after the last dose of any COVID-19 monovalent vaccine.[9-11] Bivalent BA.1 is authorised for individuals aged 12 years and older, while bivalent BA.4-5 is authorised for individuals aged 5 years and older.[12] On 06 December 2022, the Emergency Task Force of the European Medicines Agency (EMA) concluded that bivalent BA.4-5 may be used as a primary vaccination series in adults and children who have not been previously vaccinated against SARS-CoV-2.[1]

Safety data reported from Vaccine Adverse Event Reporting System (VAERS) and v-safe in the United States up to October 2022 demonstrate a similar safety profile for the bivalent booster as that reported for the monovalent booster.[13] However, to gain a more complete understanding, safety monitoring of the bivalent Pfizer-BioNTech COVID-19 booster vaccines is needed in European countries.

This non-interventional study is designated as a Post-Authorisation Safety Study (PASS) and is a commitment to the EMA. This study complies with the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance ENCePP[14] and Accelerated Development of VAccine benefit-risk Collaboration in Europe (ADVANCE)[15] codes of conduct.

8. RESEARCH QUESTION AND OBJECTIVES

8.1. Research question

The research question is: Is there an increased risk of selected adverse events of special interest (AESIs) after vaccination with bivalent BA.1 or bivalent BA.4-5 compared with no vaccination against COVID-19 among individuals with comparable vaccination history?

8.2. Primary objective

The primary objective is to determine whether there is an increased risk of pre-specified AESIs following the administration of bivalent BA.1 or bivalent BA.4-5 compared with not receiving any COVID-19 vaccine during follow-up.

9. RESEARCH METHODS

9.1. Study design

This study will use a matched cohort design. Additionally, a self-controlled risk interval (SCRI) design will be implemented for specific outcomes (specified in Table 1). Both designs will follow ACCESS specifications for vaccine safety studies.[16]

9.1.1. Matched cohort design

The matched cohort design will be used to estimate the incidence of AESIs after receipt of a bivalent vaccine by individuals in the exposed group, compared with the incidence of the

same AESIs in an unexposed group, i.e., individuals who have not received any COVID-19 vaccine, and individuals who have received COVID-19 vaccine but who did not receive a bivalent COVID-19 vaccine, will be matched on history of COVID-19 vaccination to the exposed group.

9.1.1.1. Causal contrast or estimand

The causal contrast of interest will be the observational analogue of a per-protocol effect, i.e., the effect under complete adherence to the following vaccination strategies:

- Receive a dose of the bivalent BA.1 or bivalent BA.4-5 vaccine. During follow-up, individuals can receive Pfizer vaccines as recommended (see Section 9.1.1.4).
- Do not receive any COVID-19 vaccine during the study follow-up.

9.1.1.2. Time zero or baseline

Time zero (baseline) will be defined as the time at which the exposure status is assigned, when inclusion and exclusion criteria are applied and when study outcomes start to be counted,[17-20] and will be operationalised as follows:

- **Exposed group:** date of the first dose of bivalent BA.1 or bivalent BA.4-5.
- **Unexposed group:** calendar date on when they had not received any dose of the bivalent COVID-19 vaccine, matched to date of the first dose of bivalent BA.1 or bivalent BA.4-5 vaccine in the corresponding match. Time zero for an unexposed individual will be the same date as the corresponding date of the first dose of bivalent BA.1 or bivalent BA.4-5 in the exposed individual in the matched pairs.

9.1.1.3. Matching process

As in prior applications of observational studies of the Pfizer-BioNTech COVID-19 using an unexposed comparator,[21-24] the cohort study will match exposed to unexposed individuals using a 1:1 ratio. Matching will be done with replacement, as follows.

- On the day an exposed individual receives a bivalent vaccine dose, individuals will be matched to unexposed individuals who, on that same day, have exactly the same values for the matching variables, provided that both members of the pair meet the eligibility criteria that day. If more than one unexposed individual is found to match a bivalent vaccinated individual, one will be selected at random.
- Exposed individuals will be allowed to have served as unexposed controls in prior matches, and unvaccinated individuals will be allowed to serve as controls more than once (during different time periods).

Individuals will be matched on the following variables, that have been shown to control for confounding well in prior studies:[21,22]

- Age – 2-year age groups (consecutive years);
- Sex (male, female) – exact matching;
- Place of residence – exact matching, at the level of clinical practice, neighbourhood, or proxy as available (specific for data source);
- History of COVID-19 vaccination, containing the following categories – exact matching:
    - Never vaccinated;
    - Vaccinated with a total of *i* doses (where *i* can range from 1 to the maximum number of plausible doses at the time the matching is performed) with the last dose received in calendar month *j* (where *j* can range from the first month to the last month of the study period);
- Brand of last COVID-19 vaccine received;
- History of hospitalisation for COVID-19;
- Pregnancy status yes, no (exact matching):
    - Among pregnant women, matching will take a 'greedy matching' approach, in which a matched individual will be sought first by last menstrual period (LMP) within 7 days of each other; if no matches are found, the matching period will be extended to 30 days;
- Immunocompromised (yes/no) – exact matching;
- Number of pre-existing conditions considered by the Centers for Disease Control and Prevention (CDC) as risk criteria (0, 1, >1); [22]
- Socioeconomic status/education level (as available, exact matching);

- This selection of matching variables is based on prior observational studies.[22,24,25] The selection of variables will be tailored based on the availability of variables in each data source.

9.1.1.4. Follow-up

Individuals will be followed from time zero until the end of the study period (31 August 2024, or 30 June 2025 for pregnancy outcomes), the occurrence of the AESI being analysed, death, or disenrollment from the data source, whichever occurs first. Additionally, the

matched pair will be censored when the unexposed individual of the pair receives any COVID-19 vaccine or when the exposed individual in the pair receives a non-Pfizer vaccine.

9.1.2. Self-controlled risk interval design

The SCRI design will be used as a secondary analysis to compare the occurrence of specific AESIs for which the design is recommended.[26] Occurrence of an AESI in a specified risk interval following a dose of bivalent BA.1 or bivalent BA.4-5 will be compared with the occurrence of the same AESI in a later interval (control interval), in the same individual. AESIs that are appropriate for this study design are those that are:

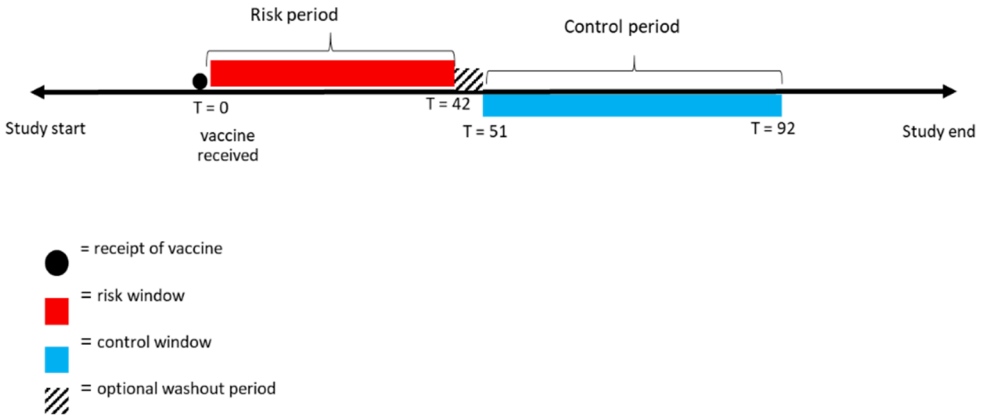
- rare, with a transient risk after exposure;
- have a constant event rate within the observation periods;
- have an occurrence that does not alter the duration of the observation window.[27]

The SCRI design will be used to compare the risk of each AESI during a prespecified period following vaccine administration during which there is a hypothesised increased risk of the outcome ('risk period') with a self-matched control period, used to assess the baseline risk of the outcome.

The risk periods proposed for each AESI and the AESIs that will be analysed using the SCRI design are summarised in Table 1.

A prespecified post-vaccination control period will be used for each outcome. The use of a control period after the risk period (rather than the use of a control period before the risk period) avoids bias from outcomes affecting the probability of exposure (eg, the outcome could be a contraindication for vaccination or could delay vaccination) (Figure 1).

Figure 1. Self-controlled risk interval design



T = time, in days.
Note: Example with a risk period of 42 days and a control period of 42 days.

9.2. Setting

9.2.1. Inclusion criteria

9.2.1.1. Cohort design

- Individuals must meet all the following criteria to be eligible for inclusion in the cohort study:
    - have a minimum of 12 months (or from birth if enrolled in the data source at birth) of continuous enrollment in the data source at time zero;
    - have complete information on the matching variables;
    - be in the population targeted by recommendations for vaccination with the bivalent BA.1 or BA.4-5 COVID-19 vaccines, either as a booster, ie, an additional dose several months or longer after the two or three-dose primary vaccine series, as per national or regional recommendations, or as primovaccination, ie, as the first ever COVID-19 vaccine received:
        - if an individual has not received any COVID-19 primovaccination, they will meet this criterion;
        - if an individual has received primovaccination, they will meet the above criterion if they:
            - meet the age threshold or other criteria for additional COVID-19 vaccine doses after primovaccination;
            - do not have any comorbidities or other characteristics for which a short interval between doses is recommended by local vaccination policies and have had prior COVID-19 vaccination at least 11 months before baseline;
            - have comorbidities or other characteristics for which a short interval between doses is recommended by local vaccination policies and have had prior COVID-19 vaccination at least 5 months before baseline.

9.2.1.2. Self-controlled risk interval design

Individuals must meet all the following criteria to be eligible for inclusion in the SCRI design:

- have received a dose of Pfizer-BioNTech bivalent BA.1 or bivalent BA.4-5 vaccine;
- have not received any COVID-19 vaccine in the three months prior to the dose of bivalent BA.1 or bivalent BA.4-5;

- have experienced the specific AESI during the risk or control interval;
- have full accrual of data used to define the event in both the risk and control interval combined, taking into account the data lag and timing of data extraction;
- have a minimum of 12 months continuous enrollment in the data source at time zero.

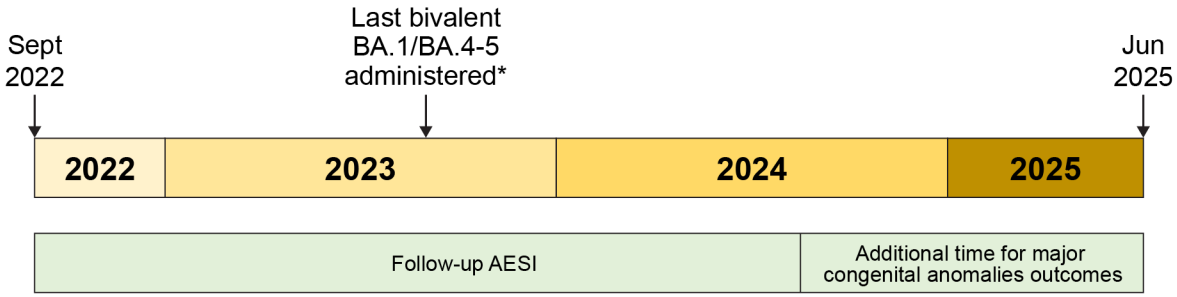
9.2.2. Exclusion criteria

Individuals who have had a diagnosis of the specific AESI being analysed within one year before time zero (to differentiate the recording of previous events from true new events) and at any time before time zero for diabetes type 1 will be excluded from the cohort and SCRI designs.

9.2.3. Study period

For both study designs, the study period will start on the date of launch of the bivalent BA.1 vaccine in each participating country, (ie, September 2022, at the earliest), since this was the first bivalent vaccine to receive authorisation in the EU, and will end on 31 August 2024 or the date of the latest data availability. Additional follow-up for accrual of data for major congenital anomalies will be included up to 30 June 2025 or the date of the latest data availability (see Figure 2).

Figure 2. Study and follow-up periods



**Exact date will depend on date of authorisation of new Pfizer-BioNTech COVID-19 vaccine*

9.3. Variables

9.3.1. Exposure

The exposure categories will be as follows (see Section 9.1.1.2 for definition of time zero):

- **Exposed group:** received a dose of the bivalent BA.1 or bivalent BA.4-5 vaccine. During follow-up, individuals can receive Pfizer vaccines as recommended.
- **Unexposed group:** did not receive any dose of the bivalent BA.1 or bivalent BA.4-5 vaccine at time zero and do not receive any COVID-19 vaccine during the study follow-up.

Exposure status will be assessed from recorded prescriptions, dispensing, or administration data for bivalent BA.1 or bivalent BA.4-5. The type of vaccine received and date of vaccination should 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 secondary data sources. Depending on the data source, vaccines may be identified via nationally used product codes where possible.

If the contributing database cannot differentiate between the monovalent, bivalent BA.1, and bivalent BA.4-5 vaccines, a proxy for vaccine identification will be used (eg, the date of each vaccine introduction in the country, or the date when it is expected that the bivalent BA.1 and bivalent BA.4-5 are mostly used).

Each contributing data source will identify vaccination as follows:

Pedianet (IT) Information on COVID-19 vaccination will include date of administration, type of vaccine, and dose. This information will be collected by the paediatrician at each contact with the individual.

PHARMO (NL): Data on vaccination will be obtained from PHARMO's General Practitioner (GP) database. Information on vaccines includes Anatomical Therapeutic Chemical (ATC) code, brand, and date of administration.

Norwegian health registers, NHR (NO): The national, electronic immunisation register (SYSVAK) records an individual's vaccination status and vaccination coverage in Norway. All vaccinations, including COVID-19 vaccinations are notified to SYSVAK without requiring participant consent. In SYSVAK, the following data are registered: individual personal identifier, vaccine name and ATC code, date of vaccination, reason for vaccination ie, health care professional versus risk-group individual, and the centre where the vaccine was administered.

EpiChron – Aragon data sources (ES): The Aragon Health System (Aragon, Spain) has implemented a specific vaccination register embedded in the electronic health record (EHR) system. COVID-19 vaccination is systematically registered in this register by health care professionals. This register will contain all the relevant information regarding the vaccination process, such as individual's identifier; date of administration and due date for next dose, if applicable; centre of administration; part of the body where vaccine was administered; name of the vaccine; brand (laboratory); dose; and vaccination criterion (risk group to which the person belongs). There is also a free-text section in which health care professionals can include their observations (eg, presence or not of an allergic reaction).

SIDIAP (ES): SIDIAP will have information available on COVID-19 vaccine administration to individuals linked to a unique and anonymous identifier for all 8 million individuals under the Catalan Institute of Health–Primary Care teams. The information will come from the electronic medical records. For each individual, SIDIAP will have date and centre of administration; dose, brand; reasons for vaccination (eg, risk of group); and other information related to vaccination.

Clinical Practice Research Datalink, CPRD (UK): CPRD contains information recorded by National Health Service (NHS) primary care general practitioners. Information on the administration of COVID-19 vaccines will be available. This will include an encrypted unique person identifier; the name of the vaccine; manufacturer; dose; administration route; date of administration; and medical observations, events, referrals, test results, and prescribed medications recorded by the GP prior to, on, or after the vaccination date. In addition, person demographic, practice-level, and staff-level information will also be available.

Standard CPRD-linked data sets can also be obtained including Hospital Episode Statistics (HES) data sets covering hospital secondary care (accident & emergency, admitted patient care, inpatient and outpatient), Office for National Statistics (ONS) data sets for death registry information, mother-baby link, and an algorithm-based pregnancy register.

9.3.2. Outcomes

The list of AESIs to be analysed as outcomes and the preferred study design are based on those proposed by the EMA -sponsored ACCESS project (vACcine COVID-19 monitoring readinESS) and also, for consistency, those included in the ongoing PASS of the Pfizer BioNTech COVID-19 monovalent vaccine (EUPAS41623). The AESIs are listed in Table 1 by body system, with the corresponding risk interval and preferred study design (Section 9.1). Clinical validation will be considered for specific AESIs.

Table 1. Safety outcomes: adverse events of special interest (AESIs) with estimated risk intervals and preferred study design

Body system / classification	Adverse event of special interest	Estimated risk interval (days)*	Preferred study design*
Autoimmune diseases	Guillain-Barré syndrome [a]	42	Cohort/SCRI
	Acute disseminated encephalomyelitis	42	Cohort/SCRI
	Narcolepsy [a]	42	Cohort/SCRI
	Acute aseptic arthritis	42	Cohort/SCRI
	Diabetes mellitus type 1	365	Cohort
	(Idiopathic) thrombocytopenia [a]	42 [a]	Cohort/SCRI
	Thrombosis thrombocytopenia syndrome (TTS) [a]	15 [a]	Cohort/SCRI
	Myositis	365	Cohort
Cardiovascular system	Acute cardiovascular injury including microangiopathy	365	Cohort
	Arrhythmia	365	Cohort
	Heart failure	365	Cohort
	Stress cardiomyopathy	365	Cohort
	Coronary artery disease	365	Cohort
	Myocarditis or pericarditis [a]	7; 14; 21	Cohort/SCRI
	Myocarditis [a]	7; 14; 21	Cohort/SCRI
	Pericarditis [a]	7; 14; 21	Cohort/SCRI
Circulatory system	Coagulation disorders: thromboembolism, haemorrhage	28	Cohort/SCRI

Table 1. Safety outcomes: adverse events of special interest (AESIs) with estimated risk intervals and preferred study design

Body system / classification	Adverse event of special interest	Estimated risk interval (days)*	Preferred study design*
	Single organ cutaneous vasculitis	28	Cohort/SCRI
	Cerebral venous sinus thrombosis	28[28]	Cohort/SCRI
Hepato-gastrointestinal and renal system	Acute liver injury	365	Cohort
	Acute kidney injury	365	Cohort
	Acute pancreatitis	365	Cohort
	Rhabdomyolysis	365	Cohort
	Glomerulonephritis	365	Cohort
Nerves and central nervous system	Generalised convulsion	42	Cohort/SCRI
	Meningoencephalitis	42	Cohort/SCRI
	Transverse myelitis [a]	42	Cohort/SCRI
	Bell's palsy	42	Cohort/SCRI
Respiratory system	Acute respiratory distress syndrome	365	Cohort
Reproductive system	Secondary amenorrhea	183	Cohort
	Hypermenorrhea	183	Cohort
Skin and mucous membrane, bone and joints	Erythema multiforme	42	Cohort
	Chilblain-like lesions	42	Cohort
Other	Anosmia, ageusia	42	Cohort
	Anaphylaxis [a]	1[31]	Cohort/SCRI
	Multisystem inflammatory syndrome	42	Cohort
	Death (any causes)	365	Cohort
	Subacute thyroiditis	365	Cohort
	Sudden death	365	Cohort
Pregnancy outcome, maternal	Gestational diabetes	Any time pregnancy	Sub-cohort
	Preeclampsia	After 20 weeks gestation	Sub-cohort
	Maternal death	Any time pregnancy	Sub-cohort
Pregnancy outcome, neonates. Define design taking trimester into account	Foetal growth restriction	Any time pregnancy	Sub-cohort
	Spontaneous abortions	At spontaneous abortion	Sub-cohort
	Stillbirth	At birth	Sub-cohort
	Preterm birth	At preterm birth	Sub-cohort
	Major congenital anomalies [a]	1 year after birth	Sub-cohort
	Microcephaly	At birth	Sub-cohort
	Neonatal death	At birth	Sub-cohort
	Termination of pregnancy for foetal anomaly	At termination	Sub-cohort

Table 1. Safety outcomes: adverse events of special interest (AESIs) with estimated risk intervals and preferred study design

Body system / classification	Adverse event of special interest	Estimated risk interval (days)*	Preferred study design*
	Small size for gestational age	Any time pregnancy	Sub-cohort
Any	Severe COVID-19[b]	Any	Cohort

SCRI = self-controlled risk interval
* Time zero corresponds to the day of vaccination (i.e., a 42-day risk interval means that individuals are followed from the day of vaccination to day 41). Risk interval based on the ACCESS project results and other publications.
a For these AESI clinical validation will occur.
b Severe COVID-19 defined as either COVID-19 hospitalisation or death

9.3.3. Covariates

The following variables will be assessed at time zero (for the cohort design) or the date of initial vaccine dose (for the SCRI design) to characterise the populations, to define subgroups, or to control for confounding. Potential covariates may include the following information, as available in each data source:

9.3.3.1. Demographics

- Age:
  - Age will be grouped in categories in line with published background incidence rates from ACCESS (0-17, 18-29, 30-39, 40-49, 50-59, 60-65, 66-69, 70-79, 80+ years);
  - The age group 0-17 years will be stratified as follows, where feasible: 0-1, 2-4,5-11, 12-15, 16-17;
- Sex;
- Pregnancy status and pregnancy trimester;
- Geographic region, as appropriate in each data source;
- Socioeconomic status (including housing, employment, and income, if available);
- Date of vaccination (categorised as appropriate, eg, by year or month).

9.3.3.2. COVID-19 history

- Previous diagnosis of COVID-19;
- Positive test result for COVID-19.

9.3.3.3. Personal lifestyle characteristics

- Smoking status;
- Body mass index.

9.3.3.4. Comorbidities

- History of anaphylaxis;
- History of allergies;
- Diabetes mellitus (types 1 and 2);
- Hypertension;
- Cardiovascular disease;
- Cerebrovascular disease;
- Chronic respiratory disease;
- Chronic kidney disease;
- Chronic liver disease;
- Cancer;
- Autoimmune disorders;
- Influenza infection or other respiratory infections;
- Charlson Comorbidity Index (CCI scales will be reported and additionally CCI component morbidities will be reported).
- CDC at risk groups

9.3.3.5. Immunocompromising conditions

- Immunodeficiencies;
- Immunosuppressant medication use;
- Human immunodeficiency virus and other immunosuppressing conditions.

9.3.3.6. Comedication use

- Comedication use during the year before time zero (prescriptions or dispensing, no over-the-counter medication use):
    - Analgesics;
    - Antibiotics;
    - Antiviral medications;
    - Corticosteroids;
    - Non-steroidal anti-inflammatory drugs;
    - Psychotropics;
    - Statins;
    - Novel oral anticoagulants;
    - Warfarin.

9.3.3.7. Healthcare utilisation

- Healthcare utilisation in the year and in the 2 weeks before time zero:
    - Number of hospitalisations;
    - Number of emergency department visits;
    - Primary care utilisation;
    - Cancer screening;
    - Other preventive health services, as appropriate
    - COVID-19 tests.

9.3.3.8. Other vaccinations

- Influenza;
- Pneumococcal;
- DTP (diphtheria, tetanus, and pertussis);
- TPV (polio);

- TV (MMR) (measles, mumps and rubella);
- Hib (Haemophilus influenzae type b);
- HB (hepatitis B virus);
- VV (varicella zoster virus);
- HZ (herpes-zoster virus);
- HPV (human papillomavirus);
- Meningococcal;
- Rotavirus.

9.3.3.9. Surrogates of frailty

The following variables will be used as surrogates of frailty (as available):

- Paralysis;
- Parkinson's disease;
- Skin ulcer;
- Weakness;
- Stroke/brain injury;
- Ambulance transport;
- Dementia;
- Difficulty walking;
- Psychiatric illness;
- Sepsis;
- Heart failure;
- Podiatric care;
- Bladder incontinence;
- Diabetes complications;

- Arthritis;
- Coagulation deficiencies;
- Vertigo;
- Lipid abnormalities.

The CDC at risk groups will be defined based on at-risk medical conditions for developing severe COVID-19 and will be reported as baseline characteristics for vaccinated and non-vaccinated individuals. These will be defined based on scientific evidence available on the US Centers for Disease Control and Prevention website and the UK National Health Services digital website. Those websites are updated regularly and provide a classification based on levels of evidence.

At-risk medical conditions that are considered as at higher risk to develop severe COVID-19 are summarised in Table 2. Medicinal products that can be considered as proxies for these conditions are also listed. At-risk subgroups will be identified using medical codes and associated dates for at-risk medical conditions characterising at-risk groups for developing severe COVID-19 as well as prescription and/or dispensing records for drug exposures which may be used as proxies for their identification. At-risk subgroups will be created for each of the at-risk medical conditions listed in Table 2. Multimorbidity, ie, individuals in more than one at-risk subgroup, will be included in each subgroup.

Table 2. Comorbidities and related medicinal products with evidence of being at high risk for developing severe COVID-19

At-risk medical conditions identified by diagnosis codes	Medicinal product proxy(ies) (ATC code)
Cancer (with chemo/immuno/radiotherapy, cancer treatment, immunosuppressant; targeted cancer treatment (such as protein kinase inhibitors or PARP inhibitors); blood or bone marrow cancer (such as leukaemia, lymphoma, myeloma))	Alkylating agents (L01A) Antimetabolites (L01B) Plant alkaloids and other natural products (L01C) Cytotoxic antibiotics and related substances (L01D) Other antineoplastic agents (L01X) Hormones and related agents (L02A) Hormone antagonists and related agents (L02B) Immunostimulants (L03) Immunosuppressants (L04)
Type 1& 2 Diabetes	Blood glucose lowering drugs A10A & A10B
Obesity (BMI > 30 kg/m²)	Peripherally acting anti-obesity products (A08AB) Centrally acting anti-obesity products (A08AA)
Cardiovascular disease/ Serious heart conditions including heart failure, coronary artery disease, cardiomyopathies	Antiarrhythmics, class I and III (C01B) Cardiac stimulants excl. Cardiac glycosides (C01C) Vasodilators used in cardiac diseases (C01D) Other cardiac preparations (C01E) Antithrombotic agents (B01A)
Chronic lung disease including COPD, asthma, bronchiectasis, interstitial lung disease, cystic fibrosis, tuberculosis.	Drugs for obstructive airway diseases (R03) Lung surfactants (R07AA) Respiratory stimulants (R07AB)
Chronic kidney disease	Erythropoietin (B03XA01)
HIV	Protease inhibitors (J05AE) Combinations to treat HIV (J05AR) NRTI (J05AF) NNRTI (J05AG)
Immunosuppression	Immunosuppressants (L04A) Corticosteroids (H02)
Sickle cell disease	Hydroxyurea (L01XX05) Other haematologic agents (B06AX)
Hypertension	anti-hypertensive drugs (C02, C03, C07, C08, C09)

9.3.4. Data extraction

Each DAP will create ETL specifications using the standard ConcePTION ETL design template (accessible via this link: https://docs.google.com/document/d/1SWi31tnNJL7u5jJLbBHmoZa7AvfcVaqX7jiXgL9uAWg/edit). Following completion of this template and review by study statisticians, each DAP will extract the relevant study data locally using their software (eg, Stata, SAS, R, Oracle). These data will be loaded into the CDM structure in comma-separated values (csv) format. These data will remain local (Figure 3).

9.3.5. Subgroups

Analyses in the matched cohort and in the SCRI study will be conducted in subgroups defined by baseline characteristics as follows:

- Pregnant women;
- Women of childbearing potential, defined as aged between 15 and 50 year old;
- Individuals who are immunocompromised;
- Individuals who have comorbidities;
- Individuals who have autoimmune or inflammatory disorders (eg, rheumatoid arthritis, inflammatory bowel disease);
- Prior influenza vaccination in the last 5 years (yes/no);
- COVID-19 primovaccination-defined subgroups:
    - Individuals who have not started primovaccination;
    - Individuals who have completed primovaccination with Pfizer's monovalent vaccine;
    - Individuals who have completed primovaccination with a Pfizer bivalent vaccine;
    - Individuals who have completed primovaccination with a non-Pfizer vaccine;
- COVID-19 booster-defined subgroups:
    - Individuals who have not received any booster;
    - Individuals whose last booster received was a Pfizer monovalent vaccine;
    - Individuals whose last booster received was a Pfizer bivalent vaccine;
    - Individuals whose last booster received was a non-Pfizer vaccine;
- Type of Pfizer bivalent vaccine received at baseline (bivalent BA.1 or bivalent BA.4-5).

9.4. Data sources

This study will be implemented using health care data source partners in VAC4EU from countries where the Pfizer-BioNTech COVID-19 bivalent vaccine will be available. A preliminary list of these partners, to be confirmed, based on availability, is as follows:

- **Italy**: PEDIANET
    - Pedianet is a paediatric general practice research data source containing health care data for children seen by approximately 140 paediatricians distributed throughout Italy. Pedianet can link to other data sources, including those containing information on routine childhood vaccination and hospitalisation. The family paediatrician's participation in the database is voluntary, and individuals and their parents provide consent for use of their data for research purposes. This database will be the only one not contributing data for the assessment of pregnancy outcomes.
- **The Netherlands**: PHARMO Database Network
    - The PHARMO Database Network, which is maintained by the PHARMO Institute for Drug Outcomes Research, is a population-based network of EHR data sources that combines anonymous data from different primary and secondary health care settings in the Netherlands. Currently, the PHARMO Database Network covers over 7 million active individuals out of a total of 17 million inhabitants in the Netherlands.[32] The General Practitioner Database contains data from 2.0 million active individuals. The records include information on diagnoses and symptoms, laboratory test results, referrals to specialists and health care product/drug prescriptions. In the PHARMO Network data on vaccinations administered by general practitioners and by the public health service are available. PHARMO general practitioner data are representative of the Dutch population in terms of demographic characteristics and diagnoses in primary care.[33]
- **Norway**: Norwegian Health Registries (NHR)
    - The Norwegian Health Registries (NHR), accessed through a partnership with the University of Oslo, cover several national health registers, ie, the Medical Birth Registry of Norway (MBRN), the National Patient Register (NPR), Norway Control and Payment of Health Reimbursement (KUHR), the Norwegian Immunisation Registry (SYSVAK), the Norwegian Prescription Database (NorPD), and Statistics Norway (SSB). These data sources cover virtually the entire population of Norway, with 5.3 million active individuals.
- **Spain:** EpiChron and SIDIAP
    - ***EpiChron:***The EpiChron Cohort Study links socio-demographic and clinical anonymised information for all the inhabitants of Aragon region, built from the BIGAN platform. The Aragon BIGAN platform integrates a technical infrastructure and a data lake

gathering individual data from the regional health service information systems, including primary care, specialised care, hospitalisations, ER episodes, drug prescriptions, image diagnosis, laboratory tests, diagnostics, vaccination, medical history and demographics from the users of the public health system of Aragon, about 2 million individuals with historic data, and an active population of 1.3 million individuals.

- ***SIDIAP:*** The Information System for the Improvement of Research in Primary Care (Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària' [SIDIAP]) SIDIAP includes data from 328 primary care centres managed by the Catalan Health Institute in Catalonia, Spain. The database contains pseudo-anonymised records for >8 million people since 2006, with 5.8 million people active in June 2021 (75% of the Catalan population). SIDIAP is representative of the general population living in Catalonia in terms of age, sex and geographic distribution [34]. SIDIAP includes data on the clinical and referral events registered by primary health care professionals and administrative staff in EHRs, demographic information, community pharmacy invoicing data, specialist referrals, and primary care laboratory test results. SIDIAP data can be linked to other data sources and registers at the local and national level.

- **UK**: Clinical Practice Research Datalink (CPRD)

  - CPRD includes data from EHRs of general practitioners in the UK who act as gatekeepers for health care and maintains individuals' life-long EHRs. Secondary care teams also provide information to general practitioners about their patients, including key diagnoses and procedures. The data recorded in the CPRD include demographic information, prescription details, clinical events, preventive care, specialist referrals, hospital admissions, and major outcomes, including death. Currently, CPRD Aurum contains data for approximately 13.3 million individuals [Current acceptable individuals (ie, registered at practices currently contributing data, excluding transferred out and deceased individuals)].

9.5. Study size

The study size will be determined by the uptake of the bivalent BA.1 and bivalent BA.4-5 vaccines in the contributing data sources during the study period. The 95% confidence intervals [35] for different potential values of the true risk ratios of the AESIs are shown in Table 3, under different scenarios of risks in the unvaccinated population and of sample sizes, assuming complete follow-up.

Table 3. Confidence interval limits for AESI risk ratios for different scenarios of true risks in the unvaccinated group and various sample sizes

Number of individuals per group (sample size)	AESI risk in unvaccinated group	Risk ratio	Lower bound of 95% CI	Upper bound of 95% CI
10,000	1 per 100,000	2	<0.01	3960.57
		5	0.01	4442.71

Table 3. Confidence interval limits for AESI risk ratios for different scenarios of true risks in the unvaccinated group and various sample sizes

Number of individuals per group (sample size)	AESI risk in unvaccinated group	Risk ratio	Lower bound of 95% CI	Upper bound of 95% CI
		10	0.02	6654.91
	1 per 50,000	2	0.01	428.70
		5	0.04	608.11
		10	0.10	991.39
	1 per 5000	2	0.37	10.92
		5	1.10	22.81
		10	2.34	42.77
100,000	1 per 100,000	2	0.18	22.06
		5	0.58	42.80
		10	1.28	78.12
	1 per 50,000	2	0.37	10.92
		5	1.10	22.82
		10	2.34	42.78
	1 per 5000	2	1.17	3.42
		5	3.09	8.08
		10	6.32	15.83
1,000,000	1 per 100,000	2	0.94	4.27
		5	2.54	9.86
		10	5.22	19.16
	1 per 50,000	2	1.17	3.42
		5	3.09	8.08
		10	6.32	15.84
	1 per 5000	2	1.69	2.37
		5	4.30	5.82
		10	8.65	11.56

AESI: adverse event of special interest. CI: confidence interval.

9.6. Data management

This study will be conducted in a distributed manner using a common protocol, common data model (CDM), and common analytics programs using existing health care data. The following steps will be implemented:

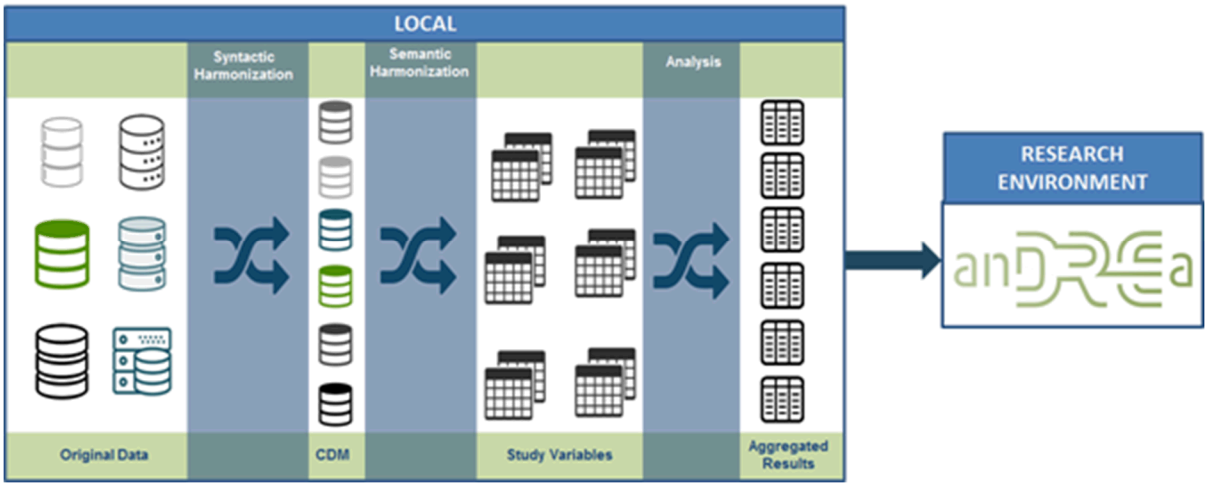
Step 1: Extraction, transformation, and loading (ETL) of data to the CDM. A shared syntactic foundation will be used to harmonise the structure of the data sets stored and maintained by each data access provider (DAP). The CDM developed during the IMI-ConcePTION project will be used.[36] In this CDM, data are represented in a common structure, but the content of the data remain in their original format. The ETL design for each study will be shared in a searchable findable, accessible, interoperable, and re-usable (FAIR) catalogue. FAIR is defined as findable, accessible, interoperable, and re-usable. The VAC4EU FAIR data catalogue is a meta-data management tool designed to contain searchable meta-data describing organisations that can provide access to specific data sources. Data quality checks

will be conducted to measure the integrity of the ETL as well as the internal consistency within the context of the CDM (see Section 9.8).

Step 2: **Reconcile differences between terminologies**. A shared semantic foundation will be built for the definition of the events to be analysed by collecting relevant concepts in a structured fashion using a standardised event definition template to reconcile differences between terminologies. The CodeMapper tool was used to create diagnosis code lists based on completed event definition templates for each AESI and comorbid risk condition in the ACCESS project.[6] Based on the relevant diagnostic medical codes and keywords, as well as other relevant concepts (eg, medications), one or more algorithms will be constructed (typically one sensitive, or broad, algorithm and one specific, or narrow, algorithm) to operationalise the identification and measurement of each event. These algorithms may differ between data sources, as the components involved in the study variables may differ. Manual review of electronic records will be conducted for a sample of the events. Specifications for both ETL and semantic harmonisation will be shared in the catalogue.

Step 3: **Conversion to harmonised study variable sets**. R and SAS programs for the calculation of incidence and prevalence rates will be distributed to DAPs for local deployment after conversion to harmonised study variable sets. The aggregated results produced by these scripts will then be uploaded to the Digital Research Environment (DRE) for pooled analysis and visualisation (Figure 3) The DRE will be made available through UMCU (University Medical Center Utrecht)/VAC4EU (https://www.andrea-cloud.eu/). 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-cloud.eu/azure-dre/).

Figure 3. Data management plan



CDM = common data model.

9.6.1. Case report forms (CRFs)/Data collection tools (DCTs)/Electronic data record

This study will use secondary data collected from EHR databases. Specific forms will be developed for the validation of selected study endpoints and securely saved in environments assuring data protection and participant confidentiality according to the requirements of each country and DAP.

As used in this protocol, the term CRF should be understood to refer to either a paper form or an electronic data record or both, depending on the data collection method used in this study.

A CRF is required and should be completed for each included participant. The completed original CRFs are the sole property of the DAPs and should not be made available in any form to third parties, except for authorised representatives of Pfizer or appropriate regulatory authorities. The DAPs shall ensure that the CRFs are securely stored at the study site in encrypted electronic form and will be password protected to prevent access by unauthorised third parties.

The DAPs have ultimate responsibility for the collection and reporting of all clinical, safety and laboratory data entered on the CRFs for the procedure of event/case verification/validation and ensuring that they are accurate, authentic/original, attributable, complete, consistent, legible, timely (contemporaneous), enduring, and available when required. The CRFs must be signed by an authorised staff member at the DAP to attest that the data contained on the CRFs are true. Any corrections to entries made in the CRFs or source documents must be dated, initialed, and explained (if necessary) and should not obscure the original entry.

The source documents are the hospital or physician's charts. In these cases, data collected on the CRFs must match those charts. Each DAP will follow its own specific validation procedures, aligned to the general VAC4EU validation plan.

9.6.2. Record retention

The final study aggregated results sets and statistical programs will be archived and stored on the DRE and the VAC4EU SharePoint site. Validation of the quality control (QC) of the statistical analyses will be documented. The final study protocol and any amendments, the final statistical analysis plan, statistical programs, and output files will be archived on a specific and secured central drive. Study records or documents may also include the analyses files, syntaxes (usually stored at the data source site), ETL specifications, and output from data quality checks.

To enable evaluations and/or inspections/audits from regulatory authorities or Pfizer, DAPs agree to keep all study-related records, including the identity of all participating individuals (sufficient information to link records, eg, CRFs and hospital records), copies of all CRFs, safety reporting forms, source documents, detailed records of treatment disposition, and adequate documentation of relevant correspondence (eg, letters, meeting minutes, and telephone call reports). The records should be retained by DAPs according to local

regulations or as specified in the vendor contract, whichever is longer. DAPs must ensure that the records continue to be stored securely for so long as they are retained.

If UMCU, for any reason becomes unable for any reason to continue to retain study records for the required period, Pfizer should be prospectively notified. The study records must be transferred to a designee acceptable to Pfizer, such as another DAP, another institution, or to an independent third party arranged by Pfizer.

Study records must be kept for a minimum of 15 years after completion or discontinuation of the study, unless UMCU and Penta and Pfizer have expressly agreed to a different period of retention via a separate written agreement. Records must be retained for longer than 15 years or as required by applicable local regulations.

UMCU must obtain Pfizer's written permission before disposing of any records, even if retention requirements have been met.

9.6.3. Data processing and transformation

The R and SAS scripts will first transform the data in the syntactically harmonised CDM to semantically harmonised study variables (Figure 2). Following creation of study variables, the data will be characterised. This characterisation will include calculation of code counts and incidence rates, as well as benchmarking within the data source (over time), between data sources and externally (against published estimates). Subsequently, R and SAS scripts to conduct analysis against semantically harmonised study variables will be distributed and run locally to produce aggregated results. The R and SAS scripts for these processing and analysis steps will be developed and tested centrally at UMCU and sent to the DAPs.

The R and SAS scripts will be structured in a modular format to ensure transparency. Functions to be used in the modules will be either standard R and SAS packages or packages specifically designed, developed, and tested for multi-database studies. Scripts will be double coded in SAS and R and quality checks will be thoroughly documented.

The DAPs will run the R and SAS scripts locally and send aggregated analysis results to the DRE using a secure file transfer protocol. In the DRE, results will be further plotted, inspected (for quality assessment), and pooled (if needed) for final reporting.

All final statistical computations will be performed on the DRE using R or SAS. DAPs will have access to the workspace for script verification.

Aggregated results, ETL specifications, and a repository of study scripts will be stored in the DRE.

9.6.4. Data access

Within the DRE, each project-specific area will consist of a separate secure folder called a workspace. Each workspace will be completely secure, and researchers will be in full control

of their data. Each workspace will have their own list of users, which will be managed by its administrators.

The DRE architecture will allow 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 will be granted access to study-specific secure workspaces. Access to these workspaces will only be possible via a double authentication process, using an identification code and password together with the user's mobile phone for authentication.

All researchers with access to the workspace within the DRE will be able to upload files. Downloading of files will only be possible after requesting and receiving permission from a workspace member with an 'owner' role.

9.7. Data analysis

Detailed methodology for summary and statistical analyses of data collected in this study will be documented in a statistical analysis plan (SAP), which will be dated, filed, and maintained by the sponsor. The SAP may modify the plans outlined in the protocol; any major modifications of primary endpoint definitions or their analyses will be detailed in a protocol amendment.

Data will be extracted and descriptive and comparative analyses (when informative) for all AESIs, except major congenital anomalies, from study start (ie, September 2022, at the earliest) up to the latest data available in each data source will be reported in the interim report (14 May 2025). For the final report (30 April 2026), additional data extraction and analyses of all outcomes, including major congenital anomalies, will be done up to 30 September 2025 or the latest data available in each data source.

9.7.1. Cohort design

9.7.1.1. Exposure assignment and follow-up

The main exposure of interest is in the receipt of at least one dose of the bivalent BA.1 or bivalent BA.4-5 vaccine. Individuals will be assigned to each vaccination category according to their data at time zero, as outlined below:

- **Exposed cohort:** Individuals will be assigned to this cohort when they receive a first dose of the bivalent BA.1 or bivalent BA.4-5 vaccine. Individuals will be censored if they receive a non–Pfizer COVID-19 vaccine during follow-up. For the matched comparative analyses individuals will be censored when their matched pair is censored.

- **Unexposed cohort:** Individuals will be assigned to this cohort if they have not received a COVID-19 vaccine at time zero. Individuals will be censored if and when

they receive a dose of any COVID-19 vaccine during follow-up. Time zero in the unexposed cohort will be the same day when they did not receive a Pfizer-BioNTech COVID-19 vaccine dose. This day will be chosen by calendar matching to the time zero of the corresponding exposed group.

Individuals will be followed from time zero (see Section 9.1.1.4) until the censoring described above, death, administrative end of follow-up, or end of study period 31 August 2024, whichever occurs first, with pregnancies followed until 30 June 2025 (see Section 9.7.1.7). For analyses of AESIs with known risk intervals, follow-up will be truncated at the end of the risk interval.

9.7.1.2. Descriptive statistics

The distributions of baseline characteristics at time zero by exposure group will be calculated to describe the study cohort and illustrate differences between the groups. Means, standard deviations, medians, and other quartiles will be estimated for continuous variables. Counts and proportions will be estimated for categorical variables. The missingness of variables will also be described. The interim report will be limited to descriptive analyses. 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.[37,38] An overall standardised difference across all levels will be calculated for multilevel categorical variables.[38] The balance will also be checked after propensity score methods are applied to control for any confounding. Further details will be described in the SAP.

9.7.1.3. Description of vaccination trajectories

Each exposure group of the matched cohort will be characterized in the following categories:

- Counts and percentages of individuals receiving a subsequent COVID-19 vaccine;
- Brand of subsequent COVID-19 vaccine;
- Time to administration of subsequent vaccine.

9.7.1.4. Crude outcome measures

For safety outcomes with a known short risk interval (Table 1) (e.g., anaphylaxis), the risk (number of events/number of individuals in a cohort) and the corresponding 95% CIs will be computed. Effect estimates will be calculated both as risk differences and as risk ratios, with their corresponding 95% CIs, for those exposed to a Pfizer-BioNTech COVID-19 bivalent vaccine compared with the unexposed group that did not receive any COVID-19 vaccine during follow-up.

For safety outcomes with unknown risk intervals or those that require long follow-up (eg, death), the cumulative incidence will be computed, which will be estimated using the 1 − Kaplan-Meier estimator. Time to outcome will be defined as the time from the baseline date (time zero) until the occurrence of the outcome or censoring (Section 9.1.1.4 and Section 9.7.1.7). The variance will be computed using approaches that account for autocorrelation (eg, the robust estimator or via bootstrapping).[39,40] Risk differences and risk ratios (and their corresponding 95% CIs) will be estimated at different time intervals, which will be adapted to the specific nature of each outcome.

9.7.1.5. Subgroup analyses

If the sample size allows for informative analyses, subgroups as defined in Section 9.3.4, will be analysed.

9.7.1.6. Adjustment for baseline imbalances

Propensity score methods will be used to estimate the adjusted risks, cumulative incidences, effect estimates and their corresponding 95% CIs to account for potential baseline confounding. Specifically, the propensity score (ie, the probability of receiving bivalent BA.1 or bivalent BA.4-5 vaccine conditional on baseline covariates listed in Section 9.3.3) will be used to construct inverse probability weights.[40,41] More details, including the variable selection and construction of weights, will be provided in the SAP. The incidence of COVID-19-related outcomes in the first few days after baseline will be evaluated as a negative control outcome for baseline exchangeability.

9.7.1.7. Estimation of the total and direct effects of vaccination

The primary objective (Section 8.2) of this study uses a simplified definition of exposure as vaccination with Pfizer-BioNTech bivalent BA.1 or bivalent BA.4-5 vaccine vs. no vaccination with a bivalent vaccine. However, there are two different effects on the incidence of AESIs compared with no vaccination that can be evaluated, i.e., the direct effect and the total effect.

The total effect includes both the direct effect that is mediated by the vaccine's preventive effect on SARS-CoV-2 infection plus its effect on the AESIs. If the total effect is assessed, it is probably that vaccine protection against acute respiratory distress syndrome will be observed since the vaccine protects against infection with SARS-CoV-2, which itself causes acute respiratory distress syndrome. If acute respiratory distress syndrome is, theoretically, an AESI, the effect of the vaccine on this event could be masked by the protective effect of the vaccine against SARS-CoV-2 infection and therefore COVID-19.

Many of the observed or hypothetical post-vaccination AESIs may also be associated with COVID-19 infection. Previous studies that evaluated associations between various aspects of COVID-19 infection (e.g., COVID-19 hospitalisation, diagnosis, or positive test) and selected AESIs reported increased risks of acute kidney injury and arrhythmia,[22] deep vein thrombosis/pulmonary embolism,[22,42] intracranial hemorrhage,[22] myocardial infarction,[22,42-44] ischemic stroke,[42-44] and myocarditis/pericarditis[22] in various populations.

In addition, the incidence of the events (AESIs) mediated by SARS-CoV-2 infection and COVID-19 will depend on the background rates of SARS-CoV-2 infection and COVID-19 in the study population. Thus, this will make the estimate of the total effect highly dependent on the level of pandemic activity during the follow-up period.

Individuals would need to be censored at the time of a documented SARS-CoV-2 infection or COVID-19 diagnosis in order to estimate the direct effect only. The likely underreporting of SARS-CoV-2 infections and COVID-19 during the time when bivalent vaccination is available will mean that some of the estimated direct effects could be mediated through the prevention of SARS-CoV-2 infection or COVID-19 by the vaccination. It is expected that this will be important limitations to evaluating the direct effect. Therefore, the main analysis of this study will evaluate the total effect. In addition, sensitivity analyses (Section 9.7.1.11) will be carried out to estimate the direct effect. An additional censoring criterium, i.e., a documented SARS-CoV-2 infection or COVID-19, will be implemented to estimate the direct effect. The assumption behind this is that uncensored individuals will have a similar risk for the AESI as that of the censored individuals if they had not been infected, which is conditional on the matched covariates and other variables that will be adjusted for.

9.7.1.8. Censoring

Censoring to analyse individuals with complete follow-up and complete adherence to the vaccination strategies will be done to estimate the total effect of bivalent BA.1 or BA.4-5 vaccination on the risk for the AESIs.

The following censoring will be applied to estimate the total effect under complete follow-up and complete adherence to the vaccination strategies, as well as the corresponding assumptions for validity:

- **Censoring after loss to follow-up** (i.e., disenrollment from the data source). This censoring will be implemented to estimate the effect under complete follow-up. The assumption behind this is that individuals who do not disenroll from the data source have a similar risk of the AESI as the individuals who do disenroll from the data source if they had remained in follow-up, conditional on the matched covariates and other variables that were adjusted for.

- **Censoring due to unexposed individual receiving COVID-19 vaccination**. Censoring of the matched pair when the individual in the unexposed cohort receives a COVID-19 vaccine or when the individual in the exposed cohort receives a non-Pfizer vaccine will occur. This censoring will be implemented to estimate the effect under complete adherence to the vaccination strategies. The assumption behind this is that the risk of the AESI for individuals that did not deviate from the vaccination strategy would be similar to that of the individuals who deviated from the vaccination strategy, had they not deviated, conditional on adjustment of the matched variables and other variables.

9.7.1.9. Meta-analysis

A combined effect will be estimated using estimates from each data source and appropriate random-effects meta-analytical methods. The heterogeneity across data sources will be analysed, and forest plots will be produced containing the individual risk estimates plus the overall risk estimate.

9.7.1.10. Missing data handling

Several approaches for handling missing data will be considered (eg, inverse probability weighting of the complete case population, complete case analysis), based on the amount of missing data and the most reasonable assumption on the pattern of how the data are missing. Additional details on when and which method will be used will be described in the SAP.

9.7.1.11. Sensitivity analysis

One sensitivity analysis will estimate the direct effect of vaccination on the AESIs by repeating the main analyses and additionally censoring in the event of a documented SARS-CoV-2 infection. Another sensitivity analysis will exclude individuals who have had contact with the health care system in the 7 days before time zero. Contact with the health care system serves as an indicator of a health event not related to subsequent vaccination that could reduce the probability of receiving the vaccine.[22]

If the homogeneity of the data allows, additional sensitivity analyses will be done to estimate the combined pooled effect across data sources.

9.7.2. Self-controlled risk interval design

9.7.2.1. Descriptive statistics

The distributions of individuals' characteristics at the time of vaccination with bivalent BA.1 or bivalent BA.4-5 will be calculated to characterise the study sample. Means, standard deviations, medians, and other quartiles will be estimated for continuous variables. Counts and proportions will be estimated for categorical variables. The missingness of variables will also be described.

9.7.2.2. Measures of association

Conditional Poisson regression will be used to estimate incidence rate ratios and their corresponding 95% CIs. AESI for which the SCRI design will be a secondary approach for AESIs, with their risk windows, are specified in Table 1. The control period that will be the same length will follow the risk interval; more details will be provided in the SAP.

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 may be included as covariates in regression models.

Analyses will be conducted in subgroups defined by demographic and clinical characteristics, dose number, and other covariates of interest when feasible.

9.7.2.3. Meta-analysis

As in the cohort design, a combined effect will be estimated using estimates from each data source using appropriate random-effects meta-analytic methods. The heterogeneity across data sources will be analysed, and forest plots will be produced to display the individual and pooled estimates.

9.8. Quality control

Data transformation into the CDM will be conducted by each subcontracted research partner in its associated data source, using the processes described below. 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 statistical analysis plans; and
- requirements for scientific review by senior staff.

9.8.1. University Medical Center Utrecht and RTI Health Solutions

At UMCU, the scientific coordinating centre responsible for central data management and analysis, all documents will 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 RTI Health Solutions (RTI-HS), the project coordinating centre and scientific coleader centre, all key study documents will undergo QC review, senior scientific review, and editorial review. Senior reviewers with expertise in the appropriate subject matter areas 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.

Rigorous QC will be performed for all deliverables.

9.8.2. Pedianet (IT)

Pedianet data processing will include QC steps that will verify the correspondence between diagnostic codes and their open-text descriptor through manual validation of clinical histories, in addition to standardised procedures in Structured Query Language (SQL) and Microsoft Access to extract data from the database. Quality control checks of individual general data will be conducted through the detection of outlier values and validation rules; grouping of diseases; and regular monitoring of aggregate clinical and drug data. All

transformations in the data will be logged in R scripts. To ensure code reliability, double programming in R and Stata or Python will be in place for all scripts.

9.8.3. PHARMO (NL)

PHARMO adheres to high standards throughout the research process based on robust methodologies, transparency, and scientific independence. PHARMO conducts studies in compliance with the ENCePP *Guide on Methodological Standards in Pharmacoepidemiology*[45] and the *ENCePP Code of Conduct* [14]. PHARMO is ISO 9001:2015 certified. Standard operating procedures, work instructions, and checklists will be used to guide the conduct of the 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.

9.8.4. NHR (NO)

The University of Oslo will have centralised information security policies in place to ensure the confidentiality, integrity and availability of the organisation's systems and data. All data will be stored and analysed within the "Tjeneste for Sensitive Data" (TSD) platform, a service for sensitive data managment at the University of Oslo.[46,47] Only authorised researchers will have access to manipulate the data within the TSD, via a two-step authentication process. The study will be conducted in compliance with the *Guidelines for Good Pharmacoepidemiology Practices (GPP)*[48] and the *ENCePP Code of Conduct.*[14] Data quality is a high priority at the Norwegian Health Registries; updated data are released regularly for research purposes after centralised quality control. The University of Oslo will use their rules for secure and confidential data storage and analysis, as well as their rules for data cleaning, linkage, and programming.

9.8.5. EpiChron (ES)

The EpiChron cohort will be built from the BIGAN platform which integrates a technical infrastructure and a data lake, collecting individual data from the regional health service information systems. The BIGAN platform includes several mechanisms to control and improve the quality of data, mainly in the ETL processes for capture and persistence in the data lake. These mechanisms include validation rules (for example., for dates and time intervals) and cross-checks with master tables, requiring that certain coded data exist in a standardised dictionary. Analyses of the distribution of variables will also be carried out periodically, to detect 'outliers' that identify errors in the data capture or transformation processes. Generally, records that do not pass the quality assurance procedures are kept in a 'holding' area for review and decision to discard or reprocess. The resulting databases will be pseudonymised to encrypt individual-level identification codes, protecting individuals' privacy and complying with data protection laws. They will be stored on a central computer server, with access restricted to the members of the research group, via a two-step authentication process. The research group will comprise a multidisciplinary qualified team

including public health specialists, epidemiologists, clinicians, pharmacists, statisticians, and data managers, who are all trained in data management and data protection.

9.8.6. SIDIAP (ES)

QC processes will be implemented at each phase of the data flow cycle. QC checks will be performed at the extraction and uploading steps. To assess data completeness, the presence of elements will be described by geographical area, registering physician, time and the distribution of values. The accuracy of the data will be assessed by validity checks on outliers, out of range values, formatting errors and logical date incompatibilities. Completeness and accuracy measures will be used to inform decisions on the required transformations to improve data quality (eg, harmonisation, normalisation, and clean-up) and the fitness for purpose of the data for use in this study.

9.8.7. CPRD (UK)

The Drug Safety Research Unit (DSRU) will have information security policies in place to preserve the confidentiality, integrity and availability of the organisation's systems and data. These will include ensuring that the premises provide suitable physical and environmental security, all equipment will be secure and protected against malicious software, the network will be accessed only by authorised staff, telecommunication lines to the premises will be protected from interception by being routed overhead or underground, and personnel will receive training on security awareness. The study will be conducted in compliance with the International Society for Pharmacoepidemiology *Guidelines for Good Pharmacoepidemiology Practices (GPP)*[48] and the *ENCePP Code of Conduct.* [14] Data quality is a high priority at the DSRU and will be assured through a number of methods based on staff training, validated systems, error prevention, data monitoring, data cleaning, and documentation, including the following:

- staff training on data processing standard operating procedures;
- data management plan for every research study outlining the legal basis for data collection, data flows, data access rights, data retention periods, etc.;
- routine data cleaning to screen for errors, missing values, and extreme values and diagnose their cause;
- system process logs to document staff access, etc.

9.9. Limitations of the research methods

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

A data-related limitation of this study is the reliance on the accuracy of codes and algorithms to identify outcomes. Some outcomes and their dates of occurrence will be validated, but the extent of validation may be limited because of only partial access to medical records. It is

possible that some AESIs may be due to immunisation errors occurring during the administration of the bivalent BA.1 or bivalent BA.4-5 vaccines. This information is not regularly collected and it will not be possible to take these events into account with the current protocol. Exposure identification may be based on pharmacy dispensing records, general practice records, immunisation registers, medical records, or other secondary data sources. How specific COVID-19 vaccine products and dates of vaccination in these data sources will be identified is described in Section 9.3.1. It is possible that vaccination of individuals outside the health care system will not be recorded in secondary EHR databases, thereby leading to potential bias because of exposure misclassification. Additionally, while every effort will be made to capture data on all variables, some data may not be identified or recorded in the data sources.

A study design-related limitation of both the cohort and SCRI designs is that any uncertainty regarding risk periods will lead to misclassification and underestimation of the risk. The estimation of cumulative incidence curves with a time zero starting on the date of the exposure assignment and eligibility assessment will minimise the impact of unknown risk periods, because there is an unvaccinated comparator and evaluation is done over the whole available follow-up.

A limitation of the cohort design is the potential for residual or unmeasured confounding, as it is unlikely that the data sources will have information on all potential confounders. To address potential confounding, the SCRI, which automatically adjusts for time-invariant confounders, will be used as a secondary approach. However, the SCRI is not well-suited to study outcomes with gradual onset, long latency, or risk periods that are not well known. It also may be subject to bias for outcomes that affect the probability of exposure. The SCRI design will be used as a secondary approach to the cohort design for prespecified AESI with defined risk intervals. Of note, several safety and efficacy studies comparing vaccination vs. no vaccination using a matched cohort design with matching variables similar to those that will be used in this study have been reported, and good adjustment for confounding was obtained.[21-25]

The matching procedure in the cohort analysis produces a study population (ie, a set of matched pairs) with a distribution of matching variables representative of the exposed by matching unexposed individuals to exposed individuals based on a prespecified set of baseline variables. Therefore, the cohort analysis estimates the average causal effect in the vaccinated (ie, in a population that has the distribution of matching variables of the vaccinated). If further adjustment via inverse probability weighting is applied, because the weights are estimated and applied to the matched population, the estimated effect will still be the causal effect in a population that has the distribution of matching variables of the vaccinated. The average causal effect in the vaccinated and the average causal effect in the whole population (vaccinated and unvaccinated) should differ only (apart from random variation) if effect modification by any baseline variable exists. This will have to be considered when comparing effect estimates with those from other studies.

Finally, the extrapolation of the study results to other European countries that not participated in the current analyses and non-European countries should be done cautiously if relevant differences in the populations exist.

9.10. Other aspects

Not applicable.

10. PROTECTION OF HUMAN SUBJECTS

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

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

10.1. Participant information

This study will involve mainly data that will be in an anonymised structured format that will contain no participant personal information. In compliance with DAPs policies[1], in all cells with less than 5 counts for AESIs the actual number will be replaced by <5 to avoid potential identification of individuals.

All parties will comply with all applicable laws, including laws regarding the implementation of organisational and technical measures to ensure protection of participants' personal data. Such measures will include omitting participants' names or other directly identifiable data in any reports, publications, or other disclosures, except where required by applicable laws.

Participant personal data will be stored by the DAPs in encrypted electronic form and will be password protected to ensure that only authorised study staff can access them.

DAPs will implement appropriate technical and organisational measures to ensure that personal data can be recovered in the event of disaster. In the event of a potential personal data breach, DAPs shall be responsible for determining whether a personal data breach has in fact occurred and, if so, providing breach notifications as required by law.

To protect the rights and freedom of natural persons in terms of the processing of personal data, when study data are compiled for transfer to Pfizer and other authorised parties, any participant names will be removed and will be replaced by a single, specific, numerical code. All other identifiable data transferred to Pfizer or other authorised parties will be identified by this single, patient-specific code. In the case of data transfer, Pfizer will maintain high standards of confidentiality and protection of individuals' personal data consistent with the vendor contract and applicable privacy laws.

[1] The data sources concerned by this rule are : PHARMO, NHR, EpiChron, SIDIAP and CPRD.

10.2. Participant consent

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

10.3. Institutional review board (IRB)/Independent ethics committee (IEC)

There will be prospective approval of the study protocol, protocol amendments, and other relevant documents (e.g., informed consent forms if applicable) from the relevant IRBs/IECs, where applicable. All correspondence with the IRB/IEC must be retained. Copies of IRB/IEC approvals must be forwarded to Pfizer.

10.4. Ethical conduct of the study

The study will be conducted in accordance with legal and regulatory requirements, as well as with scientific purpose, value, and rigour, and will follow generally accepted research practices described in the *Guidelines for Good Pharmacoepidemiology Practices (GPP)*[49] and has been designed in line with the ENCePP *Guide on Methodological Standards in Pharmacoepidemiology*.[45]

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 E2E*[50] and provided in the EMA *Guideline on Good Pharmacovigilance Practices (GVP) Module VIII: Post-Authorisation Safety Studies (Rev 3),*[2] and with the 2012 EU pharmacovigilance legislation, adopted 19 June 2012.[51] The *ENCePP Checklist for Study Protocols*[52] has been completed (see ANNEX 2).

The study will be registered in the EU PAS Register prior to the start of data collection.[53]

The research team and study sponsor will adhere to the general principles of transparency and independence in the *ENCePP Code of Conduct*[14] and the ADVANCE Code of Conduct.[15]

10.5. Independent scientific advisory committee

An **independent scientific advisory committee** (SAB) will be appointed, comprising independent experts in vaccine safety studies external to the study. This committee will have a consultative role and will have the expertise and knowledge necessary to advise the study principal investigator on scientific and technical matters of the research questions of the study. These experts will act in their personal capacity. The SAB will meet for the review of the interim and final reports upon request from the study PI.

11. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

11.1. Structured data analysis

This study involves a combination of existing structured data and unstructured data, which will be converted to structured form during the implementation of the protocol solely by a computer using automated/algorithmic methods, such as natural language processing.

In these data sources, it is not possible to link (ie, identify a potential association between) a particular product and medical event for any individual. Thus, the minimum criteria for reporting an adverse event (AE) (ie, identifiable person, identifiable reporter, a suspect product, and event) cannot be met.

11.2. Human review of unstructured data

This study protocol requires human review of individual-level unstructured data; unstructured data refer to verbatim medical data, including text-based descriptions and visual depictions of medical information, such as medical records, images of physician notes, neurological scans, x-rays, or narrative fields in a database. The reviewer is obligated to report AEs with explicit attribution to any Pfizer drug that appear in the reviewed information (defined per the participant population and study period specified in the protocol). Explicit attribution is not inferred by a temporal relationship between drug administration and an AE but must be based on a definite statement of causality by a health care provider linking drug administration to the AE.

The requirements for reporting safety events on the non-interventional study (NIS) adverse event monitoring (AEM) Report Form to Pfizer Safety are as follows:

- All serious and non-serious AEs with explicit attribution to **any Pfizer drug** that appear in the reviewed information must be recorded on the chart abstraction form and reported, within 24 hours of awareness, to Pfizer Safety using the NIS AEM Report Form.

- Scenarios involving drug exposure, including exposure during pregnancy, exposure during breast feeding, medication error, overdose, misuse, extravasation, lack of efficacy, and occupational exposure associated with the use of a Pfizer product, must be reported, within 24 hours of awareness, to Pfizer Safety using the NIS AEM Report Form.

- For exposure during pregnancy in studies including pregnant women, data on the exposure to Comirnaty Original/Omicron BA.1 or Comirnaty Original/Omicron BA.4-5 vaccines during pregnancy, are not reportable unless associated with serious or non-serious adverse events.

- For these AEs with an explicit attribution or scenarios involving exposure to a Pfizer product, the safety information identified in the unstructured data reviewed is

captured in the Event Narrative section of the report form, and constitutes all clinical information known regarding these AEs. No follow-up on related AEs will be conducted.

All the demographic fields on the NIS AEM report form may not necessarily be completed, as the form designates, since not all elements will be available due to privacy concerns with the use of secondary data sources. While not all demographic fields will be completed, at the very least, at least one person identifier (eg, gender, age as captured in the narrative field of the form) will be reported on the NIS AEM report form, thus allowing the report to be considered a valid one in accordance with pharmacovigilance legislation. All identifiers will be limited to generalities, such as the statement, "A 35-year-old female..." or "An elderly male..." Other identifiers will have been removed.

Additionally, the onset/start dates and stop dates for "Illness," "Study Drug," and "Drug Name" may be documented in month/year (mmm/yyyy) format rather than identifying the actual date of occurrence within the month /year of occurrence in the day/month/year (DD/MMM/YYYY) format.

All research staff members must complete the following Pfizer training requirements:

- *"Your Reporting Responsibilities (YRR) with Supplemental Topics*."

These trainings must be completed by research staff members prior to the start of data collection. All trainings include a "Confirmation of Training Statement" (for signature by the trainee) as a record of completion of the training, which must be kept in a retrievable format. Copies of all signed training statements must be provided to Pfizer.

Re-training must be completed on an annual basis using the most current YRR with Supplemental Topics training materials. Where Pfizer issues an updated safety training program, including during the course of a calendar year, vendor shall ensure all vendor personnel complete the updated safety training within sixty (60) calendar days of issuance by Pfizer.

12. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

In compliance with the EMA Guideline on Good Pharmacovigilance Practices (GVP) Module VIII: Post-Authorisation Safety Studies (Rev 3),[2] the study and its protocol will be registered in the EU PAS Register prior to the start of data collection.

Results of analyses and interpretation will be delivered in report form.

- The interim report will include descriptive results and comparative analyses, when the data exist, and are accessible to support the analyses with sufficient precision.

- The final report will include the analysis and interpretation of each outcome, including pregnancy outcomes.

- The final study report will be posted in the EU PAS Register. Manuscripts using data from interim and final study data will be prepared for publication, preferably in a relevant peer-reviewed journal.

- Communication via other appropriate scientific venues will be considered.

Study results will be submitted for publication following guidelines, including those for authorship, established by the ICMJE.[54] When reporting results of this study, the appropriate Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist will be followed.[55] 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 Good Pharmacovigilance Practices (GVP) Module VIII: Post-Authorisation Safety Studies (Rev 3).*[2]

In the event of any prohibition or restriction imposed (eg, clinical hold) by an applicable competent authority in any area of the world, or if the investigator party responsible for collecting data from the participant is aware of any new information that might influence the evaluation of the benefits and risks of a Pfizer product, Pfizer should be informed immediately.

13. REFERENCES

1. European Medicines Agency (EMA). ETF concludes that bivalent original/Omicron BA.4-5 mRNA vaccines may be used for primary vaccination. 2022. Available from: https://www.ema.europa.eu/en/news/etf-concludes-bivalent-original-omicron-ba4-5-mrna-vaccines-may-be-used-primary-vaccination. Accessed on: 24 February 2023.
2. EMA. Guideline on good pharmacovigilance practices (GVP). Module VIII – Post-authorisation safety studies (EMA/813938/2011 Rev 3). 13 October 2017. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-gvp-module-viii-post-authorisation-safety-studies-rev-3_en.pdf. Accessed on: 17 January 2020.
3. World Health Organization (WHO). Update on Omicron. 28 November 2021 2021. Available from: https://www.who.int/news/item/28-11-2021-update-on-omicron.. Accessed on: 14 November 2022.
4. World Health Organization (WHO). Tracking SARS-CoV-2 variants. 2022. Available from: https://www.who.int/activities/tracking-SARS-CoV-2-variants. Accessed on: 14 November 2022.
5. Vitiello A, Ferrara F, Auti AM, et al. Advances in the Omicron variant development. J Intern Med. 2022;292(1):81-90.
6. Abu-Raddad LJ, Chemaitelly H, Ayoub HH, et al. Effect of mRNA vaccine boosters against SARS-CoV-2 Omicron infection in Qatar. N Engl J Med. 2022;386(19):1804-16.
7. Callaway E. Omicron likely to weaken COVID vaccine protection. Nature. 2021;600(7889):367-68.
8. Gram MA, Emborg H-D, Schelde AB, et al. Vaccine effectiveness against SARS-CoV-2 infection or COVID-19 hospitalization with the Alpha, Delta, or Omicron SARS-CoV-2 variant: A nationwide Danish cohort study. PLOS Medicine. 2022;19(9):e1003992.
9. European Medicines Agency (EMA). First adapted COVID-19 booster vaccines recommended for approval in the EU. 01 September 2022 2022. Available from: https://www.ema.europa.eu/en/news/first-adapted-covid-19-booster-vaccines-recommended-approval-eu. Accessed on: 14 November 2022.
10. European Medicines Agency (EMA). Adapted vaccine targeting BA.4 and BA.5 Omicron variants and original SARS-CoV-2 recommended for approval. 12 September 2022 2022. Available from: https://www.ema.europa.eu/en/news/adapted-vaccine-targeting-ba4-ba5-omicron-variants-original-sars-cov-2-recommended-approval. Accessed on: 14 November 2022.
11. European Medicines Agency (EMA). COVID-19 vaccines: authorised. 2022. Available from: https://www.ema.europa.eu/en/human-regulatory/overview/public-health-threats/coronavirus-disease-covid-19/treatments-vaccines/vaccines-covid-19/covid-19-vaccines-authorised#adapted-covid-19-vaccines-section. Accessed on: 12 December 2022.
12. European Medicines Agency (EMA). Comirnaty. 2023. Available from: https://www.ema.europa.eu/en/medicines/human/EPAR/comirnaty. Accessed on: 24 February 2023.

13. Hause AM, Marquez P, Zhang B, et al. Safety monitoring of bivalent COVID-19 mRNA vaccine booster doses among persons aged ≥12 years — United States, August 31–October 23, 2022. 2022. Available from: https://www.cdc.gov/mmwr/volumes/71/wr/mm7144a3.htm?s_cid=mm7144a3_w. Accessed on: 23 November 2022.
14. ENCePP. The ENCePP code of conduct for scientific independence and transparency in the conduct of pharmacoepidemiological and pharmacovigilance studies (Revision 4). 15 March 2018. Available from: http://www.encepp.eu/code_of_conduct/. Accessed on: 17 January 2020.
15. Kurz X, Bauchau V, Mahy P, et al. The ADVANCE Code of Conduct for collaborative vaccine studies. Vaccine. 2017;35(15):1844-55.
16. Willame C, Dodd C. Safety evaluation of COVID-19 vaccines in electronic healthcare databases: a protocol template from the ACCESS project. 2020. Available from: https://vac4eu.org/wp-content/uploads/2021/02/3c.Protocol_ACCESS_Safety-Evaluation-EHR.pdf. Accessed on: 23 November 2022.
17. García-Albéniz X, Hsu J, Hernán MA. The value of explicitly emulating a target trial when using real world evidence: an application to colorectal cancer screening. European journal of epidemiology. 2017;32(6):495-500.
18. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. Am J Epidemiol. 2016;183(8):758-64.
19. Hernán MA, Sauer BC, Hernández-Díaz S, et al. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. Journal of clinical epidemiology. 2016;79:70-75.
20. Suissa S. Immortal time bias in observational studies of drug effects. Pharmacoepidemiol Drug Saf. 2007;16(3):241-9.
21. Dagan N, Barda N, Kepten E, et al. BNT162b2 mRNA COVID-19 vaccine in a nationwide mass vaccination setting. N Engl J Med. 2021;384(15):1412-23.
22. Barda N, Dagan N, Ben-Shlomo Y, et al. Safety of the BNT162b2 mRNA Covid-19 Vaccine in a Nationwide Setting. New England Journal of Medicine. 2021;385(12):1078-90.
23. Dagan N, Barda N, Biron-Shental T, et al. Effectiveness of the BNT162b2 mRNA COVID-19 vaccine in pregnancy. Nature Medicine. 2021;27(10):1693-95.
24. Barda N, Dagan N, Cohen C, et al. Effectiveness of a third dose of the BNT162b2 mRNA COVID-19 vaccine for preventing severe outcomes in Israel: an observational study. The Lancet. 2021;398(10316):2093-100.
25. Magen O, Waxman JG, Makov-Assif M, et al. Fourth dose of BNT162b2 mRNA COVID-19 vaccine in a nationwide setting. New England Journal of Medicine. 2022;386(17):1603-14.
26. Cadarette SM, Maclure M, Delaney JAC, et al. Control yourself: ISPE-endorsed guidance in the application of self-controlled study designs in pharmacoepidemiology. Pharmacoepidemiol Drug Saf. 2021;30(6):671-84.
27. Kawai A, Arana A. (2020). Safety protocol for hospital case–based monitoring of specific adverse events following COVID-19 vaccines: a protocol template from the ACCESS project Available from: https://vac4eu.org/covid-19-vaccine-monitoring/. Accessed on: 7 April 2023.

28. Klein NP, Lewis N, Goddard K, et al. Surveillance for Adverse Events After COVID-19 mRNA Vaccination. JAMA. 2021;326(14):1390-99.
29. Rausch C, Yue Q-Y. Covid-19 vaccines and hearing loss and tinnitus. WHO Pharmaceuticals Newsletter. 2022;1(1):18-25.
30. Padniewski JJ, Jacobson-Dunlop E, Albadri S, et al. Stevens-Johnson syndrome precipitated by Moderna Inc. COVID-19 vaccine: a case-based review of literature comparing vaccine and drug-induced Stevens-Johnson syndrome/toxic epidermal necrolysis. Int J Dermatol. 2022;61(8):923-29.
31. Yih WK, Lee GM, Lieu TA, et al. Surveillance for adverse events following receipt of pandemic 2009 H1N1 vaccine in the Post-Licensure Rapid Immunization Safety Monitoring (PRISM) System, 2009-2010. Am J Epidemiol. 2012;175(11):1120-8.
32. Kuiper JG, Bakker M, Penning-van Beest FJA, et al. Existing data sources for clinical epidemiology: The PHARMO Database Network. Clin Epidemiol. 2020;12:415-22.
33. Overbeek JA, Swart KMA, Houben E, et al. Completeness and Representativeness of the PHARMO General Practitioner (GP) Data: A Comparison with National Statistics. Clin Epidemiol. 2023;15:1-11.
34. Recalde M, Rodríguez C, Burn E, et al. Data Resource Profile: The Information System for Research in Primary Care (SIDIAP). Int J Epidemiol. 2022;51(6):e324-e36.
35. Rothman KJ, Greenland S. Planning study size based on precision rather than power. Epidemiology. 2018;29(5):599-603.
36. ConcePTION. (2020). Report on existing common data models and proposals for ConcePTION (D7.5). Available from: https://zenodo.org/record/5829417#.Y_shtB_MKUl. Accessed on: 26 February 2023.
37. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Stat Med. 2009;28(25):3083-107.
38. Yang D, Dalton JE. A unified approach to measuring the effect size between two groups using SAS. 2012. Available from: https://support.sas.com/resources/papers/proceedings12/335-2012.pdf. Accessed on: 22 October 2020.
39. Austin PC, Cafri G. Variance estimation when using propensity-score matching with replacement with survival or time-to-event outcomes. Stat Med. 2020;39(11):1623-40.
40. Austin PC. Variance estimation when using inverse probability of treatment weighting (IPTW) with survival analysis. Stat Med. 2016;35(30):5642-55.
41. Hernan MA, Robins JM. IP weighting and marginal structural models. In *Causal inference*. Edited by. Boca Raton: Chapman & Hall/CRC; 2020.
42. Ho FK, Man KKC, Toshner M, et al. Thromboembolic Risk in Hospitalized and Nonhospitalized COVID-19 Patients: A Self-Controlled Case Series Analysis of a Nationwide Cohort. Mayo Clin Proc. 2021;96(10):2587-97.
43. Modin D, Claggett B, Sindet-Pedersen C, et al. Acute COVID-19 and the incidence of ischemic stroke and acute myocardial infarction. Circulation. 2020;142(21):2080-82.

44. Katsoularis I, Fonseca-Rodríguez O, Farrington P, et al. Risk of acute myocardial infarction and ischaemic stroke following COVID-19 in Sweden: a self-controlled case series and matched cohort study. Lancet. 2021;398(10300):599-607.
45. ENCePP. Guide on methodological standards in pharmacoepidemiology (EMA/95098/2010 Rev. 7). July 2018. Available from: http://www.encepp.eu/standards_and_guidances/methodologicalGuide.shtml. Accessed on: 22 October 2020.
46. Øvrelid E, Bygstad B, Thomassen G. TSD: A research platform for sensitive data. Procedia Computer Science. 2021;181:127-34.
47. University of Oslo. (2023). Services for sensitive data (TSD). Available from: https://www.uio.no/english/services/it/research/sensitive-data/. Accessed on: 26 February 2023.
48. ISPE. Guidelines for good pharmacoepidemiology practices (GPP). Revision 3. June 2015. Available from: https://www.pharmacoepi.org/resources/policies/guidelines-08027/. Accessed on: 20 January 2021.
49. International Society for Pharmacoepidemiology. (2015). Guidelines for good pharmacoepidemiology practices (GPP). Revision 3. International Society for Pharmacoepidemiology, Updated June. Available from: https://www.pharmacoepi.org/resources/policies/guidelines-08027/. Accessed on: 26 September 2022.
50. ICH. Pharmacovigilance planning. E2E. 2004. Available from: https://database.ich.org/sites/default/files/E2E_Guideline.pdf. Accessed on: 22 October 2020.
51. European Commission. Commission implementing Regulation (EU) No 520/2012 of 19 June 2012 on the performance of pharmacovigilance activities provided for in Regulation (EC) No 726/2004 of the European Parliament and of the Council and Directive 2001/83/EC of the European Parliament and of the Council. 20 June 2012. Available from: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:159:0005:0025:EN:PDF. Accessed on: 22 October 2020.
52. ENCePP. ENCePP checklist for study protocols (revision 4). 15 October 2018. Available from: http://www.encepp.eu/standards_and_guidances/checkListProtocols.shtml. Accessed on: 22 October 2020.
53. ENCePP. The European Union electronic register of post-authorisation studies (EU PAS Register). 20 December 2018. Available from: http://www.encepp.eu/encepp_studies/indexRegister.shtml. Accessed on: 22 October 2020.
54. ICMJE. Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals. December 2019. Available from: http://www.icmje.org/recommendations/. Accessed on: 22 October 2020.
55. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Journal of clinical epidemiology. 2008;61(4):344-9.

14. LIST OF TABLES

15. LIST OF FIGURES

ANNEX 1. LIST OF STAND ALONE DOCUMENTS

None.

ANNEX 2. ENCEPP CHECKLIST FOR STUDY PROTOCOLS

ENCePP Checklist for Study Protocols (Revision 4)

Study title: Post-Authorisation Safety Study of Comirnaty Original/Omicron BA.1 and Comirnaty Original/Omicron BA.4-5 in Europe

EU PAS Register® number:
Study reference number (if applicable): C4591052

Section 1: Milestones	**Yes**	**No**	**N/A**	**Section Number**
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection[2]	☒	☐	☐	6
1.1.2 End of data collection[3]	☒	☐	☐	6
1.1.3 Progress report(s)	☐	☐	☒	
1.1.4 Interim report(s)	☒	☐	☐	6
1.1.5 Registration in the EU PAS Register®	☒	☐	☐	6
1.1.6 Final report of study results.	☒	☐	☐	6

Comments:

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? (eg 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?	☒	☐	☐	8
2.1.3 The target population? (ie population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	9.1.1.1
2.1.4 Which hypothesis(-es) is (are) to be tested?	☒	☐	☐	8
2.1.5 If applicable, that there is no *a priori* hypothesis?	☐	☐	☒	

Comments:

[2] 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.

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

Section 3: Study design		**Yes**	**No**	**N/A**	**Section Number**
3.1	Is the study design described? (eg cohort, case-control, cross-sectional, other design)	☒	☐	☐	9.1
3.2	Does the protocol specify whether the study is based on primary, secondary or combined data collection?	☒	☐	☐	9.4, 9.6.1
3.3	Does the protocol specify measures of occurrence? (eg, rate, risk, prevalence)	☒	☐	☐	9.1, 9.7.1.2, 9.7.2.1
3.4	Does the protocol specify measure(s) of association? (eg risk, odds ratio, excess risk, rate ratio, hazard ratio, risk/rate difference, number needed to harm (NNH))	☒	☐	☐	9.1, 9.7.1.4, 9.7.2.2
3.5	Does the protocol describe the approach for the collection and reporting of adverse events/adverse reactions? (eg adverse events that will not be collected in case of primary data collection)	☒	☐	☐	11

Comments:

Section 4: Source and study populations		**Yes**	**No**	**N/A**	**Section Number**
4.1	Is the source population described?	☒	☐	☐	9.4
4.2	Is the planned study population defined in terms of:				
	4.2.1 Study time period	☒	☐	☐	9.2.3
	4.2.2 Age and sex	☒	☐	☐	9.2.1
	4.2.3 Country of origin	☒	☐	☐	9.4
	4.2.4 Disease/indication	☒	☐	☐	9.2.1
	4.2.5 Duration of follow-up	☒	☐	☐	9.1.1.4
4.3	Does the protocol define how the study population will be sampled from the source population? (eg event or inclusion/exclusion criteria)	☒	☐	☐	9.2.1, 9.2.2

Comments:

Section 5: Exposure definition and measurement		**Yes**	**No**	**N/A**	**Section Number**
5.1	Does the protocol describe how the study exposure is defined and measured? (eg operational details for defining and categorising exposure, measurement of dose and duration of drug exposure)	☒	☐	☐	9.3.1
5.2	Does the protocol address the validity of the exposure measurement? (eg precision, accuracy, use of validation sub-study)	☒	☐	☐	9.3.1
5.3	Is exposure categorised according to time windows?	☐	☒	☐	
5.4	Is intensity of exposure addressed? (eg dose, duration)	☐	☒	☐	

Section 5: Exposure definition and measurement		**Yes**	**No**	**N/A**	**Section Number**
5.5	Is exposure categorised based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	☒	☐	☐	9.3.2
5.6	Is (are) (an) appropriate comparator(s) identified?	☒	☐	☐	9.1.1.1

Comments:

Section 6: Outcome definition and measurement		**Yes**	**No**	**N/A**	**Section Number**
6.1	Does the protocol specify the primary and secondary (if applicable) outcome(s) to be investigated?	☒	☐	☐	9.3.2
6.2	Does the protocol describe how the outcomes are defined and measured?	☒	☐	☐	9.3.2
6.3	Does the protocol address the validity of outcome measurement? (eg precision, accuracy, sensitivity, specificity, positive predictive value, use of validation sub-study)	☒	☐	☐	9.6.1
6.4	Does the protocol describe specific outcomes relevant for Health Technology Assessment? (eg HRQoL, QALYs, DALYS, health care services utilisation, burden of disease or treatment, compliance, disease management)	☐	☐	☒	

Comments:

Section 7: Bias		**Yes**	**No**	**N/A**	**Section Number**
7.1	Does the protocol address ways to measure confounding? (eg confounding by indication)	☒	☐	☐	9.7.1.6
7.2	Does the protocol address selection bias? (eg healthy user/adherer bias)	☒	☐	☐	9.7, 9.9
7.3	Does the protocol address information bias? (eg misclassification of exposure and outcomes, time-related bias)	☒	☐	☐	9.7, 9.9

Comments:

Section 8: Effect measure modification		**Yes**	**No**	**N/A**	**Section Number**
8.1	Does the protocol address effect modifiers? (eg collection of data on known effect modifiers, sub-group analyses, anticipated direction of effect)	☒	☐	☐	9.9

Comments:

Section 9: Data sources		**Yes**	**No**	**N/A**	**Section Number**
9.1	Does the protocol describe the data source(s) used in the study for the ascertainment of:				
	9.1.1 Exposure? (eg pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview)	☒	☐	☐	9.3.1
	9.1.2 Outcomes? (eg clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics)	☒	☐	☐	9.3.2
	9.1.3 Covariates and other characteristics?	☒	☐	☐	9.3.3
9.2	Does the protocol describe the information available from the data source(s) on:				
	9.2.1 Exposure? (eg date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☒	☐	☐	9.3.1
	9.2.2 Outcomes? (eg date of occurrence, multiple event, severity measures related to event)	☒	☐	☐	9.3.2
	9.2.3 Covariates and other characteristics? (eg age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	☒	☐	☐	9.3.3
9.3	Is a coding system described for:				
	9.3.1 Exposure? (eg WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	☒	☐	☐	9.3.1
	9.3.2 Outcomes? (eg International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))	☐	☒	☐	
	9.3.3 Covariates and other characteristics?	☐	☒	☐	
9.4	Is a linkage method between data sources described? (eg based on a unique identifier or other)	☒	☐	☐	9.4

Comments:

Section 10: Analysis plan		**Yes**	**No**	**N/A**	**Section Number**
10.1	Are the statistical methods and the reason for their choice described?	☒	☐	☐	9.7
10.2	Is study size and/or statistical precision estimated?	☒	☐	☐	9.5
10.3	Are descriptive analyses included?	☒	☐	☐	9.7.1.2, 9.7.2.1
10.4	Are stratified analyses included?	☒	☐	☐	9.7.1.5
10.5	Does the plan describe methods for analytic control of confounding?	☒	☐	☐	9.7.1.6
10.6	Does the plan describe methods for analytic control of outcome misclassification?	☒	☐	☐	9.6.1
10.7	Does the plan describe methods for handling missing data?	☒	☐	☐	9.7.1.8
10.8	Are relevant sensitivity analyses described?	☒	☐	☐	9.7.1.9

Comments:

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Section Number**
11.1 Does the protocol provide information on data storage? (eg software and IT environment, database maintenance and anti-fraud protection, archiving)	☒	☐	☐	9.6
11.2 Are methods of quality assurance described?	☒	☐	☐	9.8
11.3 Is there a system in place for independent review of study results?	☒	☐	☐	10.5

Comments:

Section 12: Limitations	**Yes**	**No**	**N/A**	**Section Number**
12.1 Does the protocol discuss the impact on the study results of:				
12.1.1 Selection bias?	☒	☐	☐	9.9
12.1.2 Information bias?	☒	☐	☐	9.9
12.1.3 Residual/unmeasured confounding? (eg anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods).	☒	☐	☐	9.9
12.2 Does the protocol discuss study feasibility? (eg study size, anticipated exposure uptake, duration of follow-up in a cohort study, patient recruitment, precision of the estimates)	☒	☐	☐	9.9

Comments:

Section 13: Ethical/data protection issues	**Yes**	**No**	**N/A**	**Section Number**
13.1 Have requirements of Ethics Committee/ Institutional Review Board been described?	☒	☐	☐	10
13.2 Has any outcome of an ethical review procedure been addressed?	☐	☐	☒	
13.3 Have data protection requirements been described?	☒	☐	☐	9.6.5, 10.1

Comments:

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

Comments:

Section 15: Plans for communication of study results	**Yes**	**No**	**N/A**	**Section Number**
15.1 Are plans described for communicating study results (eg to regulatory authorities)?	☒	☐	☐	6, 12
15.2 Are plans described for disseminating study results externally, including publication?	☐	☐	☐	12

Comments:

Name of the main author of the protocol: Xabier Garcia de Albeniz

Date: dd/Month/year 24/April/2023

Signature: Xabier Garcia de Albeniz

ANNEX 3. ADDITIONAL INFORMATION

Not applicable

Document Approval Record

Document Name:	C4591052_PROTOCOL AMENDMENT 2_V3.0_08JAN2024
Document Title:	C4591052_PROTOCOL AMENDMENT 2_V3.0_08JAN2024

Signed By:	Date(GMT)	Signing Capacity
Younus, Muhammad	12-Jan-2024 13:52:43	Final Approval
De Bernardi, Barbara	12-Jan-2024 21:20:49	EUQPPV Approval