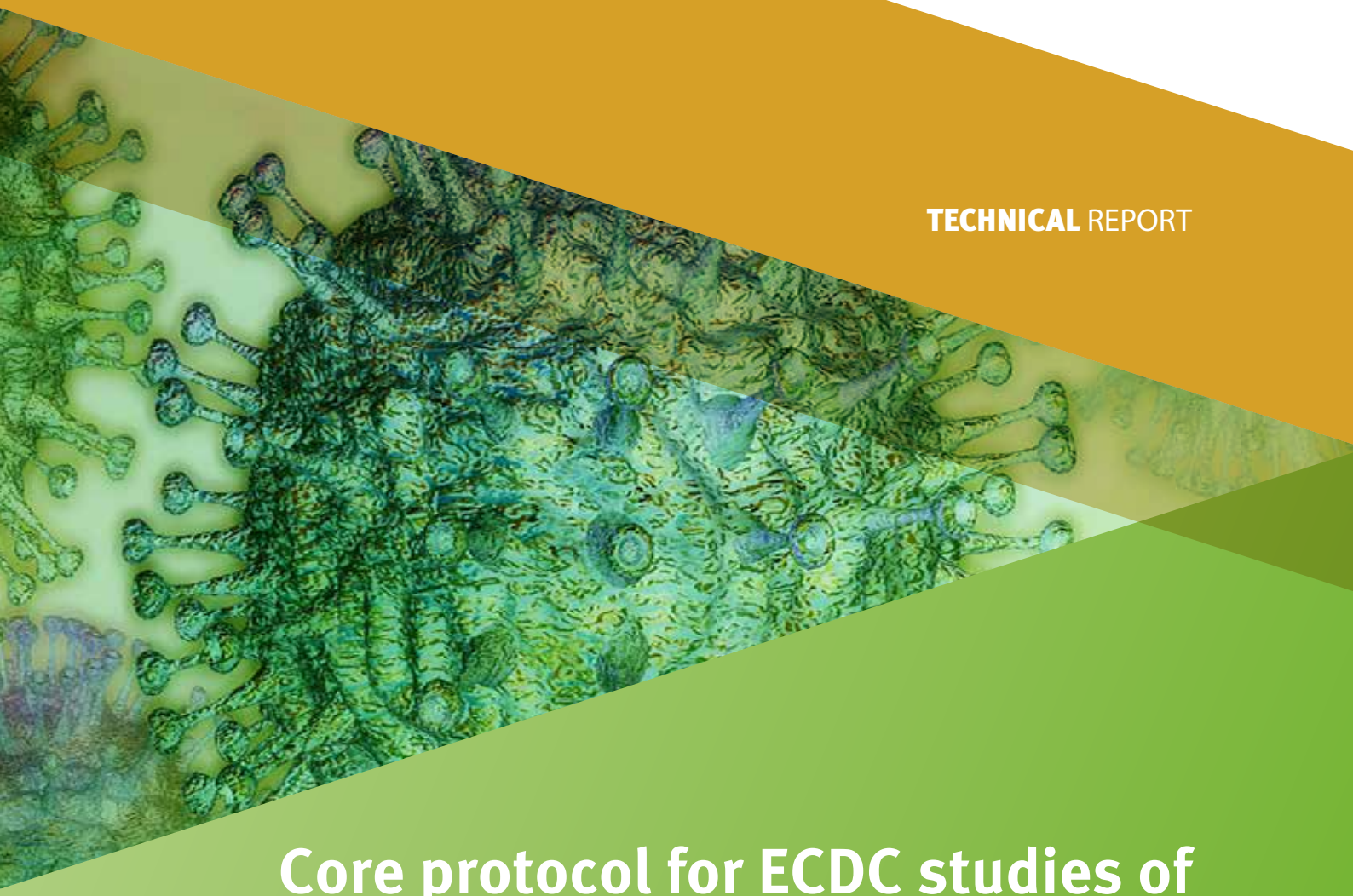


**TECHNICAL** REPORT

# 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

www.ecdc.europa.eu

**ECDC** TECHNICAL REPORT

# 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

This core protocol was commissioned by the European Centre for Disease Prevention and Control (ECDC), as part of the activities referring to Direct Contract N. ECD.11486 'Developing an infrastructure and performing vaccine effectiveness studies for COVID-19 vaccines in the EU/EEA', coordinated by Sabrina Bacci and Christiana Carstairs.

*Authors (alphabetical order)*

Esther Kissling, Alain Moren, Anthony Nardone, Angie Rose, Marta Valenciano (Epiconcept)

*Acknowledgements*
The following experts either participated in the first ECDC Expert meeting on COVID-19 vaccine effectiveness studies on 29 January 2021 or received the protocol for consultation:

Nathalie Bossuyt (Sciensano, Belgium), Natalie Fisher (Sciensano, Belgium), Sarah Denayer (Sciensano, Belgium), Isabelle Thomas (Sciensano, Belgium), Lucie Seyler (Universitair Ziekenhuis Brussel, UZ, Belgium), Thomas Demuyser (UZ, Belgium), Virgini Van Buggenhout (UZ, Belgium), Goranka Petrovic (Croatian Institute of Public Health (Hrvatski zavod za javno zdravstvo, HZJZ), Croatia), Zvjezdana Lovrić (HZJZ, Croatia), Irena Tabain (HZJZ, Croatia), Petr Husa (University Hospital Brno, Czechia), Hana Orliková (National Institute of Public Health, NIPH, Czechia), Klára Labská (NIPH, Czechia), Hanna Sepp (Health Board, Estonia), Anneli Uuskula (Health Board, Estonia), Hanna Nohynek (Finnish Institute for Health and Welfare, THL, Finland), Alessandro Pini (Santé publique, France), Daniel Levy-Bruhl (Santé publique, France), Anna Maisa (Santé Publique, France), Weibke Hellenbrand (Robert Koch Institute, RKI, Germany), Buda, Silke (RKI, Germany), Stella Michelaki (National Public Health Organization, EODY, Greece), Marina Amerali (EODY, Greece), Suzanne Cotter (Health Service Executive, HSE, Ireland), Lisa Domegan (HSE, Ireland), Melissa Brady (HSE, Ireland), Naomi Petty-Saphon (Health Protection Surveillance Centre, HPSC, Ireland), Matt Robinson (HPSC, Ireland); Françoise Berthet (National Health Directorate, Luxembourg), Myriam Alexandre (Luxembourg Institute of Health, LIH, Luxembourg), Nassera Aouali (LIH, Luxembourg), Manon Gantenbie (National Health Directorate, Luxembourg), Guy Fagherazzi (National Health Directorate, Luxembourg), Joaquin Baruch (Ministry of Health, Malta), Maria-Louise Borg (Ministry of Health, Malta), Adam Meijer (National Institute for Public Health and the Environment, RIVM, the Netherlands), Annabel Niessen (RIVM, the Netherlands), Mirjam Knoll (RIVM, the Netherlands), Pedro Pinto Leite (Direção-Geral da Saúde, DGS, Portugal), Ausenda Machado (Instituto Nacional de Saúde Doutor, INSA, Portugal), Vânia Isabel da Silva Gaio (INSA, Portugal), Raquel Guiomar (INSA, Portugal), Amparo Larrauri (Institute of Health Carlos III, ISCIII, Spain), Clara Mazagatos (ISCIII, Spain), Carmen Olmedo (Ministry of Health, MOH, Spain), Susana Monge (MOH, Spain), Antonio Rezusta, (Unizar, Spain), Miriam Latorre (Hospital Miguel Servet, Zaragoza, Spain), Nick Jewell (London School of Hygiene and Tropical Medicine, United Kingdom), Mark Page (National Institute for Biological Standards and Control, United Kingdom), Diogo Marques (Epiconcept), Gabrielle Breugelmans (Coalition for Epidemic Preparedness Innovations), Miwako Kobayashi (Centers for Disease Control and Prevention, CDC, United States), Mark Tenforde (CDC, United States), Jennifer Verani (CDC, United States), Richard Pebody (World Health Organization Regional Office for Europe), Francisco Nogareda (Pan American Health Organization), Rita Figueira (European Commission), Manuela Mura (European Medicines Agency, EMA), Vanessa Seguin (EMA), Catherine Cohet (EMA).

The following ECDC staff provided substantial input during the production/revision of the document (alphabetical order): Sabrina Bacci, Piotr Kramarz, and Nathalie Nicolay.

This core protocol is based on: Scientific literature; I-MOVE generic influenza protocol for hospitalised older adults 2019–2020; I-MOVE-COVID-19 WP4 protocol for investigation into risk factors against severe COVID-19 among hospitalised patients; Guidance document from the WHO Regional Office for Europe: COVID-19 VE against SARI hospitalisations associated with lab-confirmed SARS-CoV-1 version 5 (20 January 2021), first ECDC Expert meeting on COVID-19 vaccine effectiveness studies (29 January 2021) SARI COVID-19 Vaccine Effectiveness questionnaire v7 from the WHO Regional Office for Europe (03 February 2021), I-MOVE-COVID-19 protocol: European study of COVID-19 vaccine effectiveness against hospitalised SARI patients laboratory confirmed with SARS-CoV-2 (version 8 February 2021).

The current version 1.0 of this core protocol corresponds to version 4.0 used to implement the Direct Contract ECD.11486.

Suggested citation: European Centre for Disease Prevention and Control. 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.

Stockholm, October 2021

ISBN: 978-92-9498-546-0
doi: 10.2900/991808
Catalogue number TQ-08-21-303-EN-N

# Contents

# Abbreviations

| | |
|---|---|
| COVID-19 | Coronavirus disease 2019 |
| Ct | Cycle threshold |
| CVE | COVID-19 vaccine effectiveness |
| EEA | European Economic Area |
| EU | European Union |
| GP | General practitioner |
| ICD | International classification of diseases |
| I-MOVE | Influenza – Monitoring Vaccine Effectiveness in Europe |
| MS | Member States |
| OR | Odds ratio |
| RF | Risk factor |
| RT- PCR | Reverse-transcription polymerase chain reaction |
| SARI | Severe acute respiratory infection |
| SARS-CoV-2 | Severe acute respiratory syndrome – coronavirus 2 |
| SES | Socioeconomic status |
| VE | Vaccine effectiveness |

# Executive summary

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). As of week 38 2021, over 37 million cases and more than more than 750 000 deaths had been reported in the European Union/European Economic Area (EU/EEA) [1].

As of week 38 2021, four vaccines (Comirnaty, Spikevax [previously COVID-19 vaccine Moderna], Vaxzevria, and Janssen) 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 [2].

In 2020, the European Commission emphasised the importance of continuously monitoring the safety and effectiveness of vaccines in the EU/EEA and called on ECDC and EMA to develop a structured post-authorisation monitoring platform for vaccines, prioritising COVID-19 vaccines. In November 2020, the European Commission proposed to the European Parliament and the Council of the EU a change to the mandates of EMA and ECDC in the context of its COVID-19 lessons learned package and the creation of a European Health Union, empowering the two agencies to jointly coordinate independent vaccine monitoring studies.

As a result, at the end of 2020, utilising the lessons learned from other vaccine effectiveness studies, ECDC started building infrastructure to perform COVID-19 vaccine effectiveness studies. The infrastructure aims to build a system to regularly monitor vaccine effectiveness and perform studies in different settings, and depending on the setting, to provide information on different outcomes (severe disease, moderate disease, transmission, etc).

This 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, presents the main elements for a multicentre (multi-country) hospital-based study of COVID-19 vaccine effectiveness in patients hospitalised with Severe Acute Respiratory Infection, outlining the agreed methods for collecting data related to COVID-19 and SARS-CoV-2 at country level, and includes a plan for the pooled analysis. The combination of data from multiple sites will allow for studies with more statistical power to meet more specific objectives.

This core protocol, therefore, is primarily intended to guide the implementation of ECDC-funded studies. However, ECDC would like to encourage the use of this protocol as a basis for other vaccine effectiveness studies in countries/hospitals that do not currently plan to participate in ECDC-funded studies. The use of consistent protocols will facilitate the comparability of study results across studies, countries, and study sites.

This document presents version 1.0 of the core protocol, which is planned to be updated and revised on a regular basis and subsequently implemented at country level.

# 1 Background

## 1.1 Context

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). As of week 38 2021, over 37 million cases and more than more than 750 000 deaths had been reported in the EU/EEA [1].

Since the beginning of the vaccination rollout in the EU and as of week 38, four COVID-19 vaccines have received conditional marketing authorisation by the European Commission (EC) based on the scientific opinion of the European Medicine Agency (EMA) [2]: Comirnaty (BNT162b2) Spikevax (mRNA-1273), Vaxzevria (AZD1222), and COVID-19 Vaccine Janssen (Ad26.COV 2.5). All vaccine products authorised in the EU were initially registered for use in people aged 18 years and above, with the exception of Comirnaty, which was approved for use in people aged 16 years and above. Corminarty and Spikevax indications were recently extended to include children aged 12 to 15 years and 12 to 17 years, respectively [3] .

By January 2021, all 30 EU/EEA countries had started COVID-19 vaccination campaigns, and different COVID-19 vaccine products have been gradually introduced as they became available through the EU Vaccines Strategy.

Efficacy measured in vaccine trials under experimental conditions may differ from effectiveness measured in real life in a more diverse population under different conditions. Real life conditions include, in addition to different populations than those recruited for the clinical trials/pre-authorisation phase, assessment of the vaccine effectiveness with incomplete schedules, mixed vaccine products, and varying proportions of SARS-CoV-2 variants circulating. Studies on the effectiveness and impact of COVID-19 vaccines and vaccination programmes will be critical to understanding how the vaccines will affect morbidity and mortality against several different outcomes, and so determine their value and scope as a public health intervention in the context of the pandemic and beyond it. Additionally, the possible emergence of more transmittable variants of concern that could partially escape the immunity induced by the vaccination makes it imperative to monitor vaccine effectiveness over time.

## 1.2 ECDC COVID-19 vaccine effectiveness studies

In 2018, the Council recommendation on Strengthened Cooperation against Vaccine-preventable Diseases (2018/C 466/01) called on the European Commission to work with the Member States and with the support of the European Medicines Agency (EMA) and in cooperation with ECDC to 'continuously monitor the benefits and risks of vaccines and vaccinations at EU level including through post-marketing authorisation studies'.

In 2020, the European Commission emphasised the importance of continuously monitoring the safety and effectiveness of vaccines in EU/EEA and called on ECDC and EMA to develop a structured post-authorisation monitoring platform for vaccines, prioritising COVID-19 vaccines. In November 2020, the European Commission proposed to the European Parliament and the Council of the EU a change to EMA's and ECDC's mandates in the context of its COVID-19 lessons learned package and the creation of a European Health Union, empowering the two agencies to jointly coordinate independent vaccine monitoring studies.

As a result, at the end of 2020, utilising the lessons learned from other vaccine effectiveness studies, ECDC started building infrastructure to perform COVID-19 vaccine effectiveness studies. The infrastructure aims to build a system to regularly monitor vaccine effectiveness and perform studies in different settings, and depending on the setting, to provide information on different outcomes (severe disease, moderate disease, transmission, etc). The multi-country approach of the effectiveness studies is also one of the key features that characterizes the studies, with a foreseen approach of progressive inclusion of more countries over time.

One of the first studies implemented, and for which the core ECDC protocol is presented in this document, is a multi-country study aimed at estimating COVID-19 vaccine effectiveness (CVE) against severe disease, by assessing it in individuals hospitalised for Severe Acute Respiratory Infection (SARI).

## 1.3 Aim of the core protocol

This 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, presents the main elements for a multicentre (multi-country) hospital-based study of COVID-19 vaccine effectiveness in patients hospitalised with Severe Acute Respiratory Infection, outlining the agreed methods for collecting data related to COVID-19 and SARS-CoV-2 at country level, and includes a plan for the pooled analysis. The combination of data from multiple sites will allow for studies with more statistical power to meet more specific objectives.

With the final aim of putting in place a system for the regular monitoring of vaccine effectiveness, and taking advantage of the parallel efforts to enhance SARI surveillance in EU/EEA and countries of the Western Balkan

region[1], ECDC has worked closely with EU Member States to recruit hospitals capable of applying the core protocol and therefore contributing to the EU-level monitoring of COVID-19 vaccine effectiveness. Specifically, each study site has been identified through a process involving ECDC's relevant National Coordinator.[2]

This core protocol therefore is primarily intended to guide the implementation of ECDC-funded studies. However, ECDC would like to encourage the use of this protocol as a basis for other vaccine effectiveness studies in countries/hospitals that do not currently plan to participate in ECDC-funded studies. The use of consistent protocols will facilitate the comparability of study results across studies, countries and study sites.

This document presents version 1.0 of the core protocol, which is planned to be updated and revised on a regular basis and consequently implemented at country level.

This core protocol is complemented by a questionnaire, a list of variables to be collected and their coding, all of which are available upon request at vpd.vpd@ecdc.europa.eu.

Under each paragraph, arrow marks with italicised text indicate the points that countries/hospitals/study sites could further expand/detail when creating a country-specific protocol using the core ECDC protocol.

[1] Western Balkans, i.e. Albania, Bosnia and Herzegovina, Montenegro, Serbia, North Macedonia, and Kosovo (this designation is without prejudice to positions on status, and is in line with UNSCR 1244/1999 and the ICJ Opinion on the Kosovo declaration of independence).

[2] https://www.ecdc.europa.eu/sites/portal/files/media/en/aboutus/governance/competent-bodies/Documents/coordinating-competent-bodies-structures-terms-of-reference-and-interactions-w-Annexes.pdf

# 2 Objectives

## 2.1 Primary objective

The primary objective is to measure, within each European participating site/country and in a pooled, multi-country analysis, the direct effect (effectiveness) of overall and product-specific COVID-19 vaccines against SARI due to laboratory-confirmed SARS-CoV-2 in hospitalised patients, in order to provide up-to-date information on the ability of COVID-19 vaccines to prevent severe disease under real conditions of use.

## 2.2 Secondary objectives

The secondary objectives are:

- To measure overall and product-specific COVID-19 VE against against SARI due to laboratory-confirmed SARS-CoV-2 in hospitalised patients by participating study site/country, risk group (e.g. specific chronic conditions), sex, age group (18-49 years, 50-64 years, 65-79 years, 80+ years), COVID-19 vaccination target group, time since vaccination and regularly over calendar time, vaccine doses when applicable;
- To measure overall and product-specific COVID-19 VE among SARI patients requiring hospitalisation against specific genetic variant(s) of laboratory-confirmed SARS-CoV-2, more severe outcomes (ICU admission, invasive ventilation, in-hospital mortality); and
- To identify potential factors that may modify COVID-19 VE: prior SARS-CoV-2 infection, chronic conditions, the role of influenza vaccination, the role of settings such as long-term care facilities (LTCFs), the role of long-term medications (depending on availability of these data in the participating country).

These three secondary objectives aim at understanding the duration of protection of vaccine and identifying any differences in CVE among each of these strata, potential target groups for vaccination, and key SARS-CoV-2 virus phenotypic or genotypic evolutions that could affect vaccine performance.

- *Each study site/country to specify the (additional) secondary objectives of their study.*

# 3 Methods

## 3.1 Study design

- At study site level: hospital-based, test-negative, case–control study in each participating hospital
- At EU/EEA level: multicentre hospital-based, test-negative, case–control study, using pooled data from several countries/regions.

In addition to the test-negative controls, some sites/countries may be able to recruit a second additional control group (e.g. non-SARI hospitalised patients matched to cases by date of admission; see section 3.5.5).

- *Each study site/country to specify if, in addition to the test-negative controls, other control groups are selected.*

## 3.2 Study population

This study is intended to be conducted primarily in countries with pre-existing SARI surveillance systems, to recruit patients for hospital-based CVE studies. The study population for the CVE study will therefore consist of individuals of all ages, belonging to the target group for vaccination, hospitalised with SARI symptoms in one of the participating hospitals/services, with no contra-indication for COVID-19 vaccination.

- *Study sites/countries to describe the setting (number of hospitals included, number of beds, number and type of wards/specialties/services included).*
- *Study sites/countries to describe the existing SARI surveillance system in place.*
- *Study sites/countries to describe the study population.*
- *Study sites/countries to describe target group(s) for vaccination and order/timeline of vaccination by group (when known).*
- *Study sites/countries to describe the epidemiological situation (incidence, number of COVID-19 hospitalisations, mortality).*

## 3.3 Study period

The study period starts when the COVID-19 vaccine becomes available in each of the participating countries and when SARS-CoV-2 is circulating. The study period is defined initially for each priority vaccination group and begins for each vaccination group when the vaccination campaign in this group begins. For the general population, the study period will begin when the vaccination campaign is extended beyond target groups to the general population.

Participating hospitals carry out the study throughout the year.

- *Each study/country to define the beginning of the pandemic CVE study period (day/month/year).*
- *Each study site/country to specify the date of the start of their vaccination campaign, by target group and for the general population (when known).*

## 3.4 Outcome

The outcome of interest for the primary analysis is SARS-CoV-2 infection in patients of all ages hospitalised with SARI symptoms, laboratory-confirmed by PCR documented either on admission to hospital or within 14 days before admission.

Secondary outcomes of interest, in the same patient group, are laboratory-confirmed infections with genetic variants of SARS-CoV-2 and confirmed SARS-CoV-2 patients with severe outcomes (ICU admission, invasive ventilation, death).

## 3.5 Definitions

### 3.5.1 Hospitalised patient

A hospitalised patient is as a SARI patient who has been admitted to one of the participating hospitals during the study period and has not been discharged to their home or home equivalent within 24 hours.

### 3.5.2 SARI patient (possible COVID-19 case): WHO SARI case definition

A SARI patient (possible COVID-19 case) can be defined using the World Health Organization (WHO) SARI case definition[3] as a hospitalised person with acute respiratory infection, with

- a history of fever or measured fever of ≥ 38 C°;
- and cough;
- with onset within the last 10 days.

### 3.5.3 SARI patient (possible COVID-19 case): ECDC possible COVID-19 case definition

A SARI patient (possible COVID-19 case) can also be defined using ECDC's clinical case definition for a possible COVID-19 case[4] who is hospitalised. In this situation, a possible COVID-19 case will be defined as

a hospitalised person with **at least one** of the following symptoms:

- cough;
- fever;
- shortness of breath; **or**
- sudden onset of anosmia, ageusia or dysgeusia.

SARI patients with onset of symptoms within 14 days prior to hospital admission will be included in the study. Note that hospitals already participating in SARI surveillance systems should not modify the SARI inclusion criteria for surveillance. However, for the CVE analysis, only those patients with onset of symptoms 14 days prior to hospital admission will be included.

In a later protocol version, a cut-off for days between onset of symptoms and swabbing may be decided (if appropriate).

### 3.5.4 SARI patients confirmed as COVID-19 (confirmed cases)

A confirmed COVID-19 case will be defined as a hospitalised patient with SARI symptoms fulfilling either the WHO or ECDC possible case definition, with a respiratory sample positive for SARS-CoV-2 by PCR, [4] either on admission to hospital or documented within 14 days prior to hospital admission.

### 3.5.5 Controls

#### *Test negative controls: SARI cases testing negative for SARS-CoV-2*

A control will be defined as a patient hospitalised with SARI symptoms fulfilling either the WHO or ECDC possible case definition, with a respiratory sample negative for SARS-CoV-2 by PCR on admission to hospital.

It would be beneficial if countries test for SARS-CoV-2 and influenza (during influenza season) [5,6], as well as for all other respiratory viruses (as appropriate depending on time of year), if possible. If this is not feasible, then at least all samples that are negative for SARS-CoV-2 should also be tested for influenza during influenza season, if not already tested at primary care level.

Controls who are negative by PCR but have Ct results suggestive of COVID-19[5], and those with prior SARS-CoV-2 infection in the three months prior to admission, may be excluded as controls in sensitivity analyses (see section 3.5.6 Exclusion criteria below).

- *Each study site/country to indicate which SARI case definition they will use (WHO or ECDC).*
- *Each study site/country to indicate which testing strategy they will use (testing all samples for both SARS-CoV-2 and influenza, or only testing for influenza in those negative for SARS-CoV-2).*
- *Each study site/country to indicate whether they can test for other respiratory viruses, or only SARS-CoV-2 and influenza.*

[3] WHO SARI case definition: https://apps.who.int/iris/bitstream/handle/10665/333912/WHO-2019-nCoV-Surveillance_Case_Definition-2020.1-eng.pdf?sequence=1&isAllowed=y

[4] ECDC possible COVID-19 case definition: https://www.ecdc.europa.eu/en/covid-19/surveillance/case-definition

[5] https://www.who.int/publications/i/item/diagnostic-testing-for-sars-cov-2

### *Other control groups (Optional)*

During periods of low circulation of respiratory viruses, the number of SARI patients testing negative for SARS-CoV-2 may be limited. Therefore, some sites may wish to include other control groups, either in addition to SARI controls or as an alternative, which must be recruited throughout the whole study period.

Example of other control groups include:

- Patients hospitalised with non-SARI related symptoms matched by time, age group and, if possible, underlying conditions.
  - Example of source of hospitalised non-SARI cases: hospital wards admitting patients without COVID-19
- Primary care: selection of GP patients belonging to the hospital catchment area and vaccination target group matched by time and age-group.
  - Example of source of GP patients: contact the GP of the case, and select patients from his/her list (matching by GP).
- Community controls
  - Random selection of community controls belonging to the vaccination target group matched by time and age group (e.g. vaccine registry, telephone random survey, other planned survey
  - Vaccination coverage in cases will be compared to vaccination coverage in the vaccination target population (screening method). Vaccination coverage should be available by time, age group and comorbidities.
    - Vaccination coverage among GPs patients in the hospital catchment area: random sample of GPs in the hospital catchment area can be used to compute the proportion of GP patients who are vaccinated.
    - Vaccination coverage using vaccination centres in the hospital catchment area: the vaccination coverage can be computed by dividing the number of individuals vaccinated (by age group, target group) by the number of individuals eligible for vaccination in the hospital catchment area (by age group, target group). Several methods can be used to estimate the population in the hospital catchment area.

- All control groups should represent the vaccination coverage of the population giving rise to the cases. As the circulation of SARS-CoV-2 and vaccination coverage changes over time, it is recommended to match cases and controls by time (e.g. onset of SARI symptoms) or adjust by time in the analysis.

  ➢ *For study sites/countries including other control groups, define control group, how controls will be selected, representativeness (vaccination coverage in the population giving rise to the cases) and potential limitations.*

## 3.5.6 Exclusion criteria

The patient is not enrolled in the study if she or he:

- is unwilling to participate or unable to communicate and give consent (the consent may also be given by her/his legal representative, or by specific consent procedures, acceptable according to the local ethical review process);
- has a contraindication for the COVID-19 vaccine;
- cannot be swabbed due to severe septum deviation, obstruction or other conditions that contra-indicate swabbing;
- has a history of hospitalisation within the 14 days immediately prior to this admission (including transfers from another hospital).

Information will be collected on these and other potential exclusion factors and patients will be excluded from primary analyses according to available evidence (not all available at time of writing) on these factors.

In sensitivity analyses, the CVE will be estimated:

- with different cut-offs of numbers of days between onset and swabbing, onset and hospitalisation, and between vaccination and onset of symptoms;
- excluding those positive to a seasonal coronavirus (e.g. HCoV-NL63, HCoV-229E, HCoV-OC43 and HCoV-HKU1);
- excluding those who are a current control (SARS-CoV-2 negative) but were positive by PCR or serology in the previous year before the current hospitalisation, or reported clinically confirmed COVID-19, so as to determine the best cut-off period for having had a previous positive test during the previous year vs 'any previous positive test' regardless of date;
- excluding those who have received antivirals ≤14 days prior to swabbing (to avoid false negatives; the exact cut-off and types of antivirals will be determined as more research becomes available).

Please see Annex 2 for further analysis.

- *Study sites/countries to define how they obtain informed consent from those who are too unwell at time of recruitment (e.g. oral consent with witness for those in isolation until written consent possible, and/or consent of next of kin by telephone, etc).*

## 3.6 Initial restriction to priority groups for vaccination

SARI patients are only included in the analysis if they are part of a target group for COVID-19 vaccination, for which vaccination rollout has begun in their country. This way, all SARI patients included in the study will have had the chance to be vaccinated. SARI patients swabbed prior to rollout of the COVID-19 vaccination campaign in their particular country's target groups will not be included, as they will not be eligible for vaccination.

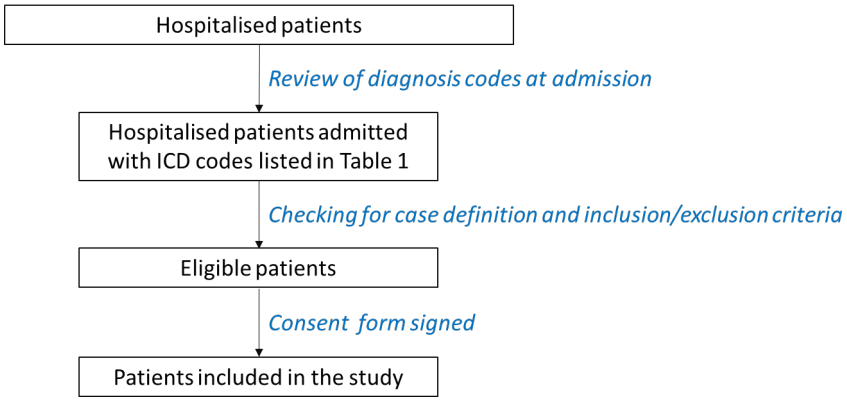
### 3.6.1 SARI patient identification – algorithm for patient inclusion

The SARI patients are identified among patients hospitalised for at least 24 hours in one of the participating hospitals. SARI patients should be enrolled and swabbed within 48 hours of hospital admission and should belong to a vaccination target group for which vaccination has already begun.

### 3.6.2 Recruitment strategies

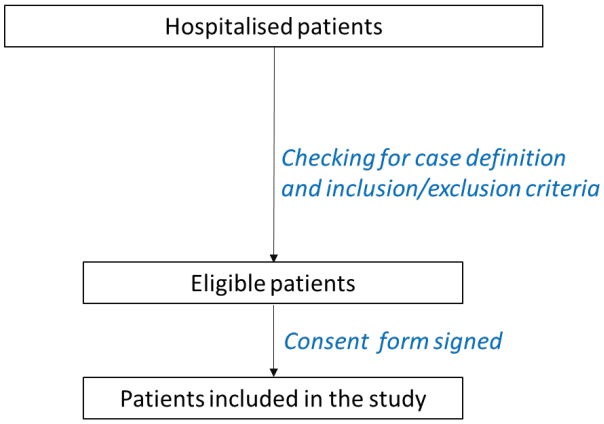
For hospitals with electronic patient records and/or diagnosis codes commonly displayed, SARI-related ICD codes (or other codes used for SARI surveillance) may be used. Patients admitted with any of the ICD codes listed in Table 1 will be approached; those meeting the SARI case definition and the inclusion criteria will be invited to be part of the study and sign an informed consent form (Figure 1).

**Figure 1. Proposed inclusion algorithm for hospitals/services relying on common use of ICD codes, hospital-based COVID-19 vaccine effectiveness study**




For hospitals where ICD codes at admission are not systematically collected or accessible, the systematic screening of all patients admitted will be organised. This should be carried out by sensitisation of the medical staff at the beginning of the study (Figure 2), followed by regular study coordinator review.

**Figure 2. Proposed inclusion algorithm for hospitals/services systematic screening of all admitted patients, hospital-based COVID-19 vaccine effectiveness study**




Retrospective recruitment (or 'catching-up' already diagnosed patients) is not recommended for the CVE study, as not all COVID-19 patients exhibit SARI symptoms and it will be difficult to determine retrospectively the reasons for testing.

- *Each study site/country to describe procedures used to identify study participants.*

In case of test scarcity, extreme workloads, or budget limiting inclusion to a threshold of patients, the study sites/countries may need to switch from exhaustive to systematic sampling (e.g. inclusion of patients every second day, or only on certain days in the week). Systematic sampling procedures should be planned ahead by the study sites. During the period of systematic selection, the study sites will make sure to document the sampling fraction.

- *Study sites/countries not testing all SARI cases to describe the systematic sampling procedure. If systematic sampling is not done, explain criteria for testing.*

**Table 1. List of diagnosis codes for which patients could be screened for onset of SARI symptoms, COVID-19 hospital-based VE study**

| Category | Morbidity | ICD-9 | ICD-10 |
|---|---|---|---|
| Influenza-like illness | Cough | 786.2 | R05 |
| | Difficulty breathing | 786.05 | R06 |
| | Sore throat | 784.1 | R07.0 |
| | Dysphagia | 787.20 | R13 |
| | Fever | 780.6 | R50.9 |
| | Headache | 784.0 | R51 |
| | Myalgia | 729.1 | M79.1 |
| | Fatigue/malaise | 780.79 | R53.1, R53.81, R53.83 |
| Cardiovascular diagnosis | Acute myocardial infarction or acute coronary syndrome | 410-411, 413-414 | I20-23, I24-25 |
| | Heart failure | 428 to 429.0 | I50, I51 |
| Respiratory diagnosis | Emphysema | 492 | J43.9 |
| | Chronic obstructive pulmonary disease | 496 | J44.9 |
| | Asthma | 493 | J45 |
| | Myalgia | 729.1 | M79.1 |
| | Dyspnoea/respiratory abnormality | 786.0 | R06.0 |
| | Respiratory abnormality | 786.00 | R06.9 |
| | Shortness of breath | 786.05 | R06.02 |
| | Tachypnoea | 786.06 | R06.82 |
| | Other respiratory abnormalities | 786.09 | R06.00, R06.09, R06.3, R06.89 |
| Infections | Pneumonia and influenza | 480-488.1 | J09-J18 |
| | Other acute lower respiratory infections | 466, 519.8 | J20-J22 |
| | Viral infection, unspecified | 790.8 | B34.9 |
| | Bacterial infection, unspecified | 041.9 | A49.9 |
| | Myocarditis | 429.0 | I40.9 |
| | Bronchitis | 490, 491 | J40, 41 |
| Inflammation | SIRS* non-infectious without acute organ dysfunction | 995.93 | R65.10 |
| | SIRS* non-infectious with acute organ dysfunction | 995.94 | R65.11 |
| Abdominal symptoms | Vomiting | 787.0 | R11 |
| | Diarrhoea | 009.3, 787.91 | A07.9, K52.9 |
| | Abdominal pain | 789.0 | R10 |
| Diagnoses related to deterioration of general condition or functional status | General physical deterioration, lethargy, tiredness | 780.79 | R53.1, R53.81, R53.83 |
| | Anorexia | 783.0 | R63.0 |
| | Feeding difficulties | 783.3 | R63.3 |
| | Abnormal weight loss | 783.21 | R63.4 |
| | Other symptoms and signs concerning food and fluid intake | 783.9 | R63.8 |
| | Disorientation/altered mental status | 780.97 | R41.0 |
| | Dizziness and giddiness | 780.4 | R42 |
| | Infective delirium | 293.0, 293.1 | F05 |
| | Coma | 780.01 | R40.2 |
| | Transient alteration of awareness | 780.02 | R40.4 |

| Category | Morbidity | ICD-9 | ICD-10 |
|---|---|---|---|
| | Other alteration of consciousness (somnolence, stupor) | 780.09 | R40.0, R40.1 |
| | Febrile convulsions (simple), unspecified | 780.31 | R56.00 |
| | Complex febrile convulsions | 780.32 | R56.01 |
| Other | Anosmia, ageusia, myalgia | 781.1, 729.1 | R43.0, R43.2, M79.1 |

** SIRS: Systemic inflammatory response syndrome*

## 3.7 Laboratory methods

Study nurses or physicians will collect respiratory specimens (see Section 4.4) from all eligible patients, respecting safety standards for COVID-19 and following WHO biosafety guidelines.[6]

- *Each study site/country to describe the type (nasopharyngeal/oropharyngeal or both) and number of swabs taken for each patient.*

PCR should be run including an internal quality control, which can check presence of cells in the respiratory specimens. In the absence of cells, a negative result should be considered inconclusive and a second swabbing should take place if possible.

The ECDC-recommended SARS-CoV-2 laboratory confirmation is by viral RNA detection with nucleic acid amplification tests, such as RT-PCR. Isolates will undergo molecular analysis for currently circulating SARS-CoV-2 virus. During the influenza season, it is recommended that influenza virus tests should also be performed, as long as there is circulation of influenza viruses in the community. [5,6] This is especially important here in order to identify any potential association between influenza and SARS-CoV-2.

Following the procedures outlined by each study, a systematic sample of isolates (or all isolates) will undergo gene sequencing. The sampling procedure can include sequencing all isolates, or a systematic sample thereof. The systematic sample should be representative of cases and be large enough to provide reasonable precision when calculating proportions of virus change over time. Guidance on random sampling for selection of samples for sequencing is provided in Annex 5.

Gene sequences, if sequencing is performed, should also be uploaded to GISAID's open access EpiCoV platform. Gene sequence information can be provided directly to the coordinating central hub for the study, or the GISAID EpiCoV accession number can be provided alongside the study unique identifier to link these data (see Annex 3, after the section on random sampling). Processed genetic information, e.g. name of genetic clade, can also be included within the epidemiological database.

- *Each study site/country to describe the laboratory procedures (samples taken, storage, transport).*
- *Each study site/country to describe the tests and the kits used (and their sensitivity, specificity, PPV) for COVID-19 and, if needed, other respiratory virus detection.*
- *Each study site/country to describe if the laboratory participates in QA/QC (Quality Assurance/Quality Control) schemes.*
- *Each study site/country to describe the selection of specimens and the procedures for genetic and antigenic characterisation, where appropriate.*
- *Each study site/country to describe genetic and antigenic analyses and specify sequencing methods*
- *Study sites/country to describe whether specimens are tested for other respiratory viruses (e.g. whether influenza continues to be tested systematically during the season and stops once the influenza season is over, or is only tested when the COVID-19 result is negative, etc).*

[6] Any non-propagative diagnostics (e.g. sequencing, RT-PCR) should be conducted at a facility using procedures equivalent to biosafety level 2 (BSL-2), while propagative work (e.g. virus culture, isolation or neutralisation assays) should be conducted at a containment laboratory with inward directional airflow (BSL-3). Patient specimens from suspected or confirmed cases should be transported as UN3373, 'biological substance category B'. Viral cultures or isolates should be transported as category A, UN2814, 'infectious substance, affecting humans'.[4]

# 3.8 Exposure (vaccination)

## 3.8.1 Definition of vaccination status

### *Current pandemic COVID-19 vaccine*

An individual will be considered as vaccinated against COVID-19 with a product-specific vaccine (see section 'COVID-19 vaccination status ascertainment') during the current pandemic under the following categories:

- **Fully vaccinated** (two-dose vaccine): to be defined according to vaccine product recommendations, but most likely patients will be considered fully vaccinated if they have **received both doses** at least 14 days* before onset.
- **Fully vaccinated** (single-dose vaccine): to be defined according to vaccine product recommendations, but most likely patients will be considered fully vaccinated if they have **received one dose** at least 14 days* before onset.
- **Partially vaccinated** (two-dose vaccine only): to be defined according to vaccine product recommendations, but most likely a patient will be considered partially vaccinated if they have **received one of two doses** at least 14 days* before onset.
- A SARI patient will be considered as ***unvaccinated*** if s/he **did not receive COVID-19 vaccine or** if s/he **was vaccinated after onset** of symptoms.

** The exact number of days will depend on the vaccine; this number may change and the protocol will be updated when more information is available. In case of any patients being vaccinated earlier or later than recommended by manufacturers, sensitivity analyses will be performed for differing delays between first and second dose.*

It is crucial that the vaccination status and date of vaccination variables are collected with the utmost care to ensure data completeness and quality. The definition of vaccination status will be updated in light of evolving decisions related to vaccination programmes, such as the use of different products for subsequent doses as well as additional doses.

## 3.8.2 COVID-19 vaccination status ascertainment

The main exposure of interest in this study is vaccination history with any COVID-19 pandemic vaccine. The vaccination history includes date of administration, type of vaccine and brand name, and number of doses. Documenting the batch codes (where this is feasible) will allow identification of the vaccine brand, the vaccine content, and the dose.

The sources of information for the vaccination status may include:

- vaccination registry (preferred option)
- consultation of the patient's vaccination card/certificate or patient's hospital notes
- interview with the patient's GP
- interview with the patient's pharmacist
- data from the patient's insurance company showing evidence of pharmacy delivery or re-imbursement for COVID-19 vaccine during the relevant period
- interview with the patient and/or his/her relatives
- self-reported vaccination status via photo on cell phone.

  - *Each study site/country to describe how vaccination status ascertainment will be performed and validated.*
  - *Each study site/country to document*
    - *vaccine products used;*
    - *places of vaccination (GPs, specific vaccination centres, etc.);*
    - *precise mode of vaccine ascertainment (self-report, card, registry, etc.);*
    - *if no precise dates of vaccination collected, the variable allowing a patient to be defined as vaccinated or unvaccinated;*
    - *vaccine status ascertainment validation.*

# 3.9 Potential confounding factors and effect modifiers

## 3.9.1 Pre-existing chronic conditions

Patients (in particular, those who are ultimately included as controls) with underlying conditions may be included due to an exacerbation of these conditions, unrelated to SARI. These patients may be more likely to be infected with SARS-CoV-2, or to develop more severe disease than the source population. Furthermore, these patients may be more likely to be vaccinated against COVID-19 than the source population. We may therefore underestimate the CVE. We will document the presence of underlying conditions among all recruited SARI patients.

Underlying conditions which could be potential confounding factors/effect modifiers are shown in Table 2 (for their ICD codes, see Annex 1).

**Table 2. List of mandatory and optional pre-existing conditions as potential effect modifiers or confounding factors**

| Mandatory | Optional |
|---|---|
| Asthma | Anaemia/chronic haematologic disease |
| Immunodeficiency (including HIV infection) and organ transplant | Asplenia |
| Cancer (solid organ and haematological) | Chronic liver disease/cirrhosis |
| Diabetes mellitus | Dementia |
| Heart disease (excluding hypertension) | Neuromuscular disorders |
| Hypertension | Renal disease (exclude acute renal failure) |
| Lung disease | Rheumatologic diseases |
| Obesity ***or*** | Stroke |
| • height and weight, ***or*** | tuberculosis |
| • BMI[7] (sites to include whichever is feasible/available) | |

- *Each study site/country to keep a list to include pre-existing conditions defining target groups for vaccination in your country.*
- *Each study site to define the list of chronic conditions to be included and state whether they are used to define target groups for vaccination, as well as any pre-existing medications being taken, and describe what the sources of information for these will be.*

## 3.9.2 Severity of underlying condition/healthcare utilisation

The severity of an underlying condition could be an effect modifier or a confounding factor, i.e. not just presence of underlying condition. To document and control for healthcare-seeking behaviour in control groups and the severity of the underlying conditions, information on the number of hospital admissions due to underlying conditions in the 12 months prior to inclusion in the study will be collected. Information on the number GP consultations (face-to-face or telephone/video) in the previous 12 months, should be collected, where available.

## 3.9.3 Ethnicity (optional)

Some studies have shown that certain ethnic groups may be at higher risk, either for becoming infected with SARS-CoV-2, or for developing severe COVID-19. Uptake of, or access to vaccination may also be linked to ethnicity. Not all sites/countries collect, or are able to collect, this information. Even if all sites /countries were able to collect information on ethnic group, each group may be defined differently in different countries. The definitions for each ethnic group will need to be standardised across study sites before we can robustly collect this information and use it to investigate VE by ethnic group for pooled data. However, any site(s)/countries who can or do already collect this information may find it useful for examining their national VE estimates by ethnic group.

[7] Obesity defined as BMI>29.

- *Each study site/country to indicate whether they can or do collect information on ethnic group, and define their ethnic groups and (proposed) method of collecting the data (e.g. by self-reported ethnicity, from patient notes, or any other method).*

### 3.9.4 Medication status for chronic condition(s) (optional)

The use of specific types of chronic medications prior to vaccination or illness may modify or confound the effect of the vaccine.

**Definition of medication status for pre-existing chronic condition(s):**

- An individual is considered as 'on medication' if s/he has received more than one dose of the medication during the six months before:
  - the first dose of pandemic COVID-19 vaccination (if date of medication use available: for the analysis measuring the effect of chronic medications on CVE); or
  - onset of SARI symptoms (for unvaccinated individuals: for the analysis measuring the effect of chronic medications on COVID-19); or
  - hospitalisation.
- An individual is considered as 'not on chronic medication' if s/he did not receive medication:
  - before the periods specified above in the protocol to document such medication use.

### 3.9.5 Chronic medication use status ascertainment (optional)

The medication history includes the date the patient started the medication(s) where known; otherwise just the year, if the patient was known to have been on medication before vaccination or symptom onset, or if the precise date is unknown. If both these are unknown, then a simple yes/no response to whether the patient was on the chronic medication before hospitalisation will be used. In addition, medication history will include medication brand name, dose, and number/ frequency of doses.

The sources of information for chronic medication status may include:

- consultation of the patient's hospital record
- interview with the patient's GP
- interview with the patient's pharmacist
- data from the patient's insurance company showing evidence of pharmacy delivery or re-imbursement for chronic medication during the relevant period
- interview with the patient and/or his/her relatives

- *Each site will define chronic medication use based on data collected.*
- *Each site to describe how chronic medication status ascertainment will be performed.*

### 3.9.6 Pregnancy status

Pregnancy status will be collected and coded for women aged 15–55 years as follows: pregnant (yes/no/do not know), and if yes: trimester (1/2/3/do not know).

### 3.9.7 Smoking history

Smoking history will be collected and coded as follows: never-smoker, former smoker (stopped smoking at least one year before inclusion in the study), current smoker (smoking currently or stopped within a year from study recruitment).

### 3.9.8 Healthcare worker

The definition of a healthcare worker for the purposes of this study is anyone working (paid or on a regular voluntary basis) in healthcare who has contact with any type of patient) during his/her work. This includes: doctors, nurses, emergency medical personnel, medical and nursing students having contact with patients, as well as porters and cleaners. It also includes anybody working with resident contact in a nursing/residential home for the elderly. Study sites should collect information on healthcare worker status where possible.

- *Each study site/country to indicate whether they can collect information on healthcare worker status.*

### 3.9.9 Other occupations (optional)

As some occupations predispose to greater exposure, and may be a proxy for attitudes towards vaccination, where possible countries may collect additional information on occupation for stratification in analysis by type of occupation (optional).

- *Each study site/country to indicate whether they intend to/can also collect information on occupation in general (optional).*

### 3.9.10 Other vaccinations

It is important to collect information on vaccinations received for influenza (including date of vaccination) and for pneumococcal disease (pneumococcal polysaccharide vaccine, PPV, and pneumococcal conjugate vaccine, PCV), where available.

#### *Previous influenza and pneumococcal vaccinations*

Vaccination against influenza in the current influenza season at the time of recruitment (where this information is available) and year of vaccination against pneumococcal disease is part of the data collection.

The sources of information for these vaccinations may include:

- vaccination registry
- consultation of the patient's vaccination card/certificate
- interview with the patient's GP
- interview with the patient's pharmacist
- data from the patient's insurance company showing evidence of pharmacy delivery or re-imbursement of influenza vaccine during the relevant period
- interview of the patient and/or his/her relatives
- self-reported vaccination status via photo on mobile phone

- *Each study site/country to indicate whether vaccination information (including date of vaccination) will be available for influenza and/or pneumococcal disease.*
- *Each study site/country to describe how previous influenza and pneumococcal vaccination status is documented.*

### 3.9.11 Antiviral administration

The use of antivirals[8] prior to swabbing may lead to misclassification biases. Sensitivity analyses excluding patients who were administered antivirals prior to swabbing are planned, as well as documenting whether the patients received any antiviral treatment in the two weeks preceding symptom onset and the type (curative or preventive) of antivirals received.

- *Each study site/country to list any antivirals administered.*

### 3.9.12 Functional impairment/frailty

Frailty may be associated with both vaccination and the risk of developing severe symptoms in case of COVID-19 infection. There are different ways in which countries may capture the presence of functional impairment related to the ability of patients to do a range of daily activities without assistance. Where possible, the Barthel Index [7] should be used. If this is not possible, countries may use simple questions related to the ability of patients to do a range of daily activities (e.g. walking, bathing) without assistance. Finally, in the absence of these, a proxy for frailty may be used, such as inclusion of a question on residence, with response options for long-term care facility and residence at home/not at home but 'with support' or 'without support'.

- *Each study site/country to describe the measure(s) used to capture functional impairment/frailty.*

[8] Types of antivirals will be determined as more research becomes available.

### 3.9.13 Presence of influenza and other respiratory virus infections

It is important to document the presence of influenza and other respiratory viruses in cases and controls. Analysis will document the presence of other respiratory infections (e.g. influenza) among patients testing negative for SARS-CoV-2, as well as in those who are positive.

- *Each study site/countries to list the other respiratory infection viruses tested for (including influenza).*

### 3.9.14 Setting (LTCF vs community)

Older and vulnerable populations, already at greater risk of severe disease, are often situated in localised settings such as those for long-term care, where they are more at risk for localised outbreaks than residents in the general community (as was observed in the early phase of the pandemic, when many hospitalisations arose in long-term care populations). It is also possible, however, that an entire LTCF (residents and staff) is vaccinated, providing them with less chance of exposure than the general population. Stratifying by setting (LTCF vs community) will help to adjust for differences between SARI patients who are LTCF residents and those who are not.

- *Each study site to ensure setting is captured, particularly for SARI patients aged over 64 years.*

### 3.9.15 Socioeconomic status or deprivation (optional)

Individuals with lower socioeconomic status (SES), who may be living in crowded conditions and have less access to good nutrition and potentially more co-morbidities, will be at greater risk of infection and severe disease, and may also be less able to access vaccination services. Stratifying by SES, if collected, will allow comparison of CVE between those of lower and higher SES.

- *Each study site to decide on the best way to represent SES for their population and to describe the measure(s) used to capture SES or level of deprivation.*

### 3.9.16 Previous SARS-CoV-2 infection

Individuals who have been previously infected may have a greater response to the vaccine or be less likely to be reinfected even if unvaccinated. Study participants should be asked about prior test results and prior symptoms to elucidate possibility of prior infection. A sensitivity analysis will be conducted by excluding individuals with prior infection.

- *Each study site/country to collect and describe measure(s) used to capture prior SARS-CoV-2 infection.*

### 3.9.17 Practice of non-pharmaceutical interventions (optional)

Individuals not practising routine use of non-pharmaceutical interventions (NPIs) will be at greater risk of infection and may also be less likely to access vaccination services. Stratifying by NPI use, if collected, will allow comparison of CVE between those regularly using and not using these key NPIs. Key NPIs proposed are: Use of a face mask in public, frequently washed hands with soap and water for at least 20 seconds, used sanitiser when soap and water were unavailable, ensured physical distancing in public (remaining at least two metres from others), with the following options: always, usually, sometimes, never, not applicable.

- *Each study site/country to define the feasibility of collecting these variables in their population. It is particularly important that these questions are asked without judgement.*
- *Sites/countries to adjust distance etc. to match national recommendations.*

## 3.10 Sample size

Providing VE estimates for each separate study is one of the objectives of this project. Therefore, the minimum sample size should be estimated for each study to obtain precise VE estimates. The pooled analyses should not prevent study teams at country level from including a big enough sample size to obtain exact estimates for each country-specific study.

- *Each study site /country to specify the minimum sample size calculation.*

- In VE estimation, sample size estimation is different from sample size estimation in hypothesis testing. Rather than being concerned about whether a VE estimation includes 0% or not, we are more concerned with the precision around the estimate. For example, if we have a VE of 70%, a lower boundary confidence interval of 1% does not provide us with a very informative VE estimate, even if the confidence interval does not include 0%. We are more concerned to have a VE estimate that is precise around the point estimate of 70% (e.g. with a lower boundary of 50%). Indeed, if we have a low VE estimate, which

can be the case in particular stratified analyses, we would need a very large sample size to provide a VE estimate where the confidence interval does not include 0%. For example, if the true VE is 5–10%, then a study providing a lower boundary not including 0% would be unreasonably large.

- The following sample size estimates focus on precision of the VE estimate (Table 3).
- The odds ratio (OR) has confidence intervals that are symmetrical around the point estimate in the log scale. However, due to the properties of the OR, these are asymmetrical around the point estimate on the arithmetic scale. The length of the confidence interval of an odds ratio on the arithmetic scale will therefore always be shorter at the lower limit (which will never reach zero). When converted to VE (VE=1-OR), this means that the length of the confidence interval will always be shorter at the upper limit. Hence we focus on the precision of the lower limit of the confidence interval, and for Table 3 below we select a potential range for this limit of between 10 and 30% (or 0.1–0.3). We also assume a case to control ratio of 1:1 for Table 3, where we include varying vaccine coverage among the source population between 30% and 90%, and varying VE (with the OR between 0.1 and 0.7).

**Table 3. Sample size calculations**

| Lower CI boundary | Controls/ case Case | Detectable OR | Vaccine coverage in source population/ controls | Number of cases | Number of controls | CVE | CI |
|---|---|---|---|---|---|---|---|
| 0.3 | 1 | 0.1 | 0.3 | 60 | 60 | 90 | 60–98 |
| 0.3 | 1 | 0.2 | 0.3 | 85 | 85 | 80 | 51–92 |
| 0.3 | 1 | 0.3 | 0.3 | 118 | 118 | 70 | 40–85 |
| 0.3 | 1 | 0.4 | 0.3 | 157 | 157 | 60 | 30–77 |
| 0.3 | 1 | 0.5 | 0.3 | 203 | 203 | 50 | 20–69 |
| 0.3 | 1 | 0.6 | 0.3 | 255 | 255 | 40 | 10–60 |
| 0.3 | 1 | 0.7 | 0.3 | 314 | 314 | 30 | 0–51 |
| 0.2 | 1 | 0.1 | 0.3 | 96 | 96 | 90 | 70–97 |
| 0.2 | 1 | 0.2 | 0.3 | 148 | 148 | 80 | 60–90 |
| 0.2 | 1 | 0.3 | 0.3 | 216 | 216 | 70 | 50–82 |
| 0.2 | 1 | 0.4 | 0.3 | 299 | 299 | 60 | 40–73 |
| 0.2 | 1 | 0.5 | 0.3 | 395 | 395 | 50 | 30–64 |
| 0.2 | 1 | 0.6 | 0.3 | 507 | 507 | 40 | 20–55 |
| 0.2 | 1 | 0.7 | 0.3 | 633 | 633 | 30 | 10–46 |
| 0.1 | 1 | 0.1 | 0.3 | 241 | 241 | 90 | 80–95 |
| 0.1 | 1 | 0.2 | 0.3 | 433 | 433 | 80 | 70–87 |
| 0.1 | 1 | 0.3 | 0.3 | 681 | 681 | 70 | 60–77 |
| 0.1 | 1 | 0.4 | 0.3 | 985 | 985 | 60 | 50–68 |
| 0.1 | 1 | 0.5 | 0.3 | 1 346 | 1 346 | 50 | 40–58 |
| 0.1 | 1 | 0.6 | 0.3 | 1 764 | 1 764 | 40 | 30–49 |
| 0.1 | 1 | 0.7 | 0.3 | 2 240 | 2 240 | 30 | 20–39 |
| 0.3 | 1 | 0.1 | 0.4 | 42 | 42 | 90 | 60–98 |
| 0.3 | 1 | 0.2 | 0.4 | 63 | 63 | 80 | 49–92 |
| 0.3 | 1 | 0.3 | 0.4 | 91 | 91 | 70 | 40–85 |
| 0.3 | 1 | 0.4 | 0.4 | 125 | 125 | 60 | 30–77 |
| 0.3 | 1 | 0.5 | 0.4 | 165 | 165 | 50 | 20–69 |
| 0.3 | 1 | 0.6 | 0.4 | 212 | 212 | 40 | 10–60 |
| 0.3 | 1 | 0.7 | 0.4 | 265 | 265 | 30 | 0–51 |
| 0.2 | 1 | 0.1 | 0.4 | 68 | 68 | 90 | 70–97 |
| 0.2 | 1 | 0.2 | 0.4 | 111 | 111 | 80 | 60–90 |
| 0.2 | 1 | 0.3 | 0.4 | 168 | 168 | 70 | 50–82 |
| 0.2 | 1 | 0.4 | 0.4 | 238 | 238 | 60 | 40–73 |
| 0.2 | 1 | 0.5 | 0.4 | 323 | 323 | 50 | 30–64 |
| 0.2 | 1 | 0.6 | 0.4 | 421 | 421 | 40 | 20–55 |
| 0.2 | 1 | 0.7 | 0.4 | 534 | 534 | 30 | 10–46 |
| 0.1 | 1 | 0.1 | 0.4 | 170 | 170 | 90 | 80–95 |
| 0.1 | 1 | 0.2 | 0.4 | 323 | 323 | 80 | 70–87 |
| 0.1 | 1 | 0.3 | 0.4 | 528 | 528 | 70 | 60–77 |
| 0.1 | 1 | 0.4 | 0.4 | 786 | 786 | 60 | 50–68 |
| 0.1 | 1 | 0.5 | 0.4 | 1 098 | 1 098 | 50 | 40–58 |

| Lower CI boundary | Controls/ case Case | Detectable OR | Vaccine coverage in source population/ controls | Number of cases | Number of controls | CVE | CI |
|---|---|---|---|---|---|---|---|
| 0.1 | 1 | 0.6 | 0.4 | 1 466 | 1 466 | 40 | 30–49 |
| 0.1 | 1 | 0.7 | 0.4 | 1 891 | 1 891 | 30 | 20–39 |
| 0.3 | 1 | 0.1 | 0.5 | 32 | 32 | 90 | 60–98 |
| 0.3 | 1 | 0.2 | 0.5 | 51 | 51 | 80 | 51–92 |
| 0.3 | 1 | 0.3 | 0.5 | 77 | 77 | 70 | 40–85 |
| 0.3 | 1 | 0.4 | 0.5 | 109 | 109 | 60 | 30–77 |
| 0.3 | 1 | 0.5 | 0.5 | 148 | 148 | 50 | 20–69 |
| 0.3 | 1 | 0.6 | 0.5 | 193 | 193 | 40 | 10–60 |
| 0.3 | 1 | 0.7 | 0.5 | 246 | 246 | 30 | 0–51 |
| 0.2 | 1 | 0.1 | 0.5 | 51 | 51 | 90 | 70–97 |
| 0.2 | 1 | 0.2 | 0.5 | 90 | 90 | 80 | 60–90 |
| 0.2 | 1 | 0.3 | 0.5 | 142 | 142 | 70 | 50–82 |
| 0.2 | 1 | 0.4 | 0.5 | 208 | 208 | 60 | 40–73 |
| 0.2 | 1 | 0.5 | 0.5 | 289 | 289 | 50 | 30–64 |
| 0.2 | 1 | 0.6 | 0.5 | 384 | 384 | 40 | 20–55 |
| 0.2 | 1 | 0.7 | 0.5 | 495 | 495 | 30 | 10–46 |
| 0.1 | 1 | 0.1 | 0.5 | 129 | 129 | 90 | 80–95 |
| 0.1 | 1 | 0.2 | 0.5 | 262 | 262 | 80 | 70–87 |
| 0.1 | 1 | 0.3 | 0.5 | 447 | 447 | 70 | 60–78 |
| 0.1 | 1 | 0.4 | 0.5 | 687 | 687 | 60 | 50–68 |
| 0.1 | 1 | 0.5 | 0.5 | 983 | 983 | 50 | 40–58 |
| 0.1 | 1 | 0.6 | 0.5 | 1 337 | 1 337 | 40 | 30–49 |
| 0.1 | 1 | 0.7 | 0.5 | 1 751 | 1 751 | 30 | 20–39 |
| 0.3 | 1 | 0.1 | 0.6 | 26 | 26 | 90 | 60-98 |
| 0.3 | 1 | 0.20 | 0.60 | 45 | 45 | 80 | 50-92 |
| 0.3 | 1 | 0.30 | 0.60 | 71 | 71 | 70 | 40-85 |
| 0.3 | 1 | 0.40 | 0.60 | 103 | 103 | 60 | 30-77 |
| 0.3 | 1 | 0.50 | 0.60 | 143 | 143 | 50 | 20-69 |
| 0.3 | 1 | 0.60 | 0.60 | 191 | 191 | 40 | 10-60 |
| 0.3 | 1 | 0.70 | 0.60 | 247 | 247 | 30 | 0-51 |
| 0.2 | 1 | 0.10 | 0.60 | 41 | 41 | 90 | 70-97 |
| 0.2 | 1 | 0.20 | 0.60 | 78 | 78 | 80 | 60-90 |
| 0.2 | 1 | 0.30 | 0.60 | 130 | 130 | 70 | 50-82 |
| 0.2 | 1 | 0.40 | 0.60 | 197 | 197 | 60 | 40-73 |
| 0.2 | 1 | 0.50 | 0.60 | 280 | 280 | 50 | 30-64 |
| 0.2 | 1 | 0.60 | 0.60 | 380 | 380 | 40 | 20-55 |
| 0.2 | 1 | 0.70 | 0.60 | 497 | 497 | 30 | 10-46 |
| 0.1 | 1 | 0.10 | 0.60 | 104 | 104 | 90 | 80-95 |
| 0.1 | 1 | 0.20 | 0.60 | 229 | 229 | 80 | 70-87 |
| 0.1 | 1 | 0.30 | 0.60 | 410 | 410 | 70 | 60-78 |
| 0.1 | 1 | 0.40 | 0.60 | 651 | 651 | 60 | 50-68 |

| Lower CI boundary | Controls/ case Case | Detectable OR | Vaccine coverage in source population/ controls | Number of cases | Number of controls | CVE | CI |
|---|---|---|---|---|---|---|---|
| 0.1 | 1 | 0.50 | 0.60 | 953 | 953 | 50 | 40-58 |
| 0.1 | 1 | 0.60 | 0.60 | 1 322 | 1 322 | 40 | 30-49 |
| 0.1 | 1 | 0.70 | 0.60 | 1 760 | 1 760 | 30 | 20-39 |
| 0.3 | 1 | 0.10 | 0.70 | 23 | 23 | 90 | 60-98 |
| 0.3 | 1 | 0.20 | 0.70 | 43 | 43 | 80 | 50-92 |
| 0.3 | 1 | 0.30 | 0.70 | 71 | 71 | 70 | 40-85 |
| 0.3 | 1 | 0.40 | 0.70 | 108 | 108 | 60 | 30-77 |
| 0.3 | 1 | 0.50 | 0.70 | 153 | 153 | 50 | 20-69 |
| 0.3 | 1 | 0.60 | 0.70 | 207 | 207 | 40 | 10-60 |
| 0.3 | 1 | 0.70 | 0.70 | 272 | 272 | 30 | 0-51 |
| 0.2 | 1 | 0.10 | 0.70 | 36 | 36 | 90 | 70-97 |
| 0.2 | 1 | 0.20 | 0.70 | 75 | 75 | 80 | 60-90 |
| 0.2 | 1 | 0.30 | 0.70 | 131 | 131 | 70 | 50-82 |
| 0.2 | 1 | 0.40 | 0.70 | 205 | 205 | 60 | 40-73 |
| 0.2 | 1 | 0.50 | 0.70 | 298 | 298 | 50 | 30-64 |
| 0.2 | 1 | 0.60 | 0.70 | 412 | 412 | 40 | 20-55 |
| 0.2 | 1 | 0.70 | 0.70 | 548 | 548 | 30 | 10-46 |
| 0.1 | 1 | 0.10 | 0.70 | 90 | 90 | 90 | 80-95 |
| 0.1 | 1 | 0.20 | 0.70 | 219 | 219 | 80 | 70-87 |
| 0.1 | 1 | 0.30 | 0.70 | 413 | 413 | 70 | 60-78 |
| 0.1 | 1 | 0.40 | 0.70 | 676 | 676 | 60 | 50-68 |
| 0.1 | 1 | 0.50 | 0.70 | 1 015 | 1 015 | 50 | 40-58 |
| 0.1 | 1 | 0.60 | 0.70 | 1 435 | 1 435 | 40 | 30-49 |
| 0.1 | 1 | 0.70 | 0.70 | 1 941 | 1 941 | 30 | 20-39 |
| 0.3 | 1 | 0.10 | 0.80 | 22 | 22 | 90 | 60-98 |
| 0.3 | 1 | 0.20 | 0.80 | 47 | 47 | 80 | 50-92 |
| 0.3 | 1 | 0.30 | 0.80 | 82 | 82 | 70 | 40-85 |
| 0.3 | 1 | 0.40 | 0.80 | 128 | 128 | 60 | 30-77 |
| 0.3 | 1 | 0.50 | 0.80 | 187 | 187 | 50 | 20-69 |
| 0.3 | 1 | 0.60 | 0.80 | 259 | 259 | 40 | 10-60 |
| 0.3 | 1 | 0.70 | 0.80 | 344 | 344 | 30 | 0-51 |
| 0.2 | 1 | 0.10 | 0.80 | 35 | 35 | 90 | 70-97 |
| 0.2 | 1 | 0.20 | 0.80 | 82 | 82 | 80 | 60-90 |
| 0.2 | 1 | 0.30 | 0.80 | 151 | 151 | 70 | 50-82 |
| 0.2 | 1 | 0.40 | 0.80 | 245 | 245 | 60 | 40-73 |
| 0.2 | 1 | 0.50 | 0.80 | 365 | 365 | 50 | 30-64 |
| 0.2 | 1 | 0.60 | 0.80 | 514 | 514 | 40 | 20-55 |
| 0.2 | 1 | 0.70 | 0.80 | 694 | 694 | 30 | 10-46 |
| 0.1 | 1 | 0.10 | 0.80 | 89 | 89 | 90 | 80-95 |
| 0.1 | 1 | 0.20 | 0.80 | 241 | 241 | 80 | 70-87 |
| 0.1 | 1 | 0.30 | 0.80 | 477 | 477 | 70 | 60-78 |

| Lower CI boundary | Controls/ case Case | Detectable OR | Vaccine coverage in source population/ controls | Number of cases | Number of controls | CVE | CI |
|---|---|---|---|---|---|---|---|
| 0.1 | 1 | 0.40 | 0.80 | 808 | 808 | 60 | 50-68 |
| 0.1 | 1 | 0.50 | 0.80 | 1 242 | 1 242 | 50 | 40-58 |
| 0.1 | 1 | 0.60 | 0.80 | 1 789 | 1 789 | 40 | 30-49 |
| 0.1 | 1 | 0.70 | 0.80 | 2 458 | 2 458 | 30 | 20-39 |
| 0.3 | 1 | 0.10 | 0.90 | 30 | 30 | 90 | 60-98 |
| 0.3 | 1 | 0.20 | 0.90 | 71 | 71 | 80 | 50-92 |
| 0.3 | 1 | 0.30 | 0.90 | 129 | 129 | 70 | 40-85 |
| 0.3 | 1 | 0.40 | 0.90 | 208 | 208 | 60 | 30-77 |
| 0.3 | 1 | 0.50 | 0.90 | 310 | 310 | 50 | 20-69 |
| 0.3 | 1 | 0.60 | 0.90 | 437 | 437 | 40 | 10-60 |
| 0.3 | 1 | 0.70 | 0.90 | 591 | 591 | 30 | 0-51 |
| 0.2 | 1 | 0.10 | 0.90 | 48 | 48 | 90 | 70-97 |
| 0.2 | 1 | 0.20 | 0.90 | 124 | 124 | 0.8 | 60-90 |
| 0.2 | 1 | 0.30 | 0.90 | 238 | 238 | 0.7 | 50-82 |
| 0.2 | 1 | 0.40 | 0.90 | 397 | 397 | 0.6 | 40-73 |
| 0.2 | 1 | 0.50 | 0.90 | 605 | 605 | 0.5 | 30-64 |
| 0.2 | 1 | 0.60 | 0.90 | 868 | 868 | 0.4 | 20-55 |
| 0.2 | 1 | 0.70 | 0.90 | 1 190 | 1 190 | 0.3 | 10-46 |
| 0.1 | 1 | 0.10 | 0.90 | 121 | 121 | 0.9 | 80-95 |
| 0.1 | 1 | 0.20 | 0.90 | 361 | 361 | 0.8 | 70-87 |
| 0.1 | 1 | 0.30 | 0.90 | 751 | 751 | 0.7 | 60-78 |
| 0.1 | 1 | 0.40 | 0.90 | 1 311 | 1 311 | 0.6 | 50-68 |
| 0.1 | 1 | 0.50 | 0.90 | 2 061 | 2 061 | 0.5 | 40-58 |
| 0.1 | 1 | 0.60 | 0.90 | 3 022 | 3 022 | 0.4 | 30-49 |
| 0.1 | 1 | 0.70 | 0.90 | 4 216 | 4 216 | 0.3 | 20-39 |

*The sample size estimates above are for the crude analysis and an adjusted analysis would require a higher sample size. The sample size should also be respected for each population subgroup for which a sub (stratified) analysis (e.g. effect modification) is planned.*

See also the Analysis section on sample size requirements for analyses.

# 3.11 Data

## 3.11.1 Sources of information

Data are to be collected using a standardised questionnaire/data collection form. The source(s) of data may include:

- vaccination card/certificate;
- hospital medical recordsl
- interview with patient or his/her family;
- communication with the patient or his/her family via mobile phone;
- interview with patient's GP;
- interview with patent's pharmacist;
- vaccination register;
- laboratory records.

➢ *Each study site/country to define the sources of information used for each variable collected and the potential limitations.*

## 3.12 Collected information

Collected information falls under the following main categories:

- study identification
  - country, hospital;
  - vaccination target groups;
  - first ward of referral;
  - ICU/other ward of admission.
- patient characteristics *(ethnic group optional)*
- SARI signs, symptoms
  - current;
  - previous clinical symptoms (if no prior tests done).
- other symptoms (optional)
- dates
  - vaccination (COVID-19, influenza, pneumococcal disease);
  - onset of SARI symptoms;
  - admission, discharge;
  - swabbing.
- laboratory
  - type of swab/sample (nasopharyngeal/sputum, etc.) *(optional)*;
  - type of test;
  - results (including information from antigenic and genetic analysis, where available);
  - previous positive PCR or antigen test for SARS-CoV-2, if feasible (for sensitivity analyses).
- underlying chronic conditions, including obesity (see sections 3.9.1–3.9.4)
  - use of medications for chronic conditions *(optional)*;
  - number of hospitalisations for chronic conditions in the previous 12 months *(optional)*;
  - number of GP consultations in the previous 12 months *(optional)*.
- presence of influenza and other respiratory viruses infection (see section 3.9.10)
- vaccination and antivirals (see sections 3.9 and 3.10.10–3.10.11)
  - COVID-19 vaccination including number of doses, date, product/brand;
  - influenza vaccination from the current season;
  - pneumococcal vaccination status, type of vaccine and either date or year of vaccination *(optional)*;
  - antiviral administration.
- functional status or proxy by residence type (see section 3.10.12)
- setting (e.g. LTCF)
- SES/deprivation *(optional)*
- non-pharmaceutical interventions (*optional)* (section 3.10.17)
- outcome

COVID-19 vaccine data collected will be revised as more information on the vaccine(s) and target groups becomes available.

These are described in more detail below (see also Annex 1 for a complete variable list including coding).

### 3.12.1 Study identifiers

For each country/study sites the following characteristics need to be documented:

Country, site, priority vaccination target group(s) and their modifications

- Hospital (unique number not including hospital name, to allow adjustment by hospital in analyses);
- Patient unique ID (note: this is not a patient identifiable ID such as date-of-birth or national ID number, but a unique identifier for a pooled database).

### 3.12.2 Hospital/ward information

The following dates and other hospital information need to be documented:

- Date of onset, admission, discharge, death;
- First ward of referral;
- Any hospital stay (for pre-existing chronic condition) in previous 12 months *(optional)*;
- Date of swab/sample.

### 3.12.3 Patient characteristics

The following patient characteristics need to be documented to describe the study population.

- Age;
- Sex;
- Smoking history (see section 3.9.5);
- Pregnancy status;
- Healthcare worker status;
- Occupation *(optional)*;
- Clinical frailty score at admission (where possible; see section 3.9.10) *(optional)*;
- Ethnic group *(optional)*;
- SES/deprivation (*optional*).

  - *Each country to describe type of clinical frailty score in use, where available.*
  - *Each country to describe community measures in place to limit exposures.*

### 3.12.4 Clinical characteristics (symptoms and markers of severity)

The following clinical characteristics and markers of disease severity should be documented:

The four following key symptoms which are part of the WHO and ECDC COVID-19 case definitions:

- fever or feverishness;
- cough;
- shortness of breath;
- sore throat.

The following three symptoms which are associated with COVID-19 illness and are part of ECDC's COVID-19 case definition:

- anosmia;
- ageusia;
- dysgeusia.

In addition, for study sites/countries which will also use this protocol to measure influenza VE, information on the following four symptoms should also be collected:

- headache;
- myalgia;
- malaise;
- deterioration of general condition (asthenia, weight loss, anorexia).

It would also be helpful to collect information on whether symptoms appeared with **sudden onset.**

The following 14 symptoms are **optional** for the hospital-based COVID-19 CVE study:

- coryza, rhinitis;
- chest pain;
- chills;
- fatigue;
- nausea;
- vomiting;
- abdominal pain;
- diarrhoea;
- conjunctivitis;
- confusion;
- dizziness;
- tachypnoea or other signs of low oxygen saturation (restlessness);
- rash or other dermatological manifestation;
- palpitations/rapid heartbeat.

The following information, which can be used to indicate severity to measure CVE to prevent severe disease:

- Oxygen use;
- ICU admission;
- Invasive ventilation;

- Death.

**Date of first key symptom onset**, as well as information on COVID-19 test(s) and laboratory results, including information on antigenic and genetic analysis, when available is to be collected. It is vital that this information is collected, as well as **date of vaccination**, **date of swab or date sample obtained** (to allow estimation of and stratification by delay from swab to onset), **date of admission** (to allow estimation of and stratification by time from onset to hospitalisation, and to measure length of hospital stay), and **date of discharge/death** (to allow measurement of length of hospital stay).

### 3.12.5 Case definition

Collection of good quality symptom information is crucial for the CVE study to be able to validate the case definition used. As a minimum, there is need to collect data on the symptoms required for the WHO or ECDC case definitions. The following variables are imperative for application of the WHO SARI case definition:

- fever or feverishness
  - if fever: measured fever (with temperature), or feverishness;
- cough;
- onset date;
- admission date.

The following additional variables will also be imperative if the study site is using ECDC's case definition:

- shortness of breath;
- sudden onset;
- anosmia;
- ageusia;
- dysgeusia.

### 3.12.6 Data entry validation

For hospitals using electronic medical records, if paper questionnaires are used, a sample of them will be checked against the medical records and against the study database. The agreement between patient records/reports by study participants will be measured when/if records are available.

- *Each study site/country to specify how data are validated.*

# 3.13 Data analysis

## 3.13.1 Individual (country /site level) analysis

Briefly, cases and controls will first be described by baseline characteristics. The timing to conduct each interim analysis will depend on the time needed to reach the appropriate sample size. This will depend mainly on the incidence of hospitalisation, COVID-19 incidence, vaccination coverage, the recruitment strategy within hospital/s and the number of participating hospitals/services per hospital.

Patients will be described according to:

- sex;
- age group;
- healthcare worker status;
- time: month of symptom onset;
- COVID-19 vaccination status;
- symptoms;
- absence, presence of at least one, presence of more than one high-risk condition;
- specific chronic conditions (e.g. respiratory, cardiovascular diseases);
- pregnancy, smoking status;
- influenza and pneumococcal vaccination status;
- respiratory co-infections (where available);
- severity (ICU, oxygen use, invasive ventilation, death);
- COVID-19 variant;
- vaccination status, with vaccinated patients described by vaccine product.

An example layout of this descriptive analysis is provided in Table 4 on the next page.

This study is a case control study (test-negative design). The measure of association is an odds ratio (OR). This can be estimated by logistic regression. An OR = 1 indicates no association between an exposure and the outcome. An OR>1 indicates a potential risk factor, an OR < 1 indicates a potential protective factor, noting that the confidence interval around the OR helps with its interpretation.

For vaccination as preventive factor, the CVE can be computed as CVE = (1 – OR)*100. A 95% confidence interval is computed around the point estimate.

Univariable analysis will be carried out to measure the CVE against being a laboratory-confirmed COVID-19 SARI case. Stratified analyses (by sex and age group, for example) can follow to better understand potential effect modifiers and confounders.

**Table 4. Example of descriptive table for cases and controls; COVID-19 hospital-based vaccine effectiveness study, European multicentre study, 2021**

| Variables | Number of laboratory-confirmed COVID-19 cases/total n (%) | Number of test-negative controls/total n (%) |
|---|---|---|
| **Median age (IQR)** | X | X |
| Missing | X | X |
| **Age groups (years)** | | |
| 18–44 | x/x (x) | x/x (x) |
| 45–64 | x/x (x) | x/x (x) |
| 65–79 | | |
| ≥ 80 | x/x (x) | x/x (x) |
| Missing | X | X |
| **Sex** | | |
| Female | x/x (x) | x/x (x) |
| Missing | X | X |
| **Healthcare worker** | x/x (x) | x/x (x) |
| Missing | X | X |
| **Days between onset of symptoms and swabbing** | | |
| 0 | x/x (x) | x/x (x) |
| 1 | x/x (x) | x/x (x) |
| 2 | x/x (x) | x/x (x) |
| 3 | x/x (x) | x/x (x) |
| 4–7 | x/x (x) | x/x (x) |
| **COVID-19 vaccination** | x/x (x) | x/x (x) |
| Missing | x | x |
| **Etc.** | | |

Prior to multivariable analysis, a model development strategy will be determined and included in the plan of analysis. In the final step, multivariable analysis will be carried out to take confounding factors and potential effect modifiers into account. This will provide adjusted ORs from which the CVE can be estimated using the formula above.

### Output tables presenting CVE estimates

To present the results in the most transparent manner and to enable the reader to best understand the data, tables similar to the one illustrated by Table 5 can be used (variables presented just as example of the output format). Useful information includes numbers of cases and controls (overall and vaccinated) and presentation of results for different models.

**Table 5. Example of table showing vaccine effectiveness against COVID-19 adjusted for various covariables by sex and age group, hospital-based COVID-19 vaccine effectiveness study, European multicentre study, 2021**

| Clade/variant | Population included | Analysis scenarios/adjustments made | CVE (%) | (95%CI) |
|---|---|---|---|---|
| COVID-19 | All ages | N (cases/ vaccinated; controls/ vaccinated) | | |
| | | Crude | | |
| | | Adjusted for onset week (cubic spline) | | |
| | | Adjusted for sex | | |
| | | Adjusted for chronic condition | | |
| | | Adjusted for age (cubic spline) | | |
| | | Adjusted for onset week, age (cubic spline) | | |
| | | Adjusted for onset week, chronic condition | | |
| | | Adjusted for onset week, age (cubic spline), chronic conditions, sex | | |
| | 0–49 years | N (cases/ vaccinated; controls/ vaccinated) | | |
| | | Crude | | |
| | | Adjusted for onset month, age (cubic spline) | | |
| | 50 years and over | N (cases/ vaccinated; controls/ vaccinated) | | |
| | | Crude | | |
| | | Adjusted for onset week, age (cubic spline), chronic condition, sex | | |

### 3.13.2 Pooled analysis

See Annex 2 for details of the pooled analysis. For the pooled data, interim analyses will be conducted in different periods if appropriate and according to the available sample size.

The timing to conduct each interim analysis will depend on the time needed to reach the appropriate sample size. This will depend mainly on the incidence of hospitalisation, COVID-19 incidence, vaccination coverage, the recruitment strategy within hospitals and the number of participating hospitals/services per hospital.

The pooled analysis will be carried out in a similar way to the site-specific analysis. Country or study site will be included potentially as a fixed effect or as a random effect in a multilevel model. Statistical heterogeneity between study sites will be determined, using Q-test and the $I^2$ index.[8]

## 3.14 Personal data protection

Each study site/country conducting the study shall comply with requirements stemming from data protection legislation, and with national ethics committee requirements. Informed consent will be required from all participants or legal tutors. Where data protection legislation allows so, the national ethics committees will specify whether oral or written consent will be required. Specific consent procedures may be needed for unconscious patients and patients with deterioration of general condition or functional status, unable to sign the consent (e.g. oral witnessed consent, consent by the next of kin, etc).

## 3.15 Training

Investigators and data collectors will be trained on the study protocol before the start of the study. They will receive the protocol, questionnaires, and laboratory respiratory specimen collection procedures.

- *Each study site to describe the training to be organised.*

# 4 Logistical aspects

## 4.1 Respiratory specimen collection

By default, the respiratory specimen will be collected through nasal/nasopharyngeal swabbing or concurrent nasal and oral/oropharyngeal swabbing (or endotracheal aspirates in ICU). Personal protection equipment must be used in accordance with guidelines.

- *Each study site to describe the specimen collection procedures.*

## 4.2 Laboratory tests

High specificity is needed for COVID-19 confirmation. COVID-19 laboratory confirmation will be done using RT-PCR or multiplex RT-PCR.

- *Each study site to describe the tests and the kits used for COVID-19 and influenza; and, if needed, other respiratory virus detection.*
- *Each study site to specify sequencing methods.*

PCR should be run including an internal quality control, which can check presence of cells in the respiratory specimens. Quality control tests should systematically be run using PCR to test for presence of cells in the respiratory specimens. In addition, quality assurance of assay performance at sites should be undertaken using international, national or research standards. [9]

- *Each study site to describe quality controls for specimens.*
- *Each study site to describe genetic and antigenic analyses.*

# 5 Limitations

With any multi-centre study, there is always the potential for heterogeneity among sites. In addition, during a pandemic with such high caseloads for hospitals, there may be difficulties in collecting all data, and not all included cases will have laboratory confirmation. There is also the possibility that very severely ill patients (e.g. those who are extremely frail and/or in nursing homes) may not be admitted to hospital at all, and would be missed by the study. Potential limitations to the VE estimates for COVID-19 are discussed below.

## 5.1 Potential biases

### 5.1.1 Bias from pooled estimates

With data from several hospitals from different countries being pooled, any bias in the individual studies will influence the pooled estimate. The power of the test for the presence of heterogeneity between individual studies is low if there are few sites/countries. In this case, the test may not be able to detect heterogeneity between them, despite it being present. It is important that heterogeneity is also assessed using qualitative knowledge about differences between studies. Depending on the nature of the bias, the inclusion of biased studies in the pooled estimate could lead to over- or underestimation of the true association between COVID-19 vaccination and the outcome.

There are many conditions which could lead to bias in a single site or hospital. With this new virus, there are new and evolving surveillance systems and strategies in each participating country. There are not only different tests being used, but a variation in the number of tests used to declare an individual negative, for example. Another example is that, when under high pressure (e.g. high volume of patients to be admitted during a peak in the epidemic for any site), it is possible that some hospitals may switch to admitting only suspected COVID-19 patients, while others focus on non-COVID-19 patients. In the event of the former type of hospital participating, this could affect the recruitment of controls and result in cases being predominantly recruited from one hospital over another. If a participating site only has one hospital providing data, this could mean they are only able to provide information on cases. Conversely, if the single participating hospital was designated a non-COVID-19 admitting hospital, this site would only be able to provide information on controls.

To allow for complete assessment of heterogeneity, sites need to document all changes in their COVID-19 surveillance system during the study period.

- *Each country to document any changes in COVID-19 surveillance during the study period, including allocation of participating hospitals to COVID-19 or non-COVID-19 admission status.*

### 5.1.1. Negative confounding

Negative confounding refers to biases that reflect the fact that high-risk groups (people more likely to develop severe complications) will be more likely to be vaccinated and therefore reduce CVE. If negative confounding is present, the CVE will be underestimated. Adjustment for potential negative confounding factors documented in the study (e.g. presence of chronic diseases) will minimise negative confounding.

### 5.1.2. Positive confounding

Positive confounding refers to biases that reflect a 'healthy vaccinee effect'. People with a healthy lifestyle will be more likely to accept vaccination, thus leading to an increase of measured CVE. Or, similarly, people being in a state of 'extreme frailty' will not be offered vaccination and, because they are frail, may be more likely to have severe disease. Persons with risk-taking behaviours may also be averse to vaccination, which may also increase their exposure to disease. If positive confounding is present, CVE will be overestimated.

### 5.1.3. Unmeasured confounding

Positive and negative confounding will be minimised through stratification and multivariable analysis. It will not be possible to rule out the presence of characteristics in the study population for which no information is collected in the study questionnaire and that therefore could lead to positive or negative confounding. Therefore, some residual unmeasured confounding may remain.

- *Each study site to describe the potential limitations and representativeness of the subjects included.*

### 5.1.4. Previous infection in cases or controls; inclusion of asymptomatic controls

Individuals who have been previously infected may have a stronger response to the vaccine or be less likely to be reinfected even if unvaccinated. It is possible that some of the controls (those testing negative for SARS-CoV-

2) may have themselves been positive for SARS-CoV-2 some time before, but were asymptomatic. The proportion of these (potentially immune individuals) in each country's dataset would depend on the circulation of the virus in the community in the months before the hospitalisation of the control. Knowledge of their prior infection could affect their likelihood to be vaccinated. For example, if someone knew that they had had COVID-19, despite having no symptoms (e.g. if they had had a screening test), they may be subsequently less likely to be vaccinated. This would lower vaccination coverage among controls and increase CVE.

Ascertainment of which controls may have had previous SARS-CoV-2 infection can be attempted by asking about previous SARS-CoV-2 tests and results, as well as prior clinical symptoms. However, among the controls, there could potentially be several patients with prior SARS-CoV-2 infection. Results should be interpreted in light of this, and an estimate of a range of potential bias should be calculated around the CVE estimates. Sensitivity analyses should be conducted excluding any SARI patient with previous SARS-CoV-2 infection confirmed either by PCR or by serological tests.

As antibody tests become more widespread, then this may be included in the protocol.

### 5.1.5. Inclusion of influenza-positive controls

It is possible that SARI patients who are also influenza positive will be unsuitable controls. There is limited information on co-infection with influenza and COVID-19 from the first wave of the pandemic in Europe, partly due to the timing of the pandemic being towards or after the end of the 2019–20 influenza season in many countries. The low number of coinfections described in the literature [10,11] could be due to lack of opportunity (there being little influenza circulating at that time) or to a negative correlation between the two infections, with those positive for COVID-19 being unlikely to also be positive for influenza. In addition, those receiving COVID-19 vaccination are highly likely to have also received influenza vaccine. There is therefore the potential for a relationship between being positive for influenza and receiving COVID-19 vaccination, which introduces bias. Sensitivity analyses will be conducted excluding controls who are positive for influenza.

### 5.1.6. Validation of exposure

The vaccination status is the exposure of interest and the validity of vaccination data should therefore be checked carefully. If the vaccination status is reported by the patient only without further proof, information bias may occur. Vaccination status of cases and controls should be validated using an independent source (i.e. vaccination register, GPs).

- *Each study site to describe how the source of exposure validation and its potential limitations.*

### 5.1.7. Misclassification

The use of antivirals prior to swabbing may lead to misclassification biases. Sensitivity analyses will be run excluding patients who were administered antivirals prior to swabbing. In addition, misclassification can occur due to test performance. In analysis, sensitivity and specificity of the tests can be adjusted for. Sites may use different tests, so investigators should seek to use common international, national or research standards to address possible variation in test performance at sites. Currently the National Institute of Biological Standards and Control offers international standards for molecular and serological testing. [9]

### 5.1.8. Other potential biases

Controls could come from different source populations with varying risk for infection with SARS-CoV-2, varying probability for acquiring COVID-19 vaccination, etc. (e.g. depending on time of year). Time (onset date) will be used to adjust for seasonal differences. Analyses will also be stratified by time (e.g. onset quarter of the year).

- *Each country to describe timeline of vaccination for different target groups*

## 5.2 Representativeness of subjects included in the study

The study includes only cases that are hospitalised. Health-seeking behaviour may differ by country depending on the case management strategy (e.g. recommendation to stay at home with mild symptoms, and only seeing a GP if symptoms persist, and then hospitalisation if severe). In some cases, the management strategy will have an impact on the delay between onset of symptoms and hospitalisation. This, in turn, may have an impact on the time lag between onset and respiratory specimen collection, and may affect positivity rates between study sites. Beside the collection of dates of onset/admission/respiratory specimen collection, health-seeking behaviour and case-management strategies should be described for each study and it should be noted how these may affect the CVE estimates.

Some very severely ill patients will not be able to give informed consent. If this is the case, SARI cases included will not represent very severe cases.

Importantly, the representativeness of the controls needs consideration. (For example, if controls were to be all influenza **and** COVID-19 negative, it needs to be considered whether they are representative of the source population in terms of vaccine coverage.)

- *Each study site to describe the potential limitations in terms of representativeness of the subjects included.*
- *Each study site to describe case-management strategy in their country.*

This core protocol will be updated in light of evolving scientific evidence and methodological considerations.

# References


1. European Centre for Disease Prevention and Control. COVID-19 Situation update week 38, updated 24 September 2021. ECDC: Stockholm; 2021. Available at: https://www.ecdc.europa.eu/sites/default/files/documents/Communicable-disease-threats-report-24-september-2021.pdf
2. European Medicines Agency. COVID-19 vaccines: authorised. EMA: Amsterdam; 2021. Available at: 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
3. European Medicines Agency. First COVID-19 vaccine approved for children aged 12 to 15 in EU. 28 May 2021. EMA: Amsterdam; 2021. Available at: https://www.ema.europa.eu/en/news/first-covid-19-vaccine-approved-children-aged-12-15-eu
4. World Health Organization. Laboratory testing for 2019 novel coronavirus (2019-nCoV) in suspected human cases: Interim guidance. 19 March 2020. WHO; Geneva: 2020. Available at: https://www.who.int/publications-detail/laboratory-testing-for-2019-novel-coronavirus-in-suspected-human-cases-20200117
5. European Centre for Disease Prevention and Control. Operational considerations for influenza surveillance in the WHO European Region during COVID-19: interim guidance. 16 October 2020. ECDC; Stockholm: 2020. Available at: https://www.ecdc.europa.eu/en/publications-data/operational-considerations-influenza-surveillance-european-region-during-covid-19
6. European Centre for Disease Prevention and Control. Testing strategies for SARS-CoV-2. ECDC; Stockholm: 2020. Available at: https://www.ecdc.europa.eu/en/covid-19/surveillance/testing-strategies
7. Collin C, Wade DT, Davies S, Horne V. The Barthel ADL Index: a reliability study. Int Disabil Stud. 1988;10(2):61–3.
8. Huedo-Medina TB, Sánchez-Meca J, Marín-Martínez F, Botella J. Assessing heterogeneity in meta-analysis: Q statistic or I2 index? Psychol Methods. 2006 Jun;11(2):193–206.
9. National Institute for Biological Standards and Control. Coronavirus (COVID-19)-related research reagents available from the NIBSC. Available at: https://www.nibsc.org/science_and_research/idd/cfar/covid-19_reagents.aspx
10. Ozaras R, Cirpin R, Duran A, Duman H, Arslan O, Bakcan Y, et al. Influenza and COVID-19 coinfection: Report of six cases and review of the literature. J Med Virol. 2020 11;92(11):2657-65.
11. Antony SJ, Almaghlouth NK, Heydemann EL. Are coinfections with COVID-19 and influenza low or underreported? An observational study examining current published literature including three new unpublished cases. J Med Virol. 2020 11;92(11):2489-97.

# Bibliography

Blettner, M, Sauerbrei, W, Schlehofer, B, Scheuchenpflug T, Friedenreich C. Traditional reviews, metaanalyses and pooled analyses in epidemiology. Int J Epidemiol 28, 1–9(1999).

Lemeshow S, Hosmer D, Klar J. Sample size requirements for studies estimating odds ratios and relative risks. Stat Med 1988 ;(7) :759–64.

National Research Council. Combining information – statistical issues and opportunities for research. Washington, DC: National Research Council; 1992. doi: https://doi.org/10.17226/20865

Rodrigues L, Kirkwood BR. Case–control designs in the study of common diseases: updates on the demise of the rare disease assumption and the choice of sampling scheme for controls. Int J Epidemiol 1990 Mar;19(1):205–13.

Rothman, K, Greenland, S, Lash, T. Modern Epidemiology (ed.). Philadephia: Lippincott–Williams–Wilkins, 2008.

Smith-Warner, S.A, Spiegelman, D, Ritz, J, Albanes, D, Beeson, W.L, Bernstein, L. et al. Methods for pooling results of epidemiologic studies: the Pooling Project of Prospective Studies of Diet and Cancer. Am J Epidemiol 2006 ;163 :1053–64.

Stukel TA, Demidenko E, Dykes J, Karagas MR. Two-stage methods for the analysis of pooled data. Stat Med 2001 Jul 30;20(14):2115–30.

# Legislative references

ECDC Founding Regulation (EC) No 851/2004 of the European Parliament and of the Council of 21 April 2004 establishing a European Centre for disease prevention and Control

Decision No 1082/2013/EU of the European Parliament and of the Council of 22 October 2013 on serious cross-border threats to health and repealing Decision No 2119/98/EC

Council Recommendation of 7 December 2018 on strengthened cooperation against vaccine preventable diseases (2018/C 466/01)

Commission Implementing Decision (EU) 2018/945 of 22 June 2018 on the communicable diseases and related special health issues to be covered by epidemiological surveillance as well as relevant case definitions

Communication from the Commission to the European Parliament and the Council of 15 October 2020 on Preparedness for COVID-19 vaccination strategies and vaccine deployment, COM (2020)680 final

# Annex 1. List of ICD-9 and ICD-10 codes for pre-existing chronic conditions

| Category | ICD-9 | ICD-10 | Underlying conditions included |
|---|---|---|---|
| Anaemia | 280–285 | D50-64 | Nutritional anaemias, Haemolytic anaemias, Aplastic and other anaemias and other bone marrow failure syndromes |
| Asplenia | 746.87, 759.0 | Q89.01, Q20.6, Z90.81 | Malposition of heart, Anomalies of spleen, Isomerism of atrial appendages, Acquired and Congenital absence of spleen |
| Asthma | 493.0, 493.1, 493.9 | J45 | Extrinsic asthma, Intrinsic asthma, Predominantly allergic asthma, Non-allergic asthma, Mixed asthma, Asthma unspecified |
| Chronic liver disease | 571 | K70, K72-74, K754, K769 | Alcoholic liver disease, Hepatic failure, Chronic hepatitis, Fibrosis and cirrhosis of liver, Other inflammatory liver diseases |
| Cardiovascular diseases | 093, 112.81, 130.3, 391, 393–398, 402, 404, 410–429, 745, 746, 747.1, 747.49, 759.82, 785.2-3 | A52.01, B37.6, B58.81, I05-9, I11, I13, I20-25, I26.09, I26.9, I27, I30-51, I97.0-1, R00.1, T81.718A, T81.72XA, T82.817A, T82.818A, Q20-24, Q25.1-2, Q26.0-1, Q26.8, Q87.4, R01.1-2 | Syphilitic aneurysm of aorta, Candidal endocarditis, Toxoplasma myocarditis, Chronic rheumatic heart diseases, Ischemic heart diseases, Hypertensive heart and chronic kidney disease, pulmonary embolism with acute cor pulmonale, pulmonary heart diseases, diseases of pulmonary vessels, Other forms of heart disease (including Nonrheumatic valve disorders, pericarditis, endocarditis, myocarditis, cardiomyophathy, heart failure, block, cardiac arrhythmias, heart failure), Complication of other artery / vein following a procedure, Embolism of cardiac/vascular prosthetic devices, implants and grafts, congenital malformations of cardiac chambers and connections or heart, Coarctation or atresia of aorta, Congenital malformations of great veins, Marfan's syndrome, Cardiac murmur |
| Diabetes | 250 | E10-11 | Type 1 and Type 2 diabetes mellitus |
| Hypertension | 401, 401.0, 401.9, 405, 405.91, 405.99, | I10, I15.8, I15, I15.1, I15.2, I97.3, I27.0 | Hypertension (essential and secondary), Secondary to other [renal or endocrine] disorders, Malignant hypertension |

| Category | ICD-9 | ICD-10 | Underlying conditions included |
|---|---|---|---|
| Obesity | 27800, 278.01, 278.03 | E66.01, E66.2, E66.9 | Obesity |
| Immunodeficiency* or organ transplant | 042, 279, V08, V42 | B20, D80-84, D89.8-9, Z21, Z94 | HIV, immune deficiency, organ or tissue replaced by transplant |
| Neuromuscular disorders | 358.00-358.1, 358.8, 358.9, 378.73, 775.2 | G70-G70.01, G70.2, G70.80, G70.81, G70.9, G70.89, G73.7, | Myasthenia gravis, Myoneural disorders NEC/NOS, Neuromuscular disease strabism, Congenital and developmental myasthenia, Lambert-Eaton syndrome, Myoneural disorder NOS |
| Renal disease | 274.1, 408, 580–591, 593.71–593.73, 593.9 | M10.30, N00-19, N20.0, N28.9 | Gout due to renal impairment, Glomerular diseases, Renal tubulo-interstitial diseases, Acute kidney failure and chronic kidney disease, Calculus of kidney, Disorder of kidney and ureter, unspecified |
| Dementia | 290, 294, 331 | F01, F03, F05, G30, G31, G91, G94 | Vascular dementia, other dementia, Delirium due to known physiological condition, Alzheimer's disease, Other degenerative diseases of nervous system |
| Stroke | 348, 438 | G93, I67.83, I69 | Brain disorders, Posterior reversible encephalopathy syndrome, Sequelae of cerebrovascular disease |
| Rheumatologic diseases | 446, 710, 714 | M30-34, M35.0, M35.5, M35.8-9, M05-06, M08, M12.00 | Polyarteritis nodosa and related conditions, Other necrotising vasculopathies, Systemic lupus erythematosus (SLE), Dermatopolymyositis, Systemic sclerosis, Sicca syndrome, Multifocal fibrosclerosis, other systemic involvement of connective tissue, Rheumatoid arthritis with rheumatoid factor, Other rheumatoid arthritis, Juvenile arthritis, Chronic post-rheumatic arthropathy |
| Cancer | 140–208 | C00-96 | Malignant neoplasms and neuroendocrine tumours |
| Lung disease | 011, 490–511, 512.8, 513–517, 518.3, 518.8, 519.9, 714.81 | A15, J40–47, J60–94, J96, J99, J182, **M34.81, M05.10** | Respiratory tuberculosis, Bronchitis, not specified as acute or chronic, Chronic bronchitis, Emphysema, Other chronic obstructive pulmonary disease, Asthma, Bronchiectasis, Hypersensitivity pneumonitis due to organic dust, Pneumoconiosis, Airway disease due to specific organic dust, Hypersensitivity pneumonitis due to organic dust, Respiratory conditions due to inhalation of chemicals, gases, fumes and vapor, Pneumonitis due to solids and liquids, Respiratory conditions due to other external agents, Acute respiratory distress syndrome, Pulmonary oedema, Pulmonary eosinophilia, not elsewhere classified, Other interstitial pulmonary diseases, Abscess of lung and mediastinum, Pyothorax, Pleural effusion, Pneumothorax and air leak, Other pleural conditions, Intraoperative and postprocedural complications and disorders of respiratory system, not elsewhere classified, Other diseases of the respiratory system, Hypostatic pneumonia, unspecified organism, **Systemic sclerosis with lung involvement, Rheumatoid lung disease with rheumatoid arthritis** |
| Tuberculosis | | A15–A19 | Primary respiratory tuberculosis, Respiratory tuberculosis unspecified, Tuberculosis of nervous system, Tuberculosis of other organs, Miliary tuberculosis |

**Note: Patients who are only treated with glucocorticoids and have no other immune deficiency, are considered immune suppressed when treated with high-dose corticosteroids (≥ 20 mg/day of prednisone or equivalent for ≥2 weeks) in the last three months.*

# Annex 2. Detailed analysis plan

## Pooled analysis outline

A pooled analysis is part of the primary objectives of the ECDC study. Country will be included potentially as a fixed effect or as a random effect in a multilevel model. Statistical heterogeneity between study sites will be determined, using Q-test and the $I^2$ index. [8]

Briefly, cases and controls will be described by baseline characteristics, and uni- and multivariable analyses performed as described in section 3.14.1 for individual analysis.

## Pooled analysis plan

### Descriptive pooled analysis

The proportion of eligible hospitalised cases and controls who accepted to participate in the study will be calculated. The proportion of patients not consenting will be documented, along with reasons for no participation. Patients excluded will be described in a study flowchart.

Cases and controls will be described by baseline characteristics.

The main characteristics of each study will be summarised individually, including:

- Number of hospitals participating and catchment population
- Beginning of vaccination campaigns for pandemic vaccine
  - Beginning of the study
  - End of the study
  - Vaccine product(s) used
  - Estimated vaccine coverage in the country/region by vaccine brand, by target vaccine group
- Number of patients screened
- Number of patients excluded per reasons for exclusion.

## Measure of effect

This study is a case control study (test-negative design). The measure of association is an odds ratio (OR). This can be measured by logistic regression. An OR = 1 indicates no association between an exposure and the outcome. An OR>1 indicates a potential risk factor, an OR < 1 indicates a potential protective factor, noting that the confidence interval around the OR helps with its interpretation.

For vaccination as preventive factor, the CVE can be computed as CVE = (1 – OR)*100. A 95% confidence interval is computed around the point estimate.

### *Pooled univariable analyses*

Baseline characteristics of cases and controls will be compared using the chi-square test, Fisher's exact test, t-test or the Mann-Whitney test (depending on the nature of the variable and the sample size). The association (OR) between vaccination status and baseline characteristics will be measured for both case and control groups.

### *Stratified analysis*

The analysis by vaccine product will be further stratified according to (depending on sample size):

- sex
- age groups, e.g. 0–14 years, 15–49 years, 50–64 years, 65–79 years, 80+ years
- specific chronic conditions (e.g. respiratory, diabetes, obesity)
  - absence, presence of at least one, presence of more than one high-risk condition
- time: this will depend on timing of the pandemic in sites/countries and may just include one period at the start of the study once vaccines are available, and a specified period later on
- swab delay (0–3 days, 4–7 days; 8+ days)
- vaccination delay (<8 days, 8–14 days, >14 days, etc.)
- hospital admission delay (0–4 days, 5–9 days, 10 days +, onset after hospitalisation)
- previous vaccination against influenza and pneumococcal disease
- prior infection with influenza or COVID-19 (prior to hospital admission for SARI)
- current co-infection with influenza or other respiratory viruses
- severity (ICU admission, ventilation/oxygen, death)
- for the various groups of vaccines (if available/applicable), mode of injection (intradermal vs intramuscular)

- use of medications for chronic conditions (e.g. statins)
- Other stratifications may be included.

Virus variant-specific outcomes will be used, if available and feasible at the time of analysis.

A sufficient sample size should be planned in order to ensure enough individuals in each stratum for a precise estimate. Effect modification will be assessed comparing the OR across the strata of the potential effect modifiers. Confounding will be assessed by comparing crude and adjusted OR for each potential confounder.

## *Multivariable analysis*

A multivariable logistic regression analysis will be conducted to control for negative and positive confounding. Odds ratios and standard errors will be obtained. Variables will be tested for multicollinearity. Interactions will be tested using the likelihood ratio test or Wald's test and will be included in the model if significant at the 5% level. Factors other than statistical significance (prevalence of exposure, magnitude of OR) will also be used as criteria for inclusion of a variable or an interaction term. If possible, a variable for sex, age and for onset time should always be included in the model.

**Continuous variables**
Continuous variables in the COVID-19 datasets include age, time of onset of symptoms, GP visits in the previous three months and hospitalisations in the past 12 months. These variables can be coded as categories, e.g. age group, week of symptom onset, etc. However, when coding continuous variables as categories, information may be lost, introduce residual confounding and increase the standard error of the model. Tests will be carried out to see if these variables could be coded as a linear term, polynomial or a spline. In addition, a balance will be sought between simplicity of a model (so a non-expert can understand it), precision and a model that estimates the vaccine effect with the least bias.

**Identifying heterogeneity, testing for heterogeneity**
Country-specific crude and adjusted ORs and their confidence intervals will be plotted in separate forest plots. Following the core protocol minimises heterogeneity between studies. However, adherence to the protocol and study design and study quality characteristics will also be checked. Other study site characteristics will be assessed where feasible, such as types of circulating virus, information on health care use, organisation of the vaccination campaign. Then a qualitative decision will be taken if one or more studies are substantially different from the other and should be excluded from the pooled analysis.

Statistical heterogeneity between studies will be tested using Q-test and the $I^2$ index (see boxes for formulae below). The Q statistic follows a $Chi^2$ distribution (with k-1 degrees of freedom). The Q-test reports presence or absence of heterogeneity, while the $I^2$ index (based on the Q-statistic) quantifies the extent of the heterogeneity. According to the Higgins and Thompson classification, an $I^2$ index of around 25% indicates low, 50% indicates medium and 75% indicated high heterogeneity between studies.

$$Q = \sum w_i \ (\log(OR_i) - \log(OR_F))^2$$

Where:

$w_i = 1 / v_i$

$v_i$ is the inverse variance of the estimated log odds ratio of study *i*

$$\log(OR_F) = \frac{\sum w_i \times \log(OR_i)}{\sum w_i}$$

$$I^2 = \frac{Q-(k-1)}{Q} \times 100\% \quad \text{for } Q > (k-1)$$

$$I^2 = 0 \quad \text{for } Q \leq (k-1)$$

Formulae are given here for completeness, in practice these measures are automatically calculated by many statistical software packages as part of the meta-analysis commands.

## One-stage pooled analysis approach

If sample sizes are too small to measure vaccine effectiveness controlling for all potential confounders for each individual study site, a one-stage pooled approach will be used for analysis.

Individual study data will be pooled into one dataset and analysed as a one-stage model with study site as a fixed effect. This could provide a large enough sample size to obtain (for example) an estimate of CVE early in the study with reasonable precision. The results of this analysis should be interpreted with caution, though, as it assumes not only that the underlying true exposure effect is the same in all studies, but also that the association of all covariates with the outcome is the same in all studies.

Formal tests of interaction between study site and covariates will be carried out to determine if the effect of each covariate differs across studies, to test the assumptions of the one-stage pooled fixed effect analysis.

The significance of interaction terms are themselves influenced by sample size and should be interpreted also with caution. Particular care needs to be taken if heterogeneity is found between study sites when using a one-stage fixed effects approach (see above section). Reasons for heterogeneity need to be thoroughly investigated and the assumptions underlying the one-stage pooling approach need to be revisited.

## Controlling for hospital effect

Primary analysis will be carried out using simple logistic regression to obtain the individual study estimates. However, there could be an effect of the hospital that is related both to the exposure (propensity to vaccinate) and the outcome (in terms of swabbing behaviour). To adjust for this cluster effect, a multi-level logistic regression with each hospital as a random effect will be carried out when using one-stage pooled analysis.

Multi-level logistic regression can also be carried out for each individual study with hospital as a random effect. Then the two-stage model as outlined above will be used to obtain a summary CVE measure, using these estimates.

The same applies to stratified analyses. The point estimates and confidence intervals from the multi-level and simple logistic regression will be compared in a sensitivity analysis.

## Two-stage pooled analysis approach

If adequate sample size by study is achieved to obtain an adjusted OR, then a two-stage approach to pooled analysis will be taken.

Country-specific adjusted ORs and standard errors for the effect of COVID-19 vaccination obtained from the individual studies, will be combined in a model that incorporates random effects of the studies, to account for unmeasured country- and hospital-specific factors that differ between countries.

The country-specific exposure-disease effects (ORs) are then weighted by the inverse of their marginal variances. The marginal variance is the sum of the individual study-specific variances and the variance of the random study effects (τ2). This will give the pooled odds ratio and standard error.

$$\log(OR_R) = \frac{\sum w_i^* \times \log(ORi)}{\sum w_i^*}$$

$$wi^* = \frac{1}{vi + \tau^2}$$

The country-specific ORs and their confidence intervals, along with the pooled OR, will be presented graphically in a forest plot. This model will also be compared against a two-stage analysis with fixed study effects, to assess the effects of model assumptions.

If, despite the common protocol, covariates were not uniformly collected in the different studies, then an analysis will be carried out excluding certain studies and a comparison to the analysis including all studies will be made. In a different scenario, analyses can also be carried out excluding certain study participants for whom variables were collected differently.

## Further analyses

Where sample size allows, further analyses will be carried out. These include:

- CVE at different time points in calendar time, e.g. CVE by week or group of weeks (e.g. CVE for weeks 2–3, 4–5, 6–7, etc.)
- CVE by time since vaccination. Time since vaccination can be calculated by subtracting the date of vaccination from the date of onset. Time since vaccination can then be modelled as a continuous variable,

including correction for either stable or increased rate of COVID-19 illness over time; cumulative risk of COVID-19 illness.

- CVE for patients with previous influenza vaccination (current influenza season) vs no previous influenza vaccination.
- If negative CVE is found in some target groups
  - assess possibility of vaccine-mediated enhanced disease (VMED), which could manifest as negative CVE, by comparing severity in vaccinated and unvaccinated patients. Results should show reduced severity among vaccinated patients; findings of increased severity in vaccinated patients could suggest VMED
- As a sensitivity analysis, CVE will be calculated
  - considering those vaccinated <X days before onset of symptoms as unvaccinated (in the main analysis these records will be excluded)
  - including in the control group, SARI patients testing positive for influenza
  - including in the case group, SARI patients testing positive for influenza
  - including in the control group, SARI patients whose influenza vaccine status is unknown
  - using, as a control group, only SARI patients testing positive for at least one non-influenza respiratory virus
  - considering different restrictions according to swabbing delay (e.g. <14 days, <10 days, etc.)
  - considering the sensitivity and specificity of PCR
  - based on assumptions of previous infections
  - excluding participants who received antivirals ≤14 days prior to swabbing
  - excluding all participants with lab-confirmed influenza at any time after COVID-19 onset, to reduce bias.
- This can then be repeated using RSV as a sham outcome (if multiplex results are available for any sites); there should be no association between COVID-19 vaccination and RSV-positivity in the absence of confounding.

Time can be input as a variable in the model and assess whether it can be an effect modifier.

# Minimum sample size

Sample sizes may be very small for some sub-analyses. Different criteria can be used to determine whether the sample size is large enough to obtain a valid measure of CVE:

- There are at least 10–15 cases (or controls, whichever is smaller) in the sub-analysis for crude analyses and more for adjusted analyses (e.g. at least 10 for each parameter in the model)
- There are ≥5 records in each cell of the two-by-two table of case and vaccination status
- The precision of the estimate does not span both -200% and 90% (uninformative).

With low sample size, approached include collapsing categories, modelling continuous variables in a different way (if applicable); sensitivity analyses can be carried out using penalised logistic regression.

# Use of propensity scores

To limit the number of co-variables to include in the multivariable model, **if sample size allows,** estimates will be built and adjusted based on propensity scores. Propensity scores can be defined as the conditional probability of receiving the vaccine given a number of observed covariables.

In propensity score matching, a propensity score for vaccination is calculated for cases and controls. Cases and controls are then matched by propensity score and all non-matched patients are discarded. Variables used to calculate the propensity score will include variables related to the vaccination and outcome. Care will be taken to avoid correlation and overmatching.

# Annex 3. Genetic characterisation in vaccine effectiveness studies

## Virus selection

### Objectives

- To describe the viruses included in the ECDC vaccine effectiveness (VE) studies (overall, by site and by time period), in order to identify key SARS-CoV-2 virus genotypic evolution that could affect vaccine effectiveness.
- To measure genetic variant-specific COVID-19 VE among study sites

In order to meet these objectives, it is important that the viruses sequenced are representative of the SARS-CoV-2 viruses from cases belonging to the VE study population. In order to achieve this, either all or a random selection of viruses are sequenced (among those that are technically feasible to sequence). This way, the selection is independent from vaccination or clinical outcomes.

Study sites/country will select viruses from SARI patients included in the VE studies testing positive for SARS-CoV-2. If feasible, before virus selection, study sites/countries should verify if the SARS-CoV-2 positive cases meet the criteria used to include cases in the VE pooled analysis (e.g., target group for vaccination, vaccination status and date documented, delay symptom onset swabbing not more than 10 days, etc.).

### Proportion of SARS-CoV-2 viruses to characterise: sampling fraction

Each study site/country has different resources, different incidence and different proportion of genetic variants circulating.

Ideally, a study site will sequence all viruses that are technically feasible to sequence. If this is not feasible, then the proportion sequenced (the sampling fraction) will be based on the study site resources and on the epidemiological/virological situation. We suggest that study sites/countries sequence at a minimum 50% of viruses that are technically feasible to sequence, to be reviewed along the course of the pandemic.

If study sites are sequencing a proportion of viruses (as opposed to all viruses), the sampling fraction can change over time, depending on resources. For example, during peak incidence only 50% of viruses could be sequenced and during periods of low incidence all viruses could be sequenced. Important is than when sequencing a proportion of viruses, the random selection process is adhered to.

The proportions sampled over time should be documented in the 'Example of sampling fraction definition' Excel spreadsheet. An example is in Figure A1.

**Figure A1. Example of how to define sampling fractions over time using the Excel spreadsheet 'Example of sampling fraction definition.xlsx'**

| Time period | First date of time period | Last date of time period | Sampling fraction used | Date used for definition of time unit (onset date, swab date, other) | Comments |
|---|---|---|---|---|---|
| 1 | *01/01/2021* | *31/01/2021* | *1* | *Date of swab* | *All specimens were characterised among those technically feasible* |
| 2 | *01/02/2021* | *30/04/2021* | *0.5* | *Date of swab* | *50% of specimens were characterised among those technically feasible* |
| 3 | | | | | |

- *Each study site/country to*
  - *define the sampling fraction used for each time interval (if all viruses characterised, then indicate 100%);*
  - *document the sampling fraction for each time interval in the 'Example of sampling fraction definition.xlsx' document.*
  - *take the sampling fraction into account when measuring site-specific genetic variant-specific VE (if applicable).*

### Procedures for random selection of specimens to be characterised

If a study site/country is not genetically characterising all viruses, then the random selection proposed should be used to select the viruses included in the multicentre study. This should be done independently from any routine virological surveillance.

As it is difficult to prospectively randomise viruses to sequence, study sites can use a list of viruses by a predefined period (e.g., each week or month) to use for randomisation.

At the end of each period defined for the selection of strains, study sites will select viruses using random selection (e.g., the Bernoulli sampling method). This method ensures that each strain has the same probability of being selected.

- *Each site to define who selects the strains to sequence (e.g., team of epidemiologists, team of virologists)*

## Steps to randomly select the strains

**Step 1: Sampling frame**

- Create a list of all positive cases recruited for the period study sites/country would like want to sequence(week/month).
- Viruses already characterised by the National Reference Centre (or other laboratories) during that period will be part of the sampling frame to ensure representativity.
- Viruses with low viral load will be part of the sampling frame.
- If possible, sites/countries will exclude from the sampling frame the viruses from cases that would later be excluded from any pooled analysis (e.g. target group for vaccination, vaccination status and date documented, delay symptom onset swabbing not more than 10 days, etc).

**Step 2: Randomisation**

- List order all positive cases (viruses) by swab date.
- Assign a random number to each virus: The Excel function =RAND() can be used [this may be different if using a different language version of Excel].
- Copy the random number column and paste as values (this is important, otherwise they will keep changing) and then sort the list of cases (viruses) by random number in the pasted column (e.g. in order of high to low).

The below STATA syntax can be used. The example is selection of 50% of 88 cases (44 strains to characterise).

- *sort swabdate*
- *set seed 500*
- *gen naleat=runiform(),*
- *sort naleat*
- *gen select=0*
- *replace select=1 if _n<=44*
- *list IDnumber strain select if select==1, noobs separator (44)*

where 500 is the number used to set the seed. You can select another number but it is recommended to use the 'set seed' to be able to replicate the selection.

**Step 3: Selection**

- Based on the proportion of viruses sites/countries would like to sequence for this period (e.g. 50% or 75%), select the number of cases/viruses needed: start from the first case in the list and continue selecting the following cases until reaching the desired number (e.g. if 88 cases have been recruited in the study and the sampling fraction is 0.5, the first 44 cases in the list will be selected).

**Step 4: Replacement of viruses randomly selected, but not characterised**

- Viruses that cannot be sequenced should be replaced using the same subtype/lineage sampling frame. The strains will be replaced by the next ones in the list. So, for example, if one would have chosen 44 viruses from the randomised list and two are not feasible to sequence, then one can select the viruses in the 45th and 46th line to sequence. The reasons for not sequencing selected specimens should be documented (e.g., low viral loads) and all study sites should document their Ct threshold for sequencing (if applicable).

**Step 5: Increase in proportion sequenced if needed**

- If during the pandemic, the study site decides to increase the proportion sequenced for a given time period, the study site should go back to the sampling frame for that time period and continue selecting the following strains from the ordered original list.

## *Data collected*

For the viruses characterised, study sites/countries should fill in an Excel spreadsheet as shown below, with at least the following information available (see Figure A2 for an example):

- Country;
- Patient's study ID number;
- GISAID sequence database accession number;
- Selected for characterisation? (Y/N);
- Reasons for not characterising;
- If possible: Ct value;
- If possible: Type of sample (primary specimen or isolate).

**Figure A2. Example of information collected on viruses using an Excel spreadsheet**

| Country | ID number | Reasons for no characterisation (for those selected not characterised) | Ct value clinical specimen | GISAID Accesion ID |
|---|---|---|---|---|
| Spain | 2016128 | Sequenced | 20.48 | EPI_ISL_691732 |
| Spain | 2016451 | No product due to low viral load | 31.52 | N/A |

Screenshot of Excel spreadsheet to collect information on proportions sequenced over time (sampling fraction):

| Time period | First date of time period | Last date of time period | Sampling fraction used | Date used for definition of time unit (onset date, swab d | Comments |
|---|---|---|---|---|---|
| 1 | | | | | |
| 2 | | | | | |
| 3 | | | | | |
| 4 | | | | | |
| 5 | | | | | |
| 6 | | | | | |
| 7 | | | | | |
| 8 | | | | | |
| *Example1* | *01/01/2021* | *31/01/2021* | *1* | *Date of swab* | *(this is only an example;* |
| *Example2* | *01/02/2021* | *30/04/2021* | *0.2* | *Date of swab* | *(this is only an example;* |

**Figure A3. Example of sampling fraction definition**

Screenshot of Excel spreadsheet to collect information on proportions sequenced over time (sampling fraction):

| Country | ID number | Reasons for no characterisation (for those selected not characterised) | Ct value clinical specimen | GISAID Accesion ID |
|---|---|---|---|---|
| Spain | 2016128 | Sequenced | 20.48 | EPI_ISL_691732 |
| Spain | 2016451 | No product due to low viral load | 31.52 | N/A |

# Genetic and antigenic analysis data (example)

The minimum amount of data needed to obtain genetic data from GISAID (sequences of all viruses should be sent to GISAID's open access EpiCoV platform) is country, ECDC SARI VE study patient ID number and GISAID accession number. Additional information on Ct value and selection for characterisation and reasons for not characterising can be additionally collected (see Table A1 below).

**Table A1. Example of a data collection form for genetic data**

| | Country | Study ID number | GISAID accession ID number | Selected for character-isation? | Reasons for not character-ising? | Ct value | Type of sample (primary specimen or isolate) |
|---|---|---|---|---|---|---|---|
| *strain1* | | | | | | | |
| *strain2* | | | | | | | |

Where not all viruses were attempted to be sequenced, but only a random selection of them, additional information on sampling fraction should be provided, in order to better understand how viruses were selected for sequencing over time. An example can be seen in Table A2 below.

**Table A2. Example of documenting outlining how viruses were selected for sequencing over time**

| Period | First date of period | Last date of period | Sampling fraction | Date used for definition of time unit (onset date, swab date, other) | Comments |
|---|---|---|---|---|---|
| 1 | | | | | |
| 2 | | | | | |
| *Example1* | *01/10/2020* | *31/12/2020* | *1* | *Date of onset* | *For example: all specimens were characterised* |
| *Example2* | *01/01/2021* | *15/02/2021* | *0.2* | *Date of onset* | *For example: 20% of all specimens were characterised* |

**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

Publications Office
of the European Union

PDF ISBN 978-92-9498-530-9