ecdc
EUROPEAN CENTRE FOR DISEASE PREVENTION AND CONTROL

TECHNICAL REPORT

Protocol for a COVID-19 vaccine effectiveness study using health data registries

www.ecdc.europa.eu

ECDC TECHNICAL REPORT

Protocol for a COVID-19 vaccine effectiveness study using health data registries

Version 1.0

This pilot protocol was commissioned by the European Centre for Disease Prevention and Control (ECDC), as part of the activities within the Framework Contract (FWC) ECDC/2021/018 'Vaccine Effectiveness, Burden and Impact Studies (VEBIS) of COVID-19 and Influenza Lot 4'. Production of the protocol was coordinated by Nathalie Nicolay.

Authors
Susana Monge Corella, Amparo Larrauri (Instituto de Salud Carlos III, Spain), Irina Kislaya, Baltazar Nunes (Instituto Nacional de Saúde, Portugal), Cristina Buguri, Itziar Casado, Jesús Castilla, Iván Martínez-Baz (Instituto de Salud Pública y Laboral de Navarra), Hanne-Dorthe Emborg, Katrin Finderup, Christian Holm Hansen (Statens Serum Institut, Denmark), Jostein Starrfelt, Hinta Meijerink (Norwegian Institute of Public Health), Anthony Nardone, Alexis Sentís, Marta Valenciano (Epiconcept, France).

Suggested citation: European Centre for Disease Prevention and Control. Protocol for a COVID-19 vaccine effectiveness study using health data registries. Stockholm: ECDC; 2023.

Stockholm, January 2023

ISBN 978-92-9498-614-6
ISSN 2315-0947
doi: 10.2900/125804
Catalogue number : TQ-03-23-039-EN-N

Contents

Abbreviations

BMI	Body Mass Index
COVID-19	Coronavirus disease 2019
CVE	COVID-19 vaccine effectiveness
ECDC	European Centre for Disease Prevention and Control
EEA	European Economic Area
EMA	European Medicines Agency
EU	European Union
HR	Hazard Ratio
ICU	Intensive care Unit
PCV	Pneumococcal Conjugated Vaccine
PPV	Pneumococcal Polysaccharide Vaccine
RR	Risk Ratio
RT- PCR	Reverse-transcription polymerase chain reaction
SARI	Severe acute respiratory infection
SARS-CoV-2	Severe acute respiratory syndrome – coronavirus 2
VE	Vaccine effectiveness

Executive summary

This protocol presents a common methodology to estimate vaccine effectiveness (VE) for COVID-19, using established health data registries in participating EU/EEA Member States. This work is performed within the Vaccine Effectiveness, Burden and Impact Studies (VEBIS) framework contract. A pilot study with the participation of Denmark, Spain (Navarre) , Norway and Portugal was carried out between April and June 2022, following a common pilot protocol (published by ECDC in 2022) [1, 2]. The experience gained in the pilot informed the current protocol.

The study design is a retrospective cohort study using data collected routinely in electronic health records databases, with data on the resident community-dwelling population (i.e., excluding those living in nursing homes) who belong to an age group for whom vaccination has been universally recommended at the time of the study. Outcomes of interest include hospital admission or death due to COVID-19. Other data to be collected include socio-demographic (age, sex), clinical (comorbidities, previous history of SARS-CoV-2 infection) and COVID-19 vaccination (brand, number and dates of dose administration) variables. The protocol outlines the agreed methods for analysing available data related to COVID-19 outcomes at country level and includes a plan for the pooled analysis. The proposed approach of the study is to produce prospectively monthly VE estimates to monitor VE over time. Vaccine effectiveness will be estimated in study periods covering an eight-week follow-up time. A minimum of one month between the end of the study period and the data extraction date will be omitted, to allow consolidation of data, particularly regarding hospital discharge data in which estimates for some study sites are based. Therefore, this protocol defines a rolling-study period in which the eight-week window is moved forward one month for each successive monthly estimate.

This master protocol is primarily intended to guide the implementation of ECDC-funded studies. However, ECDC encourages the conduct of vaccine effectiveness studies, using this protocol as a basis, in countries that do not currently plan to participate in ECDC-funded studies. The use of consistent protocols will facilitate the comparability of results across studies, countries and sites.

Background

The end of 2019 saw the emergence of a novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which causes coronavirus disease 2019 (COVID-19) with substantial disease burden [3]. As of December 2022, seven vaccines (Comirnaty, COVID-19 Vaccine Valneva, Nuvaxovid [previously Novavax], Spikevax [previously COVID-19 vaccine Moderna], Vaxzevria [previously AstraZeneca], and Jcovden [previously Covid-19 Vaccine Janssen] and VidPrevtyn Beta (indicated as a booster)) and three adapted vaccines as boosters (Comirnaty Original/Omicron BA.1, Comirnaty Original/Omicron BA.4-5, Spikevax bivalent Original/Omicron BA.1) have been authorised by the European Commission based on the scientific opinion of the European Medicines Agency (EMA) for use in the European Union, and many others are under rolling review (Annex 1) [4].

ECDC COVID-19 vaccine effectiveness studies

In 2020, the European Commission stressed the importance of continuously monitoring the safety and effectiveness of vaccines in the EU/EEA in the post-authorisation phase, with particular emphasis on COVID-19 vaccines in the context of the ongoing pandemic [5]. The 2018 Council Recommendation on Strengthened Cooperation against Vaccine-preventable Diseases asked ECDC and EMA to cooperate in ensuring the continued monitoring of vaccines and vaccination in use in EU/EEA vaccination programmes [6]. Such a request was subsequently formalised as part of the extended EMA regulatory mandate [7] and ECDC's newly amended mandate [6], which request the two Agencies develop a structured and independent post-authorisation vaccine monitoring platform, initially prioritising COVID-19 vaccines, which requested the two Agencies to develop a structured and independent post-authorisation vaccine monitoring platform, initially prioritising COVID-19 vaccines. ECDC and EMA officially established and launched such a platform in May 2022, with the intention of bringing together public health and regulatory experts to discuss the studies needed to generate real-life evidence on the safety and effectiveness of vaccines in use in EU/EEA immunisation programmes.

In 2020, utilising the lessons learned from other vaccine effectiveness studies, ECDC started building an infrastructure to perform COVID-19 vaccine effectiveness (CVE) studies [1-2, 9-16] (Annex 2). Building a system to regularly monitor CVE and perform studies in different settings, and depending on the setting, to provide information on different outcomes (severe disease, moderate disease, infection, transmission, etc) is underway. At the start of 2022, ECDC expanded this infrastructure to include monitoring of VE. The multi-country approach of the effectiveness studies is also one of the key features that characterises them, as well as the inclusion of more countries over time.

One study aims to assess the vaccine effectiveness and the impact of COVID-19 vaccines through routinely collected vaccination status and outcome data using health registries. The current protocol describes the methods to implement such studies. A pilot study with the participation of Denmark, Spain (Navarre), Norway and Portugal was carried out between April and June 2022, following a common pilot protocol (published by ECDC in 2022) [1, 2]. The experience gained in the pilot informed the current master protocol.

The studies will rely on routinely collected vaccination status and COVID-19 outcome data using established health data registries across several study sites in the EU/EEA. The main objective is to set up a robust network of study sites with the capacity to use health registries under common criteria to monitor on a continuous basis COVID-19 vaccines effectiveness and their impact. The overall aim will be to provide a real-life monitoring system to detect any changes in VE over time and in population subgroups. In the future and once the system is established, it may be used to address further specific questions.

Overall aim

The overall aim of the study based on this protocol is to monitor real-time performance of COVID-19 vaccines in the community-dwelling resident population in EU/EEA countries to detect any signal of reduced vaccine effectiveness, so that public health vaccine recommendations may be adjusted accordingly. To achieve this aim, hazard ratio or rate ratio of outcomes of interest will be estimated using information routinely collected in various electronic health registries, including vaccination, population and health databases and using deterministic data linkage.

In the following sections, arrow marks with italicised text indicate the points that countries could further expand/detail when creating a country-specific protocol using the current ECDC protocol.

Objectives

Principal objective

To measure vaccine effectiveness (VE) of COVID-19 vaccines in community-dwelling resident populations in EU/EEA countries, against the following outcomes:

- Hospital admission due to COVID-19 with laboratory-confirmed SARS-CoV-2 infection
- Death due to COVID-19 with laboratory-confirmed SARS-CoV-2 infection.

Vaccine effectiveness is measured for individuals who have a complete vaccination status with primary series and for individuals who have a complete vaccination status with booster doses (first booster, second, booster or as appropriate, according to current recommendations), compared to non-vaccinated individuals.

Secondary objectives

To measure COVID-19 vaccine effectiveness (if sample size allows):

- By time since the booster dose (first, second, third or subsequent booster doses as appropriate, according to current recommendations): number of weeks elapsed between current time and the date of the booster, to evaluate the presence of waning protection.
- Relative vaccine effectiveness of the first, second, and first booster dose (if such booster doses would be recommended): Comparing individuals vaccinated with complete primary vaccination course plus a first, second, or third booster COVID-19 vaccine dose vs. individuals with complete primary vaccination course of COVID-19 vaccines only, with last primary dose administered ≥24 weeks (i.e., ≥169 days).
- Relative vaccine effectiveness of the second (or subsequent) booster(s) dose (if such booster doses would be recommended): Comparing individuals vaccinated with the second (or subsequent) booster COVID-19 vaccine dose vs. individuals eligible for the second (or subsequent) booster COVID-19 vaccine dose but who did not receive it (yet).
- Against intensive care unit (ICU) admission due to COVID-19 with laboratory-confirmed SARS-CoV-2 infection.
- By vaccine products used as a booster dose, depending on the evolving context and types of vaccines used (mRNA vaccines vs. non-mRNA vaccines, Omicron-non-specific vaccines vs Omicron-specific vaccines).

Alternatives

- Study sites can contribute to all or only a subset of the established objectives (for example, only some outcomes within the main objectives), or only to some of the secondary objectives.
- Additional objectives might be added as the situation evolves and new questions arise.

→ *Study sites/countries to specify the study objectives*

Methodology

Study design

A retrospective cohort study using data collected routinely in electronic health record databases with a comparison of the risk or rate of outcome occurrence between individuals with different vaccination status.

Study setting

A retrospective cohort study using data routinely collected from community-dwelling individuals (i.e., excluding those living in nursing homes) in national electronic health record databases. The study population should belong to an age group for whom vaccination has been universally recommended at the time of the analysis.

→ Study sites/countries to include more details on the areas covered by the study, population and representativeness. Include brief information on the vaccination rollout, such as prioritised groups and dates in which different populations were incorporated into the vaccination programme.

Study period

VE estimates will be produced monthly, to monitor VE over time. VE will be estimated in study periods covering 8-weeks follow-up time, to allow sufficient events to support the estimations (the study window may need to be increased ad hoc if, due to the evolving situation, the number of events results insufficient), as well as to be sensitive and reactive to changes in VE over time. A minimum of one month between the end of the study period and the data extraction date will be omitted, to allow consolidation of data, particularly regarding hospital discharge data in which estimates for some study sites are based. Therefore, this protocol defines a rolling-study period in which the eight-week window is moved one month forward for each successive monthly estimate.

For example, if the study is conducted with data extracted in March 2022, the follow-up period will be the eight weeks starting on 1 December 2021 (t0) to 25 January 2022 (administrative censoring time). For data extracted in April 2022, the eight-week window will be moved one month forward, therefore starting on 1 January 2022 (t0) and up to 25 February 2022. Example follow-up periods for monthly estimates are reported in Table 1.

Table 1. Follow-up period corresponding to each monthly estimate, pilot study, 2022

Data extraction and analysis *	Study period (follow-up time)
January 2022	October 1 to November 25, 2021
February 2022	November 1 to December 26, 2021
March 2022	December 1, 2021, to January 25, 2022
April 2022	January 1 to February 25, 2022
May 2022	February 1 to March 28, 2022
June 2022	March 1 to April 25, 2022
July 2022	April 1 to May 26, 2022
August 2022	May 1 to June 25, 2022
September 2022	June 1 to July 26, 2022

** When estimates are drawn retrospectively for past monthly estimates, date of data extraction and analysis may be later that in this example table*

Study population

The study population includes individuals targeted by the national vaccination plan/programme and/or the reference population registries fulfilling the following criteria during the different study periods:

- Aged between 5 and 110 years.
- Resident in an EU/EEA territory covered in the study.
- Excluding those living in a nursing home.
- Vaccinated as indicated by age (i.e., excluding those vaccinated before it was generally recommended in the corresponding age-group or, alternatively, the first 5% of persons vaccinated within each age-group – for each 5-year age bracket- as these first vaccinees may not be representative of their corresponding age group).

- Exclude individuals with inconsistent or missing data on vaccination (vaccination status unknown, any vaccination date is unknown, any vaccine brand is unknown, number of doses is unknown, interval between first and second dose is shorter than 19 days, interval between complete vaccination and booster dose or between booster doses is shorter than 90 days, number of doses higher than recommended, received any vaccine brand not approved by EMA, and the combination of vaccine brands is not a recommended schedule -may vary by age group).

→ Study sites/countries to specify and describe the study population, and specify any deviation (e.g. not possible to identify residents of nursing homes)

Definitions

Vaccination status

The vaccination status will be based on vaccine doses administered up to the date to which vaccination status is assessed (as a time-changing variable), and individual will be classified as follows:

- Non-vaccinated: has not received any vaccine dose.
- Complete vaccination with primary series of COVID-19 vaccines: individuals who received the primary series of COVID-19 vaccine doses defined as one dose of Ad26.COV2.S (Jcovden) vaccine or two doses of ChAdOx1-S (Oxford/Astra Zeneca), BNT162b2 (Pfizer-BioNTech) or mRNA-1273 (Moderna), or combination of any of the three vaccines. The two doses should be administered no less than 19 days apart. The status of completion of primary series is achieved 14 days after the date of administration of the final dose required for complete primary series vaccination (induction period).
- Complete vaccination with first booster COVID-19 vaccine dose: individuals who received an additional dose of BNT162b2, mRNA-1273 or ChAdOx1-S at least three months (90 days) after the date of complete primary series (as defined above). This status is achieved 14 days after the date of administration of the booster dose (induction period).
- Complete vaccination with second booster COVID-19 vaccine dose: individuals who received an additional dose of BNT162b2, mRNA-1273 or ChAdOx1-S at least 3 months (90 days) after the first booster COVID-19 vaccine dose (as defined above). This status is achieved 14 days after the date of administration of the booster dose (induction period).
- Complete vaccination with third booster COVID-19 vaccine dose: individuals who received an additional dose of BNT162b2, mRNA-1273 or ChAdOx1-S at least 3 months (90 days) after the second booster COVID-19 vaccine dose (as defined above). This status is achieved 14 days after the date of administration of the booster dose (induction period).

Any individual who received at least one dose of a vaccine but does not fulfil the definition of complete vaccination with primary series of COVID-19 vaccines will be considered incompletely vaccinated. This category will be analysed but will not be reported as a vaccination category for this study. Likewise, individuals who received a second vaccine dose or a booster dose within less than 14 days after the administration of the last dose, will be considered separately (i.e. not merged with the previous nor the subsequent vaccination status). They need to be included in the model for completeness, but estimation for these groups will not to be reported for this study.

Individuals who received the second dose within less than 90 days after the first vaccine dose, or that received a subsequent vaccine dose (any vaccine dose after complete vaccination) that do not fulfil the definition criteria of a booster (either the dose was administered too early or it was with a vaccine not listed above as accepted as a booster) will be considered information errors and be dropped from the risk set. Also, persons with vaccine brands or vaccination schedules not included in the national vaccination programme or unknown will be dropped (see selection criteria).

Time since vaccination with a booster dose

In response to the secondary study objective on the waning of vaccine effectiveness, the alternative exposure definition that accounts for time since uptake of each dose will be used.

Time since vaccination will be computed at each point in time by constructing a time-dependent variable. Time since the first, second or third booster (as appropriate) will be calculated, though the focus of interest may change as vaccine recommendations evolve.

Assessment of time since vaccination will start at the end of the induction period (as defined in the vaccination status definitions). This means that, for the time since the first, second or third booster vaccination, time 0 will be day 14 after the date of administration of the booster dose. Time since vaccination will be classified into three categories as follows:

- From time 0 (as defined above) up to ≤84 days after time 0 (i.e. <12 weeks).
- Days 85–168, both included, after time 0 (i.e. ≥12 weeks & <24 weeks).
- ≥169 days, after time 0 (i.e. ≥24 weeks).

Outcomes

Principal outcomes of interest are defined as:

- Hospital admission due to COVID-19:
  - admission to hospital:
  - with a laboratory-confirmed infection between 24 hours after or up to 14 days before admission, in which admission criteria are compatible with sever acute respiratory infections (SARI - based on similar criteria as in SARI surveillance, International Classification of Diseases (ICD) codes or similar).

  or

  - in which COVID-19 is the main diagnosis in the discharge record (for example, based on ICD coding or similar).
- Death due to COVID-19: death for which COVID-19 is recorded as the cause of death OR, if cause of death not available, laboratory-confirmed infection with death in the 30 days after a positive test or symptoms onset.

For every outcome, the censoring date of outcome occurrence will be the date of laboratory diagnosis (i.e. the date of the first diagnosis of the infection episode that resulted in hospital admission, ICU admission or death, respectively).

Alternatives

- Study sites may limit the objectives of the study to some of those listed above.
- Sites not being able to identify hospitalisations due to COVID-19 (using the proposed or other similar definition) may provide hospitalisations with COVID-19 instead, provided this is well documented in the site-specific protocol and the reporting methods (Annex 5).

→ *Study sites/studies to define the outcomes used and the definitions*

Stratification variables

Age group

Age will be calculated at the beginning of each study period using the date of birth and categorised into 5-years bins to adjust models. For reporting stratified results by age-group the following groups will be used 5-17, 18-49, 50-64, 65-79, 80+ years. Alternative age groups may be discussed depending on needs.

Vaccine product

Vaccine effectiveness estimates by vaccine product will be defined according to booster doses and depending on the question of interest at each time, it may be defined according to primary vaccination series.

Primary series vaccine products will be classified as:

- One dose of Ad26.COV2.S (Jcovden).
- Two doses of ChAdOx1-S (Oxford/Astra Zeneca), administrated with minimum of 19 days between first and second dose.
- One dose of ChAdOx1-S (Oxford/Astra Zeneca), followed by an mRNA vaccine (either mRNA-1273 (Moderna) or BNT162b2 (Pfizer-BioNTech) administrated with minimum of 19 days between first and second dose.

- Two doses of an mRNA vaccine (either mRNA-1273 (Moderna) or BNT162b2 (Pfizer-BioNTech), that are considered interchangeable in this protocol), administrated with minimum of respectively 19 days between first and second dose.

The specific brand of booster product will be recorded, and broader categories may be defined ad hoc (e.g. mRNA or non-mRNA, Omicron-specific or non-Omicron-specific, bivalent or monovalent, etc.).

Primary series and boosters' brands analysis will be performed separately. Further analysis will be explored to compare VE estimates, considering different vaccine products categories: 1) vaccinated (primary and booster) with mRNA vaccines, 2) vaccinated (primary and booster) with non-mRNA vaccines, and 3) vaccinated (primary and booster) with a combination of mRNA and non-mRNA vaccines. More categories may be considered depending on the sample size by category among all the participating sites.

The current classification will be updated when needed (e.g., with Omicron-specific vaccines or new vaccines).

→ Study sites/countries to specify the definitions used for stratification variables

Potential confounding variables for adjustment

Previous infection

Previous infection will be classified into two categories at the beginning of each study period:

- No previous infection: no positive SARS-CoV-2 test recorded before the first day of the study period.
- Previous infection: At least one positive SARS-CoV-2 test recorded before the first day of the study period.

New positive SARS-CoV-2 tests after that date will only be considered events if they lead to hospitalisation or death (see the Outcomes definitions) but ignored in the analysis otherwise.

If available data allow, previous infection may also be defined as a three-level variable, also including the severity of the infection:

- No previous infection: no positive SARS-CoV-2 test recorded before the first day of the study period.
- Previous infection: at least one positive SARS-CoV-2 test recorded before the first day of the study period NOT leading to hospital admission (as defined in the study outcomes).
- Previous infection: at least one positive SARS-CoV-2 test recorded before the first day of the study period leading to hospital admission (as defined in the study outcomes).

→ Study sites/countries to specify the definitions used for stratification variables

Sociodemographic

- Sex.
- Age group.
- Individual level socioeconomic status (normally not available): Educational level, occupation, unemployment, income, as available in registries.
- Area level socioeconomic condition (postal code, municipality or other): income per capita, Gross Domestic Product per capita, inequality or deprivation index [17], unemployment rate, as available in other complementary data sources.
- Others.

Comorbidities and health seeking behaviours

Different variables can be used to account for comorbidities. For homogenisation purposes, it is recommended to include comorbidities as a three-level variable:

- No comorbidities related to increased risk of COVID-19 severe outcome.
- Medium-low risk comorbidities (for example, comorbidities that are associated with risk of COVID-19 severe outcome, but different from immunocompromising conditions, or other classification decided at site level).
- High risk comorbidities (for example, immunocompromising conditions, or other classification decided at site level)

Examples of comorbidities that can be considered in each category is provided in Annex 3.

Additionally, based on which variables are relevant to account for differences between the study sites, other adjusting variables can be considered:

- Number of consultations in primary care over the last 12 months, or another relevant timeframe (0, 1, 2, ≥3 consultations).
- Number of SARS-CoV-2 tests performed in the last year, or other relevant timeframe.
- Hospitalisation in the previous year, or other relevant timeframe (yes, no).
- Vaccination with other 'respiratory' vaccines: seasonal influenza or pneumococcus (yes, no).

→ Study sites/countries to specify the definitions used for effect modifiers/confounding variables

Data sources

The study will be using routinely collected data from various population health registries available at national or subnational level. Each database should contain a unique identifier for any individual to allow data linkage between databases.

Sources of information on the reference population

- The reference population database (for example, census database, health coverage database, etc.) with individual records of the target population. It should contain variables that allow the identification of non-residents, or temporary residents.
- In any of databases it is desirable that identification of specific population such as those living in nursing homes or other institutions is possible.

Sources of information on vaccination status

- Vaccination registry or vaccination record databases with record of individuals and dates of COVID-19 vaccination including vaccine product.

→ Study sites/countries to specify and describe sources of information on the vaccination status and potential limitations

Sources of information on outcomes

Data will be extracted from different electronic health record databases:

- COVID-19 laboratory-confirmed infection.
- Epidemiological surveillance databases (for notifiable diseases).
- Primary healthcare consultation.
- Hospital admission/discharge.
- Death or mortality registries with record of cause of death.

Alternative

- Study sites may only contribute with estimations for some but not all outcomes. Therefore, sources of data will be included according to which outcomes will be included in the site-specific study protocol.

→ Study sites/countries to specify and describe sources of information for each outcome and potential limitations.

Sources of information on effect modifiers or confounders

- Electronic databases informing on comorbidities: possibly primary healthcare records, data on medication prescriptions, or any other population-based data source that can provide information on comorbidities for all cohort individuals (and not only for cases).
- Electronic databases informing on healthcare seeking behaviour: possibly healthcare administrative databases (i.e. number of consultations), laboratory records (i.e. number of test performed) or non-COVID-19 vaccination records (i.e. other vaccines administered).

→ Study sites/countries to specify and describe sources of information for each effect modifier or confounding factor and potential limitations.

Construction of the cohort

Identification of individuals and characteristics at baseline

The reference population database will be linked with the electronic databases on vaccination, comorbidity and/or health-seeking behaviours registries using the unique identifier and a deterministic data linkage procedure (no random component in the linkage procedure).

People will enter the study in their corresponding group of vaccination status based on the data available in the vaccination registry.

Variables to be measured at baseline include age, sex, region (if relevant), previous infection, comorbidities, and other socioeconomic or health-seeking behaviour variables that will be used to adjust the models.

→ Study sites/countries to specify the data linkage method used

Time-changing characteristics

Vaccination status will be assessed, and people classified into the same or updated vaccination status on a daily (preferable) or weekly basis, generating a new record in the dataset for each new assessment. The variable with the time since last vaccine dose will be updated accordingly (Table 2).

Table 2. Example of implementation of time-dependent variables

Individual ID	Start day*	End day*	Vaccination status	Time since booster dose	Other variables classified at baseline (e.g., age, sex, previous infection or comorbidities)
12345	0	15	Complete vaccination	-	Constant
12345	16	29	-	-	Constant
12345	30	56	Booster dose	<12 weeks	Constant

Identification of outcomes during follow-up

Information for identification of outcomes and their occurrence dates will be obtained by individual deterministic linkage between the cohort built previously and the databases containing the information on the outcomes.

Outcome classification for each individual will be assessed from study follow-up start date (t0) and up to the administrative censoring date (eight weeks later). The date of outcome occurrence will be defined as the day before the date of the positive SARS-CoV-2 laboratory test (since outcomes are normally assigned to the vaccination status in the previous day).

Table 3. Example of dataset with information on the outcome

Individual ID	Start day*	End day*	Vaccination status	Time since booster dose	Event
12345	0	15	Complete vaccination	-	0
12345	16	29	-	-	0
12345	30	50	Booster dose	<12 weeks	***1***

Censoring events

All individuals will be followed from the start of the observation period to the:

- Date of a positive SARS-CoV-2 test leading to the event of interest.*
- Death of any cause (on the date of death).
- Discontinuation in the administrative database (i.e., emigration).
- Administrative censoring (eight weeks after the start of the observation period).

** Depending on the outcome being analysed: events will be positive tests leading to hospitalisation, ICU admission or death, as per the previous definitions.*

Analysis plan

Description of the sample selection

The total number of individuals in each of the databases fulfilling the inclusion criteria will be collected at least once per year along with basic information on the data extraction (Table 4). The number and proportion of individuals excluded after the application of each selection criteria will be recorded in a flowchart.

It is recommended that sites/countries perform an analysis of the time needed for data consolidation, for example, by comparing the number of events in the same fixed period in different data extractions. Then, the time needed for data consolidation will be defined as the number of days from the end of the observation period to the date of the data extraction needed to have a considerable proportion (i.e. around >80%) of events recorded in the final evaluated data extraction.

Table 4. Description of data extraction and number of individuals in the source data

	Source 1	Source 2	Source 3	Source 4
Name of database				
Date of extraction				
Last date in the dataset				
Number of individuals (before selection)				

Description of the study population

Information on the number of persons, total person-time of follow up and number of events by vaccination status and age, for the different events under study will be collected in the format for periodic reporting of VE results (Annex 4).

Distribution of key variables in the vaccination status groups considered in the study will be explored and collected monthly (Table 5).

The study-site background information can include the incidence rate, the distribution of vaccination coverage and the proportion of variant circulation at different points in time. This information will be extracted from ECDC data repositories on case notifications, vaccine coverage and variant circulation.

Information on dates in which the different age groups entered the vaccination programme for first vaccine dose, first booster dose or successive doses will be collected as part of the background information on the study setting (see corresponding section in this protocol). The proportion of the missing data will be used to determine if each specific variable can be included in the model and how (e.g., missing could eventually be included in the model as a category). Data imputation for missing data in order to improve quality is not planned but will be encouraged when possible.

Finally, it is recommended that sites look at the distribution of outcomes along the observation period for the different outcomes and age-groups, and by vaccination status, to assess any systematic difference in the vaccination status groups being compared.

Table 5. Socio-demographic, microbiological and clinical characteristics of study population included by vaccination status collected through the monthly reporting template (one per country/site)

	Not vaccinated	Complete vaccination with primary series of COVID-19 vaccines among included	Complete vaccination with first booster COVID-19 vaccine dose among included	Complete vaccination with second booster COVID-19 vaccine dose among included	Complete vaccination with third booster COVID-19 vaccine dose among included
	n	***n***	***n***	***n***	***n***
Site					
Number of individuals					
Sex					
Male					
Female					
Missing					
Age-group (years)					
5 to 17					
18 to 49					
50 to 64					
65 to 79					
80+					
Country of birth					
Native					
Non-native					
Missing					
Nationality					
National					
Non-national					
Missing					
Vaccine product - primary course (only including individuals who achieved full vaccination before end of study period)					
Comirnaty	NA				
Spikevax	NA				
Vaxzevria	NA				
Jcovden	NA				
Nuvaxovid	NA				
Other					
Missing	NA				
Vaccine product - first booster (only including individuals who received the first booster before end of study period)					
Comirnaty (Original)	NA	NA			
Spikevax (Original)	NA	NA			
Comirnaty (Adapted BA.1)	NA	NA			
Spikevax (Adapted BA.1)	NA	NA			
Comirnaty (Adapted BA.4/BA.5)	NA	NA			
Spikevax (Adapted BA.4/BA.5)	NA	NA			
Other	NA	NA			
Missing	NA	NA			
Vaccine product - second booster (only including individual who received the second booster before end of study period)					
Comirnaty (Original)	NA	NA	NA		
Moderna (Original)	NA	NA	NA		
Comirnaty (Adapted BA.1)	NA	NA	NA		

	Not vaccinated	Complete vaccination with primary series of COVID-19 vaccines among included	Complete vaccination with first booster COVID-19 vaccine dose among included	Complete vaccination with second booster COVID-19 vaccine dose among included	Complete vaccination with third booster COVID-19 vaccine dose among included
Spikevax (Adapted BA.1)	NA	NA	NA		
Comirnaty Adapted BA.4/BA.5)	NA	NA	NA		
Spikevax (Adapted BA.4/BA.5)	NA	NA	NA		
Other (AZ, others,...)	NA	NA	NA		
Missing					
Vaccine product - third booster (only including individual who received the third booster before end of study period)					
Comirnaty (Original)	NA	NA	NA	NA	
Spikevax (Original)	NA	NA	NA	NA	
Comirnaty (Adapted BA.1)	NA	NA	NA	NA	
Spikevax (Adapted BA.1)	NA	NA	NA	NA	
Comirnaty (Adapted BA.4/BA.5)	NA	NA	NA	NA	
Spikevax (Adapted BA.4/BA.5)	NA	NA	NA	NA	
Other (AZ, others,...)	NA	NA	NA	NA	
Missing					
Comorbiditites					
No comorbidity					
Low-medium risk comorbidities /non-immunocompromising					
High risk comorbidities/ immunocompromising					
Missing					

Estimation of vaccine effectiveness

Groups to be compared

Vaccine effectiveness will be estimated by comparing different vaccination status groups. In order to fulfil all the study objectives, the relevant comparisons include:

- Exposed group = complete vaccination with primary series of COVID-19 vaccines; Reference group = unvaccinated. Stratified by age group (5-17, 18-49, 50-64, 65-79, 80-110 years).
- Exposed group = complete vaccination with primary series of COVID-19 vaccines + first, second and/or third booster; Reference group = unvaccinated. Stratified by age group (5-17, 18-49, 50-64, 65-79, 80-110).
- Exposed group = complete vaccination with primary series of COVID-19 vaccines + first, second and/or third booster < 84 days ago; Reference group = unvaccinated. Stratified by age group (5-17, 18-49, 50-64, 65-79, 80-110).
- Exposed group = complete vaccination with primary series of COVID-19 vaccines + first, second and/or third booster 85 – 168 days ago; Reference group = unvaccinated. Stratified by age group (5-17, 18-49, 50-64, 65-79, 80-110).
- Exposed group = complete vaccination with primary series of COVID-19 vaccines + first, second and/or third booster ≥169 days ago; Reference group = unvaccinated. Stratified by age group (5-17, 18-49, 50-64, 65-79, 80-110).
- Exposed group = complete vaccination with primary series of COVID-19 vaccines + first, second and/or third booster; Reference group = complete vaccination with primary series of COVID-19 vaccines ≥169 days ago. Stratified by age group (5-17, 18-49, 50-64, 65-79, 80-110).
- Exposed group = complete vaccination with primary series of COVID-19 vaccines + first, second and/or third booster < 84 days ago; Reference group = complete vaccination with primary series of COVID-19 vaccines ≥169 days ago. Stratified by age group (5-17, 18-49, 50-64, 65-79, 80-110).
- Exposed group = complete vaccination with primary series of COVID-19 vaccines + first, second and/or third booster 85 – 168 days ago; Reference group = complete vaccination with primary series of COVID-19 vaccines ≥169 days ago. Stratified by age group (5-17, 18-49, 50-64, 65-79, 80-110).

- Exposed group = complete vaccination with primary series of COVID-19 vaccines + first, second and third booster ≥169 days ago; Reference group = complete vaccination with primary series of COVID-19 vaccines ≥169 days ago. Stratified by age group (5-17, 18-49, 50-64, 65-79, 80-110).

Those individuals that received the first, second or third booster but who did not fulfil the definition criteria of a booster will be considered information errors and dropped from the risk set.

It is recommended that contrasts that use the same reference group are performed using a single model with vaccination status defined by categories within a single variable. For instance, group comparisons numbered 1 and 2 above could be done in a single model, with a variable that equals 0 for non-vaccinated, 1 for complete vaccination with primary series and 2 for first booster (3 and 4 as second and third booster if applicable). Similar categories can be created in a single variable to be used in a single model in the comparison groups number as 3, 4 and 5, as well as for the groups 7, 8 and 9.

For analysing successive booster doses, as relevant, similar contrasts will be performed but with the exposed group being individuals with complete vaccination with primary series of COVID-19 vaccines + first vaccine booster + second vaccine booster; and the reference group will include either the unvaccinated or individuals with complete vaccination with primary series of COVID-19 vaccines + first vaccine booster ≥169 days ago.

Other contrasts can cover other secondary objectives (for example, by vaccine brand), but this will be defined on an ad hoc basis, depending on the epidemiological context and the interest of the study question.

→ Study sites/countries to define which comparisons will be carried out

Crude vaccine effectiveness

Each individual will enter the study in the different vaccination status groups on the date they are first classified into that group. This will be the date of the beginning of the study, except for individuals that change vaccination status groups throughout follow-up, which will be censored without event in the group that they leave and are recorded as a delayed entry in the group where they are newly classified in (as in the Example Tables 2 and 3). End of follow-up will be established at the time of occurrence of any reason for censoring and will be marked as event=1 if the reason for censoring is the event of interest, or event=0 otherwise.

Vaccine effectiveness will be estimated using hazard ratio (HR) or rate ratio (RR) of defined outcome(s) in individuals with different vaccination status categories, as defined above, within the population study [18].

Crude VE = (1-HR) x 100

or

Crude VE = (1-RR) x 100

Survival Cox regression models for the estimation of HRs will be fit with calendar time as the underlying time scale, thus assigning time 0 to the first day of the observation period. If rates-based Poisson regression models are used for the estimation of RRs, the calendar time must be accounted for by adding a covariate with the calendar week (thus splitting the follow-up time of every person into as many lines as calendar weeks as covered by the time-at risk) and adjusting the model by this calendar week variable.

→ Study sites to define the analytical approach

Adjusted vaccine effectiveness

The regression analysis to estimate HR or RR will be adjusted for fixed or time-changing confounders, as appropriate, and as previously defined. First, partially adjusted HR or RR will be estimated, adjusting by age group (5 year-bins), sex and region in the country, if appropriate. Second, a fully adjusted HR or RR estimate will be produced adjusting by variables related to previous infection, socioeconomic condition, comorbidities and health-seeking behaviour, as relevant at each study-site. To help select additional variables to adjust the models at each site, it is recommended to fit models with and without each individual covariate, to see the effect that it has on the estimate of VE.

Methods for pooling estimates

Country-specific HRs/RRs and standard errors for the effect of COVID-19 vaccination obtained from the study sites will be combined in a model using meta-analysis techniques. Study sites will not report VE estimates for which the number of events is less than five, and those will not be included in the pooled estimates [19].

First, a fixed-effects approach will be used, computing a simple weighted average across studies. To do this, the site-specific vaccination status-disease effects (HRs/RRs) will be weighted by the inverse of their marginal variances (generic inverse variance method). This will give the pooled HR/RRs and a standard error. Confidence interval around the pooled effect (the range of values that contain the true average HR/RR with 95% certainty) will then be calculated.

Second, a random-effects approach will be used. Conceptually, it is possible that VE is different depending on measured or unmeasured site-specific factors. To account for the two sources of variability (intra-study and between-studies), the marginal variance is divided into two components: the individual study-specific variances and the variance of the random study effects (τ^2). I^2 represents the proportion of the total variance that is attributable to the random study effects. τ^2 and I^2 will be used to report between-studies statistical heterogeneity, along with the p-value of the heterogeneity test.

Potential factors or specific pilot sites characteristics that could be the source of qualitative heterogeneity will be described, as covered in the descriptive part of the data analysis in this protocol.

The country-specific HR/RRs and their confidence intervals, along with the pooled HR/RRs, will be presented graphically in a forest plot. The crude effect, the basic adjusted effect (age, sex, region) and the fully-adjusted effect (adding the rest of available covariates), will be compared to assess the degree of confounding by different factors and guide the hypothesis around the explanation of differences across sites.

Sensitivity analyses will be conducted for pooled estimates obtained while excluding some study participants for whom variables were collected, defined or managed differently, or who have differences in the study setting that could affect the estimates (e.g. different SARS-CoV-2 genetic variant with predominant circulation) or for whom estimates significantly differ across sites (i.e. site confidence intervals do not overlap with the pooled estimate confidence interval), particularly if the I^2 estimate is >50%.

Data checking and validation

The following data checking and data validation will be undertaken before analysis:

- Identification of inconsistencies (e.g. earlier dates for second doses than for first doses).
- Unusual values and outliers.
- Inclusion/exclusion criteria adherence.
- Missing values, missing clinical details, missing laboratory results.
- Duplicate cases and multiple admissions.
- Too long delay from date of onset of symptoms to lab specimen collection date.
- Consistency of and among dates (onset, admission, discharge, swabbing).
- Missing data for essential variables that can lead to excluding of the records from the analyses.

→ *Study sites/countries to list data checking and validation items*

Ethical requirements

Approval by an ethics committee is a requirement. All sites must conform with national and EU ethical and data protection requirements.

→ *Study sites to provide information on ethical approval*

Potential biases and limitations

- The availability and degree of detail in the study variables will be limited by what is recorded in registries whose final aim was different than performing this study. Besides, if the registry was not established with the aim of supporting VE studies, there might be an issue regarding consent. Moreover, a high degree of variability between countries is expected, making the pooled results more difficult to interpret as a homogeneous estimation.
- Control for confounding will also be limited since many relevant variables are not monitored or are not available in the data. Previous infections as a relevant cofounder may not be properly recorded and, therefore, may not be controlled for due to limitations in the availability of data and different testing policies among participating countries.

- Given the staggered rollout of COVID-19 vaccination in most European countries, there is collinearity between age, time of infection, and type of vaccine used. Therefore, the interpretation of effects by time since vaccination and by age needs to be done with caution from a causal perspective.
- The non-vaccinated individuals will be the reference group in the analysis. However, they are increasingly different from the vaccinated population. For example, they could take less preventive precautions and be more at risk and have an underlying higher risk of COVID-19, overestimating VE. On the other hand, they may be individuals with previous infection that we failed to identify in our registries (something likely after the widespread use of self-tests), thus we would underestimate VE.
- Many countries have stopped exhaustively registering COVID-19 diagnoses during the Omicron variant-dominated pandemic wave, and many are ending the systematic testing of all mild COVID-19 cases. This can make estimates against infection and/or symptomatic infection unreliable.

→ Study sites to describe the main limitations of their study and how those can affect the results

Data sharing for the purpose of pooling results at a European level

The specific information to be shared within the European collaboration, should be included in the Excel files (Annex 4 and 5). They include:

- Point estimation (HR/RR) for all analysed effects (for each outcome, in every subgroup), i.e. exponential coefficient from the Cox or Poisson regression.
- Lower and upper limits of the confidence interval of every estimation, i.e. exponential lower and upper limits of the confidence intervals estimated from the Cox or Poisson regression.
- Number of persons, of person-days of follow-up and of events in each analysis.

References

1. European Centre for Disease Prevention and Control (ECDC). Pilot protocol for a COVID-19 vaccine effectiveness study using health data registries. Stockholm: ECDC; 2022. Available at: https://www.ecdc.europa.eu/sites/default/files/documents/covid-19-pilot-protocol-vaccine-effectiveness-study.pdf
2. Sentís A, Kislaya I, Nicolay N, Meijerink H, Starrfelt J, Martínez-Baz I et al., the VEBIS-Lot4 working group. Estimation of COVID-19 vaccine effectiveness against hospitalisation in individuals aged ≥ 65 years using electronic health registries; a pilot study in four EU/EEA countries, October 2021 to March 2022. Euro Surveill. 2022;27(30):pii=2200551. https://doi.org/10.2807/1560-7917.ES.2022.27.30.2200551
3. European Centre for Disease Prevention and Control (ECDC). Communicable Disease Threats Report Week 3, 16–22 January 2023. ECDC: Stockholm; 2023. Available from: https://www.ecdc.europa.eu/sites/default/files/documents/Weekly%20Communicable%20Disease%20Threats%20Report%20Public%202023w03.pdf
4. European Centre for Disease Prevention and Control (ECDC). COVID-19 vaccines [Internet]. Sweden: ECDC; Available from: https://www.ema.europa.eu/en/human-regulatory/overview/public-health-threats/coronavirus-disease-covid-19/treatments-vaccines/covid-19-vaccines
5. EMA and ECDC join forces for enhanced post-marketing monitoring of COVID-19 vaccines in Europe. Press Release. 26 April 2022. https://www.ecdc.europa.eu/en/news-events/ema-and-ecdc-join-forces-enhanced-post-marketing-monitoring-covid-19-vaccines-europe
6. COUNCIL RECOMMENDATION of 7 December 2018 on strengthened cooperation against vaccine-preventable diseases (2018/C466/01), available at https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018H1228(01)
7. REGULATION (EU) 2022/123 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 25 January 2022 on a reinforced role for the European Medicines Agency in crisis preparedness and management for medicinal products and medical devices, Available from : https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2022:020:FULL&from=EN
8. REGULATION (EU) 2022/2370 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 23 November 2022 amending Regulation (EC) No 851/2004 establishing a European centre for disease prevention and control, available from: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32022R2370&qid=1670582843806&from=EN
9. European Centre for Disease Prevention and Control (ECDC). Core protocol for ECDC studies of COVID-19 vaccine effectiveness against hospitalisation with Severe Acute Respiratory Infection laboratory-confirmed with SARS-CoV-2, version 1.0. Stockholm: ECDC; 2021. Available from: https://www.ecdc.europa.eu/en/publications-data/core-protocol-ecdc-studies-covid-19-vaccine-effectiveness-against-hospitalisation
10. European Centre for Disease Prevention and Control (ECDC). Interim analysis of COVID-19 vaccine effectiveness against Severe Acute Respiratory Infection due to laboratory-confirmed SARS-CoV-2 among individuals aged 65 years and older, ECDC multi-country study. ECDC: Stockholm; 2021. 16vailable at: https://www.ecdc.europa.eu/sites/default/files/documents/TRP-20211004-1860_0.pdf
11. European Centre for Disease Prevention and Control (ECDC). Interim analysis of COVID-19 vaccine effectiveness against Severe Acute Respiratory Infection due to laboratory-confirmed SARS-CoV-2 among individuals aged 50 years and older, ECDC multi-country study – first update. ECDC: Stockholm; 2022. Available at : https://www.ecdc.europa.eu/sites/default/files/documents/Interim-analysis-of-COVID-19-vaccine-effectiveness-against-SARI-2022.pdf
12. European Centre for Disease Prevention and Control (ECDC). Interim analysis of COVID-19 vaccine effectiveness against Severe Acute Respiratory Infection due to laboratory-confirmed SARS-CoV-2 among individuals aged 30 years and older, ECDC multi-country study – second update. ECDC: Stockholm; 2022. Available at: https://www.ecdc.europa.eu/sites/default/files/documents/covid-19-vaccine-effectiveness-interim-analysis-second-update_0.pdf
13. European Centre for Disease Prevention and Control (ECDC). Interim analysis of COVID-19 vaccine effectiveness against Severe Acute Respiratory Infection due to SARS-CoV-2 in individuals aged 20 years and older – third update. ECDC: Stockholm; 2022. Available at: https://www.ecdc.europa.eu/sites/default/files/documents/Interim-analysis-COVID-19-vaccine-over-20.pdf
14. European Centre for Disease Prevention and Control (ECDC). Generic protocol for ECDC studies of COVID-19 vaccine effectiveness against confirmed SARS-CoV-2 using healthcare worker cohorts, v.2.0. Stockholm: ECDC; 2022. Available at : https://www.ecdc.europa.eu/sites/default/files/documents/covid-19-generic-protocol-ECDC-studies-vaccine-effectiveness-healthcare-worker-cohorts-version-2-0.pdf

15 European Centre for Disease Prevention and Control (ECDC). Interim analysis of COVID-19 vaccine effectiveness in healthcare workers, an ECDC multi-country study, May 2021–July 2022. ECDC: Stockholm; 2022. Available at : https://www.ecdc.europa.eu/sites/default/files/documents/covid-19-analysis-vaccine-effectiveness-healthcare-workers.pdf

16. Kissling E, Hooiveld M, Martínez-Baz I, Mazagatos C, William N, Vilcu A-M, , et al. Effectiveness of complete primary vaccination against COVID-19 at primary care and community level during predominant Delta circulation in Europe: multicentre analysis, I-MOVE-COVID-19 and ECDC networks, July to August 2021. Euro Surveill. 2022;27(21):pii=2101104. https://doi.org/10.2807/1560-7917.ES.2022.27.21.2101104

17. Ribeiro AI, Launay L, Guillaume E, Launoy G, Barros H. The Portuguese version of the European Deprivation Index: Development and association with all-cause mortality; 2018. PLoS ONE 13(12): e0208320. Available at: https://doi.org/10.1371/journal.pone.0208320

18. Nunes B, Rodrigues AP, Kislaya I, Cruz C, Peralta-Santos A, Lima J, et al. mRNA vaccine effectiveness against COVID-19-related hospitalisations and deaths in older adults: a cohort study based on data linkage of national health registries in Portugal, February to August 2021. Euro Surveill. 2021 Sep;26(38):2100833. Available at: https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2021.26.38.2100833

19. Chapter 10: Analysing data and undertaking meta-analyses. Cochrane Training. Available at: https://training.cochrane.org/handbook/archive/v6.2/chapter-10#section-10-3

Annex 1. COVID-19 vaccines originally authorised and in use in the EU/EEA as of December 2022

Vaccine	Age of recommendation	Primary Dose regimen	Booster dose
Vaxzevria (AZD1222)	18 years +	Two doses given between 4 and 12 weeks after the first dose.	A booster dose may be given at least 3 months after the second dose.
Comirnaty (BNT162b2) 30 micrograms per dose (12 years +) 10 micrograms per dose (5-11 years) 3 micrograms per dose (6 months – 4 years)	6 months +	Two doses 3 weeks apart in individuals aged 5 years + Three doses. In children from 6 months to 4 years of age, the first two doses are given three weeks apart, followed by a third dose given at least 8 weeks after the second dose.	May be given as a booster given at least 3 months after the second dose in individuals aged 12 years + May be given as a booster given at least 6 months after the second dose of primary vaccination in individuals aged 5-11 years.
Comirnaty Original/Omicron BA.1	12 years +	No	May be given as a booster given at least 3 months after the second dose in individuals aged 12 years + May be given as a booster given at least 6 months after the second dose of primary vaccination in individuals aged 5-11 years May be given as a second booster given in 12 years + 3 months after first booster
Comirnaty Original/Omicron BA.4-5	5 years +	May be given for primary vaccination in non-immune individuals	May be given at least 3 months after primary vaccination or a booster dose with a COVID-19 vaccine in 12 years + May be given as a booster dose given 3 months after primary vaccination or first booster in 5 years +
Jcovden (Ad26.COV 2.5)	18 years +	Single dose	May be used as a booster dose given at least 2 months after the first dose of Jcovden
Spikevax (mRNA-1273) 50 micrograms per dose (12 years +) 25 micrograms per dose (6 months- 5 years)	6 months +	2 injections 28 days apart	May be given as a booster given at least 3 months after the second dose in individuals aged 12 years +
Spikevax bivalent Original/Omicron BA.1 50 micrograms per dose (12 years +) 25 micrograms per dose (6 months- 5 years)	6 years +	No	May be given to adults and children from the age of 6 years, at least 3 months after primary vaccination or a booster dose with a COVID-19 vaccine
Spikevax bivalent Original/Omicron BA.4-5,	12 years +	No	May be given to adults and children from the age of 6 years, at least 3 months after primary vaccination or a booster dose with a COVID-19 vaccine
Nuvaxovid (NVX-CoV2373)	12 years +	Two doses given 3-weeks apart	May be used as a booster in adults who have had Nuvaxovid, an mRNA vaccine or an adenoviral vector vaccine as their primary vaccination
VidPrevtyn Beta	18 years +	No	May be used as a booster dose at least 4 months after a previous mRNA or adenoviral vector COVID-19 vaccine.
COVID-19 Vaccine Valneva (inactivated, adjuvanted) (VLA2001)	18 to 50 years	Two doses given 4 weeks apart	No

Annex 2. Type of studies and study settings within the VEBIS infrastructure as of November 2022

Setting	Type of study	Main outcome	References
Hospitals	Test negative design	Severe disease; influenza and COVID-19	[9-13]
Health care workers cohort	Cohort study	Infection, COVID-19	[14,15]
Electronic health care databases	Cohort study	Hospitalisation and other severe outcomes, COVID-19	[7,8]
Primary care	Test negative design	Moderate disease (~ARI/ILI), Influenza and COVID-19	[16]

Annex 3. Data dictionary (example)

Individual characteristics

	Variable	Type	Coding	Definition
Operational	Extract date	Date	dd/mm/yyyy	Database extraction date
Patient characteristics	Sex	Numeric	0 = female	Sex of patient
			1 = male	
			3 = other	
			8 = do not know	
	dob	Date	dd/mm/yyyy	Date of birth (only if no age; once age calculated from dob this will be dropped)
	postcode	Numeric		Postcode of residence
	residence	Numeric	0 = at home, not dependent on home support/care	Patient residence at time of event onset. Whether patient was living at home or was institutionalised, or had pre-hospital dependence on home support/care
			1 = at home, but dependent on home support/care	
			2 = institutionalised	
			3 = Do not know	

Outcome

	Variable	Type	Coding	Definition
COVID-19 case	swabdate	Date	dd/mm/yyyy	Respiratory specimen collection date
	lab_covtest	Numeric (categorical)	0 = No	Tested for SARS-CoV-2
			1 = Yes	
			8 = Do not know	
	lab_covtesttype	Numeric (categorical)	1 = RT-PCR	Type of lab test used
			2 = Serology	
			3 = Rapid test	
			4 = Other	
			8 = Do not know	
	lab_covtesttype_sp	Text		Specify other type of lab test
	lab_covid	Numeric (categorical)	0 = Negative	Laboratory result: virus type SARS-CoV-2
			1 = Positive	
			8 = Do not know	
			1 = Positive	
			8 = Do not know	
Hospital / ward information	prevhosp	Numeric (categorical)	0 = No	Prior admission to hospital (at least once in previous 12 months)
			1 = Yes	
			8 = Do not know	
	admitdate	Date	dd/mm/yyyy	Date of hospital admission
	hospitalward	Text		Ward
	dischargedate	Date	dd/mm/yyyy	Date of hospital discharge
COVID_icu	icu		0 = No	Admission to intensive care unit (ICU)
			1 = Yes	
			8 = Do not know	
	icuadmitdate	Date	dd/mm/yyyy	Date first admitted ICU
	icudisdate	Date	dd/mm/yyyy	Date last discharged from ICU
Death	death	Numeric (categorical)	0 = No	Person is deceased
			1 = Yes	
			8 = Do not know	
	deathdate	Date	dd/mm/yyyy	Date of death
	deathcause	Numeric (categorical)	1 = died from COVID-19	Cause of death
			2 = died other cause	
			8 = died unknown cause	

Vaccination status

	Variable	Type	Coding	Definition
COVID-19 vaccination	panvaccany	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Received at least one dose of COVID-19 vaccine
	Panvaccdate_i	Date	dd/mm/yyyy	Vaccination date (for each dose, i)
	Panvacctype_i	Text		Type of vaccine (for each dose, i)
	panvaccdose	Numeric	0, 1, 2	Number of doses received

Comorbidities

	Variable	Type	Coding	Definition
Underlying chronic conditions	Anaemia	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Anaemia/chronic haematologic disease
	Asplenia	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Asplenia (absence of/damage to spleen)
	Asthma	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Asthma
	Cancer	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Cancer (any)
	Hypert	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Hypertension
	Demente	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Dementia
	Diabetes	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Diabetes
	Heartdis	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Heart / cardiac disease (excluding hypertension)
	Immuno	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	HIV or other immunodeficiency or organ transplantation
	Liverdis	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Chronic liver disease (excluding cancer)
	Lungdis	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Lung disease (excluding asthma)
	Neuromusc	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Neuromuscular disorder
	Height	Numeric (integer)		Height of patient in metres
	Weight	Numeric (integer)		Weight of patient in kg
	BMI	Numeric (1 d.p)		BMI of patient (calculated using data collected on height and weight)
	Obese	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Obesity (only if height, weight and BMI not collected; can be calculated)
	Rendis	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Renal disease (excluding cancer and acute renal failure)
	Rheumat	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Rheumatologic disease

	Variable	Type	Coding	Definition
	Stroke	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Stroke
	Tuberc	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Tuberculosis

Other confounding variables

	Variable	Type	Coding	Definition
Patient characteristics	Frailty	Numeric (categorical)	To be updated with coding depending on score used	Clinical frailty score at admission (where possible) or Barthel Index
Pre-symptomatic treatment/intervention: vaccination	flu_vacc	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Received current seasonal influenza vaccination
	flu_vaccdate	Date	dd/mm/yyyy	Date of last influenza vaccination
	ppv_vacc	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Received PPV23 vaccination
	ppv_vaccdate	Date	dd/mm/yyyy	Date of last PPV23 vaccination
	pcv_vacc	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Received PCV7/10 or 13 vaccination
	pcv_vaccdate	Date	dd/mm/yyyy	Date of last PCV7/10 or 13 vaccination
	bcg_vacc	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Received BCG vaccination
	bcg_vaccyear	Numeric	yyyy	Year of BCG vaccination

Annex 4. Format for periodic reporting

(Example, the full periodic reporting tables are available in excel format upon request to ECDC)

Exposure categories	N*	person-days	Events	HR crude			HR adjusted1**			HR adjusted2***		
				Estimate	95%CI low	95%CI high	Estimate	95%CI low	95%CI high	Estimate	95%CI low	95%CI high
Non vaccinated				REF	REF	REF	REF	REF	REF	REF	REF	REF
Complete vaccination												
Complete vaccination + booster 1 (overall)												
Complete vaccination + booster 2 (overall)												
Complete vaccination + booster 3 (overall)												
Non vaccinated				REF	REF	REF	REF	REF	REF	REF	REF	REF
Complete vaccination + booster 1, ≤84 days after time 0												
Complete vaccination + booster 1, days 85 – 168 after time 0												
Complete vaccination + booster 1, ≥169 days after time 0												
Complete vaccination + booster 2, ≤84 days after time 0												
Complete vaccination + booster 2, days 85 – 168 after time 0												
Complete vaccination + booster 2, ≥169 days after time 0												
Complete vaccination + booster 3, ≤84 days after time 0												
Complete vaccination + booster 3, days 85 – 168 after time 0												
Complete vaccination + booster 3, ≥169 days after time 0												
Completely vaccinated (≥169 days)				REF	REF	REF	REF	REF	REF	REF	REF	REF
Complete vaccination + booster 1 (overall)												
Complete vaccination + booster 2 (overall)												
Complete vaccination + booster 3 (overall)												
Completely vaccinated (≥169 days)				REF	REF	REF	REF	REF	REF	REF	REF	REF
Complete vaccination + booster 1, ≤84 days after time 0												
Complete vaccination + booster 1, days 85 – 168 after time 0												
Complete vaccination + booster 1, ≥169 days after time 0												
Complete vaccination + booster 2, ≤84 days after time 0												
Complete vaccination + booster 2, days 85 – 168 after time 0												
Complete vaccination + booster 2, ≥169 days after time 0												
Complete vaccination + booster 3, ≤84 days after time 0												
Complete vaccination + booster 3, days 85 – 168 after time 0												
Complete vaccination + booster 3, ≥169 days after time 0												
Complete vaccination + booster 1 (≥90 days)				REF	REF	REF	REF	REF	REF	REF	REF	REF
Complete vaccination + booster 2 (overall)												
Complete vaccination + booster 2, ≤84 days after time 0												
Complete vaccination + booster 2, days 85 – 168 after time 0												
Complete vaccination + booster 2, ≥169 days after time 0												
Complete vaccination + booster 2 (≥90 days)				REF	REF	REF	REF	REF	REF	REF	REF	REF
Complete vaccination + booster 3 (overall)												
Complete vaccination + booster 3, ≤84 days after time 0												
Complete vaccination + booster 3, days 85 – 168 after time 0												
Complete vaccination + booster 3, ≥169 days after time 0												

** Number of individuals contributing to each group. Because exposure is time changing, the sum of N in all categories will be greater than the total sample size in the study*

*** HR adjusted1: Adjusted by age (5-year bins), sex and region accoding to country-specific protocol*

**** HR adjusted2: Additionally adjusted by the rest of confounding variables accoding to country-specific protocol*

Annex 5. Study sites methods overview

Overview of methods

Outcomes analysed	☐ Laboratory-confirmed SARS-CoV-2 infection ☐ Hospital admission due to COVID-19 ☐ ICU admission due to COVID-19 ☐ Death due to COVID-19
Secondary objectives that will be included	☐ By age-group ☐ In people with previous infection ☐ Relative effectiveness (i.e. between vaccinated in different categories of vaccination) ☐ By time since booster ☐ In people primed with mRNA
Which identifier is being used for data-linkage?	
Definition used for symptomatic infection	
Definition used for COVID-19 hospitalization	
Definition used for ICU admission due to COVID-19	
Definition used for COVID-19 death	
Definition used for the group with previous infection	
Method used to analyse data	☐ Incidence Rates and Poisson regression ☐ Survival analysis and Cox regression
Software used for the analysis	
Comments (challenges, limitations or other)	

Information of variables and data sources

(Example table, add or remove rows as needed)

Group of variables	Variables (e.g. vaccination date, region, chronic disease)	Data source	Limitations, other comments
Outcome	Laboratory-confirmed SARS-CoV-2 infection		
	Hospital admission due to COVID-19		
	ICU admission due to COVID-19		
	Death due to COVID-19		
Vaccination status	Vaccination status		
	Time since vaccination		
	Vaccine brand		
Stratification	By age group		
	Previous infection		
	Add others as appropriate		
Adjustment	Age		
	Sex		
	Health Region		
	Comorbidities		
	Add others as appropriate		

Exact definition and categorisation used in adjustment/stratification variables

Variable	Definition, categorisation, use in the model
Age groups	
Comorbidities	
Country of residence / country of birth / nationality	
Education level	
Deprivation index or similar	
Time / calendar	
Region / Health region / Geographic level	
Other vaccines uptake in the last four years (e.g., influenza, PCV7, PCV10, PCV13, PPV23)	
Number of COVID-19 tests in 2020-2022	
Other variables	

PDF TQ-03-23-039-EN-N

European Centre for Disease
Prevention and Control (ECDC)

Gustav III:s Boulevard 40, 16973 Solna, Sweden

Tel. +46 858601000
Fax +46 858601001
www.ecdc.europa.eu

An agency of the European Union
www.europa.eu

Subscribe to our publications
www.ecdc.europa.eu/en/publications

Contact us
publications@ecdc.europa.eu

Follow us on Twitter
@ECDC_EU

Like our Facebook page
www.facebook.com/ECDC.EU

ISBN 978-92-9498-614-6