World Health Organization
REGIONAL OFFICE FOR Europe

Estimating COVID-19 vaccine effectiveness against severe acute respiratory infections (SARI) hospitalisations associated with laboratory-confirmed SARS-CoV-2

An evaluation using the test-negative design

Guidance Document

Document number: WHO/EURO:2021-2481-42237-58308

Suggested citation. Estimating COVID-19 vaccine effectiveness against severe acute respiratory infections (SARI) hospitalisations associated with laboratory-confirmed SARS-CoV-2. An evaluation using the test-negative design: guidance document. Copenhagen: WHO Regional Office for Europe; 2021. Licence: CC BY-NC-SA 3.0 IGO.

Cataloguing-in-Publication (CIP) data. CIP data are available at http://apps.who.int/iris.

Contents

Acknowledgements

The WHO Regional Office for Europe led the development of this document with Angie Rose (Epiconcept), Pernille Jorgensen, Mark Katz, and Richard Pebody (WHO Regional Office for Europe). The WHO Regional Office for Europe convened an ad-hoc expert review committee that provided strategic direction and direct inputs into the protocol. The WHO Regional Office for Europe would like to thank the following members of the review committee (in alphabetical order):

1. Nick Andrews, Public Health England, United Kingdom
2. Melissa Arvay, US Centers for Disease Control and Prevention, Atlanta, GA, USA
3. Sabrina Bacci, European Centre for Disease Prevention and Control, Stockholm, Sweden
4. Ed Belongia, Marshfield Clinic, Wisconsin, USA
5. Alessandro Cassini, World Health Organization, HQ, Geneva, Switzerland
6. Ben Cowling, Hong Kong School of Public Health, Hong Kong
7. Daniel Feikin, World Health Organization, HQ, Geneva, Switzerland
8. Wiebke Hellenbrand, Robert Koch Institute, Germany
9. Kari Johansen, European Centre for Disease Prevention and Control, Stockholm, Sweden
10. Esther Kissling, Epiconcept, Paris, France
11. Arnold Monto, University of Michigan School of Public Health, Michigan, USA
12. Alain Moren, Epiconcept, Paris, France
13. Francisco Nogareda, Pan American Health Organization, Washington DC, USA
14. Hanna Nohynek, THL, Helsinki, Finland
15. Pasi Penttinen, European Centre for Disease Prevention and Control, Stockholm, Sweden
16. Sheena Sullivan, Doherty Institute, Melbourne, Australia
17. Marta Valenciano, Epiconcept, Paris, France
18. Marc-Alain Widdowson, Institute of Tropical Medicine, Antwerp, Belgium

This publication has been supported by WHO Unity Studies, a global sero-epidemiological standardization initiative.

Abbreviations

COVID-19	Coronavirus disease 2019
CVE	COVID-19 vaccine effectiveness
EEA	European Economic Area
ECDC	European Centre for Disease Prevention and Control
EU	European Union
GP	General practitioner
ICD	International classification of diseases
ILI	Influenza-like illness
I-MOVE	Influenza – Monitoring Vaccine Effectiveness in Europe
MS	Member States
OR	Odds ratio
RF	Risk factor
RT- PCR	Real-time polymerase chain reaction
SARI	Severe acute respiratory infection
SARS-CoV-2	Severe acute respiratory syndrome – coronavirus 2
VE	Vaccine effectiveness

Executive Summary

Many critical questions remain about the effectiveness of COVID-19 vaccines in real-world settings. These questions can only be answered in post-introduction vaccine effectiveness studies.

This guidance document outlines an approach to leverage existing surveillance systems for Severe Acute Respiratory Infection (SARI) to estimate COVID-19 vaccine effectiveness (VE) in preventing SARI associated with laboratory-confirmed SARS-CoV-2 using existing SARI surveillance systems. The approach uses the test-negative design to evaluate VE; cases are SARI patients who tested positive for SARS-CoV-2, and controls are SARI patients who tested negative for SARS-CoV-2.

As in current SARI surveillance, hospitalized patients with SARI at designated SARI surveillance hospitals can be enrolled if they meet the WHO case definition for SARI. A respiratory specimen should be collected from SARI patients within 48 hours of admission to the hospital. In addition, data can be collected about the patient's sociodemographic characteristics, history of acute illness, clinical course in hospital, and the COVID-19 vaccine history.

The document proposes that enhanced SARI surveillance to allow for VE estimates be implemented for a minimum of 6 months but ideally longer. The methods described in the guidance document are also appropriate for estimating influenza VE. The document also proposes that a pooled data analysis be conducted for the WHO/Europe region, which requires that SARI surveillance systems collect a minimum data set of similar variables.

1. Background

Since the emergence of the novel severe acute respiratory syndrome – coronavirus 2 (SARS-CoV-2), which causes coronavirus disease 2019 (COVID-19), in late 2019, the COVID-19 pandemic has resulted in more than 140 million cases and 3million deaths worldwide *(1)*.

After the influenza A(H1N1)pdm09 pandemic in 2009, a number of WHO/Europe Member States implemented sentinel surveillance for severe acute respiratory infections (SARI) for influenza. In these surveillance systems, case-based data (on all or a subset of patients) is systematically collected to monitor the severity and the burden of disease, identify viruses associated with severe clinical presentations, and determine risk factors for severe illness. These systems have also been used as platforms to estimate vaccine effectiveness (VE) against influenza-associated SARI hospitalizations.

WHO has advised that countries with existing hospital-based sentinel influenza surveillance adapt these systems to also monitor severe SARS-CoV-2 cases *(2)*, and collect data that allows for measuring COVID-19 VE when possible. Systematic SARS-CoV-2 testing and data collection from hospitalized SARI patients, when done as part of routine surveillance, allows for the post-licensure evaluation of COVID-19 VE in preventing severe illness. Stakeholders in the results of a SARI surveillance system that can evaluate COVID-19 VE include ministries of health, public health practitioners, clinicians, and international public health organizations.

Evaluating the performance COVID-19 vaccines post-licensure is critical as a number of factors can impact real-world VE, including variability in cold chain during transportation and storage, how vaccines are administered, advanced age and presence of underlying medical conditions, and previous SARS-CoV-2 infection. In addition, post-licensure evaluations of COVID-19 vaccines will allow public health authorities to a) understand the duration of protection of vaccines and thus the need (and frequency) for re-vaccination, b) estimate the level of protection against severe disease and death c) assess the relative effectiveness of different vaccines types and of single doses, and d) evaluate VE of new emerging virus variants.

This guidance document proposes enhanced data collection within existing SARI surveillance systems. The document describes how existing hospital-based surveillance systems for SARI can be adapted to collect data to inform estimates of vaccine effectiveness against COVID-19 in persons of all ages. This document is intended to be used for developing site-specific protocols. It outlines in details methods for collecting and analysing data on vaccinated and unvaccinated patients based on the “test negative design” (TND), where cases are SARI patients who test positive for SARS-CoV-2, and controls are SARI patients who test negative for SARS-CoV-2.

Since this document only outlines general principles, country-specific guidelines related to COVID-19 SARI surveillance to allow estimation of VE should be detailed in site-specific guidelines or protocols. This guidance document should ideally be used to complement existing SARI sentinel surveillance systems operated on a national level by the Ministry of Health or similar public health agency. SARI surveillance work on vaccine effectiveness should be conducted in accordance with the WHO guidelines regarding the ethical issues in public health surveillance *(3)*. Formal ethical committee approvals or waivers should be obtained

similar to how they were obtained for existing SARI surveillance systems. Surveillance procedures should protect the interests of patients and mitigate any foreseeable risks and harms to patients.

2. Objectives

2.1 Primary objective

The primary objective is to measure overall COVID-19 vaccine effectiveness (CVE) against laboratory-confirmed SARS-CoV-2 in hospitalised SARI patients belonging to the target group(s) for COVID-19 vaccination.

Site-specific protocols or guidelines should include language that describes which kinds of COVID-19 patients are routinely hospitalized according to national or local Ministry of Health guidelines or hospital policy.

2.2 Secondary objectives

- To estimate overall and product-specific CVE against laboratory-confirmed SARS-CoV-2 requiring hospitalisation among SARI patients in COVID-19 vaccine target groups by:
  - age group
  - sex
  - risk group (e.g. specific chronic conditions)
  - time since vaccination
  - in persons who have received one dose of vaccine compared to those who have received two doses
  - specific SARS-CoV-2 genetic variant(s)
- To estimate CVE for more severe outcomes (e.g. ICU admission, or in-hospital mortality)
- To compare CVE for different kinds of vaccines
- To identify potential factors that may modify CVE, including the following:
  - chronic medical conditions
  - living in a long-term care facility (LTCFs)
  - receiving statins or other long-term medications
  - degree of occupational or community exposure to SARS-CoV-2
  - previous SARS-CoV-2 infection.
- The methods described in this document can also be used to estimate influenza vaccine effectiveness

3. Methods

3.1 Harnessing existing SARI surveillance systems to estimate CVE

- CVE analysis is based on information gathered via existing SARI surveillance systems. The analysis is based on a test-negative, case–control design, which has been used in the past 10 years for estimating annual influenza VE. The principle behind this design is to evaluate SARS-CoV-2 laboratory results among persons who meet the standard SARI case definition and categorize those who test positive for SARS-CoV-2 as "cases" and those who test negative as "controls".

3.2 SARI surveillance population

- The surveillance population will consist of individuals of all ages who are hospitalised with SARI symptoms in a hospital that is part of the SARI surveillance network. This population has been targeted because of the interest in understanding the effectiveness of COVID-19 vaccines in preventing severe illness, and because of the interest and convenience in leveraging existing SARI surveillance systems to generate COVID-19 VE estimates.
- The CVE analysis should only include SARI patients who belong to a country's target group(s) for COVID vaccination and have no contra-indications to COVID-19 vaccination.

3.3 Surveillance period for CVE analysis

The relevant surveillance period should begin when COVID-19 vaccines are available in the country. It should continue until transmission levels decline to a low level for a sustained period.

Due to the constantly evolving COVID-19 situation (globally and locally) and the unpredictability of the intensity of transmission at the time when vaccines become available in 2021, it is recommended to plan for at least a 6-month surveillance period, although this period would ideally be longer.

3.4 Outcome

The outcome of interest for the primary analysis will be SARS-CoV-2 detection in patients of all ages hospitalised with SARI symptoms. SARS-CoV-2 infection should be laboratory-confirmed by PCR documented either within 48 hours of admission to hospital or within 14 days before admission.

Secondary outcomes include genetic variants of SARS-CoV-2 in hospitalised SARI patients of all ages. Evaluating this outcome will likely depend on capacity for genetic sequencing.

Additional outcomes include the following markers of disease severity during hospitalization:

- Length of stay (LOS)
- Use of supplemental oxygen
- Intensive Care Unit (ICU) admission
- Mechanical ventilation

- In-hospital death
- Clinical signs of pneumonia
- Severe respiratory rate > 30 breaths/min
- Oxygen saturation <90% on room air
- Severe respiratory distress
- Acute respiratory distress syndrome (ARDS),
- Sepsis
- Septic shock

3.5 SARI Case definition

3.5.1 SARI Case definition

A SARI patient will be defined using the WHO SARI case definition(4) as follows:

- a hospitalised person with acute respiratory infection, with
- a history of fever or measured fever of ≥ 38 C°
- **and** cough
- with symptom onset within the last 10 days.

Where a minimum of 24 hours in-hospital is required to be considered hospitalized.

3.5.2 Inclusion criteria

- Meets the SARI case definition
- Part of a target group the COVID-19 vaccine on the date of hospital admission
- No contraindications for COVID-19 vaccination

3.5.3 Exclusion criteria

SARI patients will not be enrolled in the CVE analysis if one of more of the following conditions are met:

- has a contraindication for the COVID-19 vaccine
- is ineligible for COVID-19 vaccine at the time of hospital admission (i.e. does not belong to a vaccine target group)
- 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 prior to this admission (including transfers from another hospital).

3.6 SARI patient identification – algorithm for patient inclusion

SARI patients will be identified among patients hospitalised for at least 24 hours in one of the participating hospitals, according to current country protocols for SARI sentinel surveillance. Patients who meet the inclusion criteria should be enrolled and have a respiratory specimen collected within 48 hours of hospital admission.

3.6.1 Recruitment strategies

Recruitment strategies should follow current recruitment strategies for existing SARI surveillance. In one approach, surveillance team members, which could include doctors, nurses or other health professionals, can approach patients to participate in surveillance. Like in routine SARI surveillance, it should be made clear to patients which aspects of data collection are related to surveillance and which aspects are related to their clinical care in the hospital.

For SARI surveillance systems that include electronic patient records where ICD codes are routinely used, potential SARI patients can be identified for further screening using relevant ICD codes (Table 1). (Figure 1).

Any patient-related information that is collected from hospital medical records, vaccine registries, or other sources should be collected in accordance with procedures outlined and approved in the protocol approved or granted waiver by the local ethical review committee, in accordance with WHO guidelines on ethical issues in public health surveillance.

For hospitals where ICD codes at admission are not systematically collected or accessible, systematic screening of all patients admitted can be conducted (Figure 1).

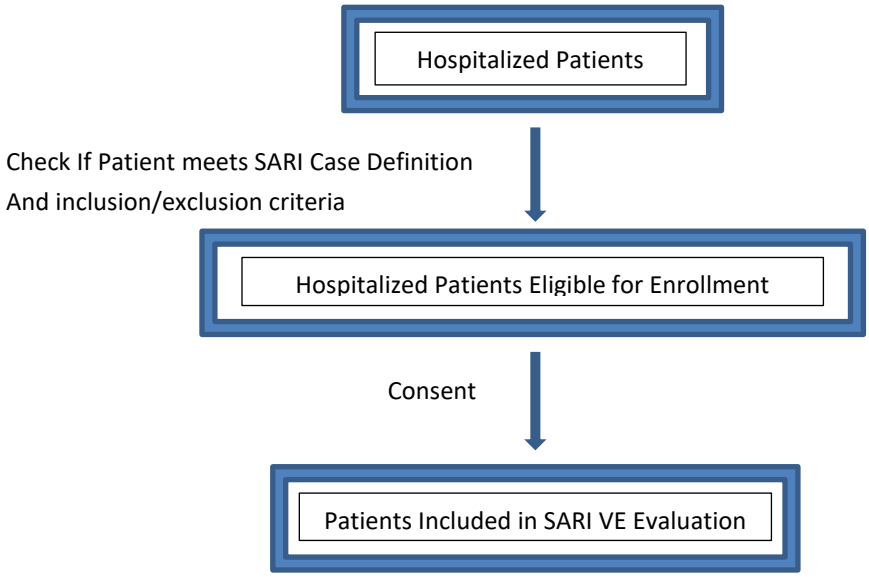


Figure 1: Algorithm for screening SARI patients

Table 1. List of diagnosis codes for which patients could be screened for presence of SARI symptoms according to the case definition.

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
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
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
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
Other	Anosmia, ageusia, myalgia	781.1, 729.1	R43.0, R43.2, M79.1

**SIRS: Systemic inflammatory response syndrome*

Following local procedures for SARI sentinel surveillance, patients meeting the SARI case definition will be asked to consent to participate in the SARI surveillance, according to local ethical requirements for informed consent. Patients who meet the SARI case definition should have a nasal, nasopharyngeal and/or throat, or

other (e.g. saliva) specimen taken for SARS-CoV-2 testing. In participating SARI surveillance hospitals, respiratory samples may be routinely collected from all hospitalized patients admitted with respiratory symptoms and tested for SARS-CoV-2, and therefore additional specimens for surveillance purposes will not need to be collected.

If resources are limited, SARI patients can be identified only on certain days of the week (e.g. recruitment of patients every second day, or only on 2 days of each week). Specific approaches to SARI patient recruitment should be planned in advance and well documented. Convenience sampling is not recommended as this can introduce bias.

Additional symptoms can be collected from SARI patients depending on the goals of the surveillance system. Evaluation could include symptoms listed as part of the most recent WHO COVID-19 case definition. *(5)*

3.7 Laboratory methods

As is the practice for routine SARI surveillance systems, trained health personnel should collect respiratory specimens (see Section 4.4) from all eligible patients, respecting safety standards for COVID-19 and following WHO biosafety guidelines *(6)*.

Samples should be tested by RT-PCR for SARS-CoV-2.

According to national testing guidelines and capacity, a systematic sample of PCR-positive samples should undergo genetic sequencing. If only a subset of the sample is sequenced, efforts should be made to ensure that these samples are representative of the SARI cases included in the surveillance *(7)*.

Genetic sequences, if sequencing is performed, should also be uploaded to GISAID's publicly accessible EpiCoV platform. Processed genetic information, such as the name of genetic clade and the source of the specimen (from sentinel or non-sentinel surveillance), can also be included within the epidemiological database.

Biological materials and related data should only be collected and stored in collaboration with hospitals and health authorities according to current procedures for SARI surveillance systems.

Most respiratory samples that will be collected for SARI surveillance to inform vaccine effectiveness estimates will be done so as part of routine clinical practice. It should be explained that management of these samples should be in accordance with hospital policies on management of clinical samples. For all respiratory samples that are collected for surveillance and not for clinical purposes, an explanation should be given to the participant about how these samples should be managed; policies should be in accordance with existing protocols for management of samples related to SARI surveillance.

Results of PCR tests for SARS-CoV-2 should be shared with participants as soon as they are available. Because the surveillance is being carried out within hospitals, dissemination of results should be carried out in accordance with hospital policies.

3.8 Vaccination history

3.8.1 COVID-19 vaccination status ascertainment

For every SARI patient, dates of vaccine administration, type of vaccine and brand name, and batch code for every dose of COVID-19 vaccine should be documented. The source of documentation of vaccine history should be recorded as well.

The sources of information for COVID-19 vaccination status and date of vaccination(s) may include:

- vaccination registry (preferred option)
- consultation of the patient's vaccination card or patient's hospital notes
- interview with the patient's GP
- interview with facility that administrated the vaccine
- data from the patient's insurance company showing evidence of delivery or re-imbursement for COVID-19 vaccine

Patient self-report should be discouraged as a sole source of information for vaccination status and vaccine details.

3.9 Additional information to collect from SARI patients

3.9.1 Pre-existing chronic conditions

Information about underlying medical conditions which are risk factors for severe COVID-19 infection should be collected (See Table 2 below).

Table 2. List of underlying medical conditions that have been identified as risk factors for severe COVID-19 infection

Chronic Medical Condition
Cancer*
Chronic cardiac disease, except hypertension*
Hypertension
Chronic kidney disease*
Chronic liver disease*
Chronic respiratory disease*
Asthma
Diabetes*
Immunocompromised, including solid organ transplant*

Neurological disease, including cerebrovascular disease*
Obesity*
Rheumatologic disease
Anemia or other blood disorder
Tuberculosis

**Risk Conditions outlined as priority groups for vaccination by WHO's European Technical Advisory Group of Experts on Immunization (ETAGE) at its online meeting on 11–12 November 2020* (8).

3.9.2 Identifying healthcare workers

Information on whether SARI patients are healthcare workers should be collected. Healthcare workers can be defined according to previously published WHO guidance documents *(9)* or according to national guidelines used to define the target population of healthcare workers eligible for COVID-19 vaccination.

Because essential or frontline workers have been shown to be at increased risk for severe COVID-19 infection, Information on essential worker status could be collected as well.

3.9.3 Residence (Long-Term Care Facility vs Private Residence)

Information on type of residence should be collected, as Long-Term Care Facility residents have been shown to be at higher risk for severe COVID-19 outcomes, and, as a result, these populations have been prioritized for vaccination.

3.9.4 Hospitalizations and ambulatory visits in the year prior to hospitalization

In order to document and control for healthcare-seeking behaviour and the severity of the underlying conditions, information on the number of hospital admissions in the previous 12 months prior to admission to the hospital should be collected.

3.9.5 Prior infections with SARS-CoV-2

Because prior infection with SARS-CoV-2 is associated with immunity to subsequent infections, previous infections with SARS-CoV-2 should be documented. Previous SARS-CoV-2 diagnoses related to the current illness episode and SARS-CoV-2 diagnoses prior to the current illness episode should be documented.

3.9.6 Previous influenza and pneumococcal vaccinations

Information on vaccinations received for influenza and pneumococcal disease [(pneumococcal polysaccharide vaccine (PPV) and pneumococcal conjugate vaccine (PCV)] including date of vaccination should be collected when available.

3.9.7 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 functional impairment can be captured. Where possible, the Barthel Index *(10)* should be used.

3.9.8 Ethnicity (optional)

Some studies have shown that certain ethnic groups may be at higher risk, either for becoming infected with, or for developing severe COVID-19. Information on ethnic group, relevant to the context of the surveillance system, could be also collected from SARI patients.

3.9.9 Medications for chronic condition(s) (optional)

The patient's medications for chronic diseases (medication name and amount and frequency of dose) in the previous 3 months should be recorded.

3.9.10 Behavior related to public health and social measures (optional)

Suboptimal adherence to public health and social measures (such as physical distancing, avoiding crowded spaces, mask-wearing, or hand hygiene), can lead to an increased risk of infection. Individuals who are not consistently observing public health and social measures may be less likely to get vaccinated against COVID-19. Compliance with physical distancing, mask-wearing and hand hygiene should be documented when possible.

3.9.11 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, have been shown to be at greater risk of infection and severe disease, and may also be less able to access vaccination services. As a result, SES status of SARI patients should be documented when possible.

3.10 Ethical considerations

Data collection for this vaccine effectiveness analysis will be conducted within the framework of ongoing SARI surveillance, which constitutes routine public health practice, However, approvals or ethical review committee waivers should be sought from the appropriate local ethical review committee(s).

3.11 Data

3.11.1 Sources of information

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

- hospital medical records
- interview with patient or his/her family
- interview with patient's GP
- interview with patent's pharmacist
- vaccination register
- laboratory

- *Local protocols should describe the sources of information used for each variable collected and the potential limitations of the data sources*

3.12 Information to be collected from SARI patients (See Annex 2)

The following information should be collected from SARI patients:

- Hospital information
  - country, hospital
  - vaccination target groups
  - first ward of referral
  - ICU/other ward admission
- patient characteristics *(ethnic group optional)*
- signs, symptoms
  - current
  - previous clinical symptoms (if no prior tests done)
- dates
  - vaccination (COVID-19, influenza, pneumococcal disease)
  - onset of symptoms
  - admission, discharge
  - swabbing
- laboratory
  - type of swab (nasopharyngeal, sputum, etc.) *(optional)*
  - type of test
  - results (including information on 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 visits in the previous 3 months *(optional)*
- presence of influenza and other respiratory viruses (see section 3.9.10)
- vaccination and antivirals (see sections 3.9.8–3.9.9)
  - pandemic vaccination including number of doses, date, product
  - current influenza vaccination
  - pneumococcal vaccination status, type of vaccine and either date or year of vaccination *(optional)*
  - antiviral administration *(optional)*
- functional status or proxy by residence type (see section 3.9.9)
- setting (e.g. LTCF)
- SES/deprivation *(optional)*

3.12.1 Information about the surveillance system and the patient

The following information should be recorded.

- Country, site, priority vaccination target group(s)
- Hospital (note: actual name of hospital will not be collected, but a unique number will be assigned by each participating country to each hospital, to be able to adjust 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 the pooled database)

3.12.2 Hospital/ward information

The following information should be recorded:

- 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 information should be recorded:

- Age
- Sex
- Smoking history (see section 3.9.5)
- Pregnancy
- Healthcare worker
- Occupation *(optional)*
- Clinical frailty score at admission (where possible; see section 3.9.10) *(optional)*
- Ethnic group *(optional)*
- Risk-of exposure
  - mask use, hand hygiene, social distancing
- SES/deprivation (optional)

3.12.4 Clinical characteristics (symptoms and markers of severity)

Clinical characteristics that comprise the SARI case definition *(4)* and the WHO suspected COVID-19 case definition *(5)* should be collected.

The following information related to the SARI hospitalization, which can be used to indicate severity, should be collected

- Length of hospital stay
- Oxygen use
- ICU admission
- Invasive ventilation
- Death

The following information related to SARI illness and hospitalization should be recorded:

- the **date of first key symptom onset**
- information on COVID-19 test(s) and laboratory results, including information on genetic analysis, when available.
- **date of vaccination**
- **date of swab** (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)
- **date of discharge/death** (to allow measurement of length of hospital stay).

3.12.5 Data entry validation

For hospitals using electronic medical records (EMR), if paper questionnaires are used, a sample of them should be validated using the EMR data and the surveillance database.

- *Local protocols should describe procedures used to validate data.*

3.12.6 Data management

Data entry and transfer

Web-based data collection methods or paper-based methods can be used. Data entry should include checks to minimise data entry errors. Double data entry is recommended if paper forms are used for data collection.

Data storage and data management for individual country analysis

Data validation, cleaning and verification should be carried out at routine intervals.

Data storage and data management for pooled analysis

Pooled analysis of data from SARI sentinel surveillance systems in the region will be performed for the WHO European Region Office. Like in routine SARI surveillance, additional data on SARI patients that is relevant for vaccine effectiveness estimates can be reported to WHO-Europe through the TESSy platform. SARI data relevant to vaccine effectiveness estimates that is submitted to WHO-Europe will be analyzed collectively, as is the case currently for pooled SARI data. WHO-Europe will provide a coordinating role in this work, as it has done with existing SARI surveillance systems in Member States.

More information on pooling is provided in Annex 3.

Irrespective of plans for pooled analyses, all data management procedures should comply with applicable national laws on data security.

3.13 Data analysis

3.13.1 SARI patients confirmed as COVID-19 (confirmed cases)

A confirmed COVID-19 case should be defined as a patient hospitalised with SARI symptoms, with a respiratory sample positive for SARS-CoV-2 by PCR only, either within 48 hours of hospital admission or documented within 14 days prior to hospital admission.

3.13.2 Primary control group: SARI patients who are negative for SARS-CoV-2

The primary control group will be defined as patients hospitalised with SARI symptoms, with a respiratory sample **negative** for SARS-CoV-2 by PCR on hospital admission, up to 48 hours after hospital admission. Controls should ideally also have a negative test within the period of 14 days prior to hospital admission.

SARI patients who test negative for SARS-CoV-2 at hospitalization (controls) who have reported prior SARS-CoV-2 infection in the previous 3 months prior to admission should be excluded as controls in sensitivity analyses. In addition, sensitivity analyses should be performed in which the control group only includes SARI patients who tested negative for SARS-CoV-2 on enrolment, and up to 7 days prior to enrolment.

3.13.3 Definition of vaccination status

COVID-19 vaccine

- An individual should be considered as vaccinated with one dose against COVID-19 with a product-specific vaccine (see section “Vaccination status ascertainment”) if s/he has received one dose of the named vaccine product more than 14 days before SARI symptom onset.
- “Fully vaccinated” should be defined according to vaccine product recommendations. Most likely patients will be considered fully vaccinated if they have received the second (final) dose at least 14 days before symptom onset
- “Partially vaccinated” should be defined according to vaccine product recommendations. Patient should likely be considered partially vaccinated if they have received only the first of two doses at least 14 days before SARI symptom onset
- A SARI patient should be considered as unvaccinated if s/he did not receive COVID-19 vaccine or if s/he was vaccinated after onset of symptoms.
- Sensitivity analyses evaluating CVE at different periods of time post-vaccination may be conducted.

Local protocols should describe the country policy regarding the timing of first and second doses of COVID-19 vaccine (when applicable).

3.13.4 VE analysis using the test-negative design

Briefly, cases and controls should first be described by baseline characteristics. Patients will be described according to:

- sex
- age group

- Occupation (healthcare worker, other essential worker)
- Date 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
- BMI
- influenza and pneumococcal vaccination status
- respiratory co-infections (where available)
- severity (use of oxygen, ICU admission, mechanical ventilation, death)
- CT values of positive specimens

An example layout of this descriptive analysis is provided in Table 3 below.

Table 3. Example of descriptive table for cases and controls

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		
0–14	x/x (x)	x/x (x)
15–44	x/x (x)	x/x (x)
45–64	x/x (x)	x/x (x)
≥ 65	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.		

VE then should be evaluated using the test-negative design approach. 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.

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

Prior to multivariable analysis, a model development strategy will be determined. 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.

WHO/EURO can provide example scripts to aid in data analysis.

Output tables presenting CVE estimates

In order 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 4 can be used (variables have been presented as an example of the output format).

Table 4. Example of table showing vaccine effectiveness against COVID-19 adjusted for various covariables by sex and age group

Type/subtype	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

In sensitivity analyses, CVE should be evaluated with different cut-offs of numbers of days between symptom onset and swabbing, between symptom onset and hospitalisation, and between vaccination and onset of symptoms. Other sensitivity analyses include exclusions of prior test positives and those positive to a seasonal coronavirus (e.g. HCoV-NL63, HCoV-229E, HCoV-OC43 and HCoV-HKU1).

3.13.5 Pooled analysis of surveillance data

SARI surveillance data, including relevant data related to vaccination status, can be submitted through the European Surveillance System platform (TESSy). WHO/EURO will be able to conduct an analysis of pooled surveillance data submitted through TESSy (see Annex 3 for details).

4. Limitations

It may be challenging to collect all of the required data on SARI patients. In addition, very severely ill patients (e.g. those who are extremely frail and/or in nursing homes) may not be admitted to hospital at all. Potential limitations to the VE estimates for COVID-19 are discussed below.

4.1 Potential biases

4.1.1 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 (e.g. presence of chronic diseases) will minimise negative confounding.

4.1.2 Positive confounding

Positive confounding refers to biases that reflect a 'healthy vaccine effect'. People with a healthy lifestyle will be more likely to get vaccinated, leading to an increase of estimated CVE. Conversely, people in a state of extreme frailty may not be offered vaccine, and, because they are frail, may be more likely to have severe disease. Persons who have a higher risk of infection (through increased exposure – e.g. not routinely wearing facemasks where recommended) may also be less likely to be vaccinated, which may also increase their exposure to disease. If positive confounding is present, CVE will be overestimated.

4.1.3 Unmeasured confounding

Positive and negative confounding will be minimised through stratification and multivariable analysis. However, not all characteristics that could lead to positive or negative confounding are captured in surveillance data . Therefore, some residual unmeasured confounding may remain.

4.1.4 Previous infection in cases or control; inclusion of asymptomatic controls

Individuals who have been previously infected may have a greater response to the vaccine or be less likely to be reinfected even if unvaccinated. Some controls (those testing negative for SARS-CoV-2) may have been previously infected symptomatically or asymptomatically with SARS-CoV-2 without having been tested. Additionally, having been previously infected could impact one's decision to be vaccinated. Previously infected individuals may be less likely to get vaccinated and less likely to get infected. This could lead to a lower percentage of vaccinated controls and increase CVE.

Sensitivity analyses should be conducted excluding any SARI patient with previous SARS-CoV-2 infection confirmed either by PCR or by antibody test.

4.1.5 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 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 there to be a relationship between being positive for influenza and receiving COVID-19 vaccination, which introduces bias. Sensitivity analyses should be conducted excluding controls who are positive for influenza.

4.1.6 Validation of exposure

The vaccination status is the exposure of interest and therefore reliable vaccination data is critical. If the vaccination status is only self-reported, without written or electronic documentation, information bias may occur. Vaccination status of cases and controls should ideally be validated using an independent source (i.e. vaccination register, GPs).

4.1.7 Bias from pooled estimates

With data from a number of different hospitals from different countries being pooled, bias in the individual studies can influence the pooled estimate. If few sites are included in the pooled analysis, the power of the test for the presence of heterogeneity between individual studies will be low and may not be able to detect heterogeneity. 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. Many aspects of surveillance systems may differ. Different tests may be used, and there may be differences in the number of tests used to declare an individual negative. Different criteria for admission may exist, and these criteria may change over time depending on the intensity of the pandemic and hospital capacity. Some hospitals may only admit

laboratory-confirmed COVID-19 patients, or, conversely, only patients who have tested negative for SARS-CioV-2 prior to admission.

To allow for complete assessment of heterogeneity, sites should document all changes in their COVID-19 surveillance system, including testing and admission policies, during the surveillance period.

4.1.8 Other potential biases

Cases and controls could come from different geographic and demographic backgrounds that are associated with different risk for SARS-CoV-2 infection, and different likelihoods of being vaccinated. Time (onset date) should be used to adjust for seasonal differences. Analyses will also be stratified by time.

4.2 Representativeness of subjects

SARI surveillance includes only cases that are hospitalised. Health-seeking behaviour and approaches to patient management may differ by country or within countries. These differences could affect the time between onset of symptoms and hospitals, and may affect positivity rates. To the extent possible, health-seeking behaviour and case-management strategies should be described for each SARI surveillance system and it should be noted how these may affect the CVE estimates.

4.3 Community Engagement

This study will engage with the community in so far as it will be conducted in hospitals, which serve their local communities. Efforts should be made to be transparent about the objectives of the study with interested community members.

5. Annexes

Annex 1. Questionnaire for SARI Patients

Section A. Administrative Information	
1. Form completion date (dd/mm/yyyy)	___/___/___
2. Unique patient ID	
3. Hospital code	

Section B. Personal Information	
1. Date of birth (dd/mm/yyyy) or age	___/___/___ or age: ____
2. Sex	☐ Female ☐ Male ☐ Do Not Know
3. What is the patient's ethnic group? (optional)	
4. Patient's primary residence prior to admission	☐ Private residence/home ☐ Long-term care facility ☐ Other_____ ☐ Do not know
4a. If private residence/home	☐ at home, not dependent on support/care ☐ at home, dependent on home support/care ☐ at home, unknown support/care
6. Is the patient a healthcare worker*?	☐ Yes ☐ No ☐ Do not know
7. What is the patient's occupation?	☐ _____ ☐ Do not know
8. What is the patient's height (in m)?	☐ _____ ☐ Do not know
9. What is the patient's weight (in kg)?	☐ _____ ☐ Do not know

*Healthcare worker can be defined according to the country definition or according to the WHO definitions proposed in the text of this guidance document

Section C. Medical History	
1. Which of the following Chronic Medical Conditions does the patient have?	
a. Cancer	☐ Yes ☐ No ☐ Do not know
b. Chronic cardiac disease, except hypertension	☐ Yes ☐ No ☐ Do not know
c. Hypertension	☐ Yes ☐ No ☐ Do not know
d. Chronic kidney disease	☐ Yes ☐ No ☐ Do not know
e. Chronic liver disease	☐ Yes ☐ No ☐ Do not know
f. Chronic respiratory disease	☐ Yes ☐ No ☐ Do not know
g. Asthma	☐ Yes ☐ No ☐ Do not know
h. Diabetes	☐ Yes ☐ No ☐ Do not know
i. Immunocompromised, including solid organ transplant and HIV	☐ Yes ☐ No ☐ Do not know
j. Neurological disease, including cerebrovascular disease	☐ Yes ☐ No ☐ Do not know
k. Rheumatologic disease	☐ Yes ☐ No ☐ Do not know
l. Anemia or other blood disorder	☐ Yes ☐ No ☐ Do not know
m. Tuberculosis	☐ Yes ☐ No ☐ Do not know
n. Obesity	☐ Yes ☐ No ☐ Do not know
o. Other	
2. Smoking status	☐ Never smoked ☐ Former smoker (stopped> 1 year ago) ☐ Current smoker, or stopped in the past year ☐ Do not know
3. Is the patient pregnant? (for female patients aged 15-55)	☐ Yes ☐ No ☐ Do not know If yes, trimester ☐ 1 ☐ 2 ☐ 3 ☐ Do not know
4. How many times was the patient hospitalized in the previous 12 months (not including the current hospitalization)?	☐ Number of hospitalizations: ___ ☐ Do not know

Section D. Vaccination	
1. Has the patient received a **COVID-19** vaccine?	
a. Dose 1	☐ Yes ☐ No ☐ Do not know If yes, date: ___/___/___
i. Name/brand: ☐ Pfizer/BioNTech ☐ AstraZeneca ☐ Moderna ☐ Sputnik V ☐ Janssen ☐ EpiVacCorona ☐ Covaxin (Bharat) ☐ Covishield ☐ Novavax ☐ CoronaVac (Sinovac) ☐ Sinopharm COVID-19 vaccine ☐ Do not know ☐ Other, specify _____	
i. Batch number (if known):	
b. Dose 2	☐ Yes ☐ No ☐ Do not know ☐ Second dose not applicable If yes, date: ___/___/___
i. Name/brand: ☐ Pfizer/BioNTech ☐ AstraZeneca ☐ Moderna ☐ Sputnik V ☐ Janssen ☐ EpiVacCorona ☐ Covaxin (Bharat) ☐ Covishield ☐ Novavax ☐ CoronaVac (Sinovac) ☐ Sinopharm COVID-19 vaccine ☐ Do not know ☐ Other, specify _____	
i. Batch number (if known):	
c. Vaccination status ascertainment: ☐ Vaccination registry ☐ Vaccination card ☐ Patient's hospital record ☐ Patient's GP ☐ Pharmacist ☐ Patient's insurance record ☐ Patient interview ☐ Relative interview ☐ Patient self-report (mobile phone) ☐ Other, specify _____	
2. Other vaccines	
a. Latest seasonal influenza *(vaccinated any time from 01 September 2020)*	☐ Yes ☐ No ☐ Do not know If yes, date: ___/___/___
If exact date unknown: was patient vaccinated **at least 14 days before onset** of current illness?	☐ Yes ☐ No ☐ Do not know
Influenza vaccination product name/brand	_____
Seasonal influenza vaccine in previous season (n-1)	☐ Yes ☐ No ☐ Do not know
Seasonal influenza vaccine in season 2 years prior to current season (n-2)	☐ Yes ☐ No ☐ Do not know
b. PPV23	☐ Yes ☐ No ☐ Do not know If yes, date: ___/___/___
c. Pneumococcal Conjugate Vaccine (PCV-10 or PCV-13)	☐ Yes ☐ No ☐ Do not know If yes, date: ___/___/___

Section E. SARS-CoV-2 – before hospitalization	
1. In the two weeks before admission to the hospital did the patient get tested for SARS-CoV-2?	☐ Yes ☐ No ☐ Do not know
a. If yes, which test was used, when was it performed and what were the results?	☐ Rapid test date _/_/_☐ pos ☐ Neg ☐ Undetermined ☐ Do not know ☐ PCR date _/_/_☐ pos ☐ Neg ☐ Undetermined ☐ Do not know ☐ Serology date _/_/_☐ pos ☐ Neg ☐ Undetermined ☐ Do not know ☐ Other kind of test _____ ☐ Do not remember what kind of test was done
b. Has the patient ever tested positive for SARS-CoV-2 prior to the acute illness associated with this hospital admission?*	☐ yes ☐ No ☐ Do not know (if source of information is hospital records and no data available)
If yes, which test was used?	☐ Rapid test date _/_/_ ☐ PCR date _/_/_ ☐ Serology date _/_/_ ☐ Do not remember what kind of test was done

*If the patient reported having a positive test in question 1, then question 2 can be skipped.

Section F. Hospital/Ward information	
1. Date of admission to hospital (dd/mm/yyyy)	___/___/___
2. Was patient admitted to ICU/HDU?	☐ Yes ☐ No ☐ Do not know
3. Date of admission to ICU/HDU (dd/mm/yyyy)	___/___/___ ☐ Do not know
4. Date of discharge from ICU/HDU (dd/mm/yyyy)	___/___/___ ☐ Do not know

Section G. Symptoms or signs, either at or prior to admission	
1. Fever	☐ Yes, measured >=38C ☐ Yes, but not measured ☐ No ☐ Do not know
2. Headache	☐ Yes ☐ No ☐ Do not know
3. Sore throat	☐ Yes ☐ No ☐ Do not know
4. Cough	☐ Yes ☐ No ☐ Do not know
5. Runny nose (coryza)	☐ Yes ☐ No ☐ Do not know
6. Shortness of breath	☐ Yes ☐ No ☐ Do not know
7. General weakness and/or fatigue	☐ Yes ☐ No ☐ Do not know
8. Muscle pains/myalgia	☐ Yes ☐ No ☐ Do not know
9. Loss of smell (anosmia)	☐ Yes ☐ No ☐ Do not know
10. Loss of taste (ageusia)	☐ Yes ☐ No ☐ Do not know
11.Vomiting or nausea or loss of appetite (anorexia)	☐ Yes ☐ No ☐ Do not know
12. Abdominal pain	☐ Yes ☐ No ☐ Do not know
13. Diarrhoea	☐ Yes ☐ No ☐ Do not know
14. Heart palpitations	☐ Yes ☐ No ☐ Do not know
15. Chest pain	☐ Yes ☐ No ☐ Do not know
16. Dizziness	☐ Yes ☐ No ☐ Do not know
17. Rash	☐ Yes ☐ No ☐ Do not know
18. Altered mental status	☐ Yes ☐ No ☐ Do not know
19. Onset date of first symptoms related to current illness	___/___/___

Section H. Clinical course during hospitalization	
1. Did the patient receive oxygen?	☐ Yes ☐ No ☐ Do not know
2. Please indicate whether the patient received any respiratory support while in hospital. *(Note: enter most invasive support that applies.)*	☐ None ☐ High-flow oxygen ☐ Mechanical ventilation ☐ ECMO ☐ Other respiratory support (*specify*): _____ ☐ Unknown
3. Patient outcome at time of data collection *(Note: this information should be updated at conclusion of hospitalization)*	☐ Died in hospital (date _/__/_) ☐ Discharged* (date __/__/__) ☐ Still on treatment (in hospital) ☐ Unknown outcome

*Discharged alive

Section I. Laboratory testing and results (SARS-CoV-2) – during hospitalization	
1. Was the patient tested for SARS-CoV-2 by PCR during his or her hospitalization?	☐ Yes ☐ No ☐ Do not know
2. If the patient was tested by PCR:	
a. Respiratory specimen collection date	___/___/___
b. Laboratory result: virus type SARS-CoV-2	☐ Negative ☐ Positive ☐ Inconclusive or undetermined ☐ Do not know
c. Ct value	
3. If first PCR was negative, was a second PCR test performed?	☐ Yes ☐ No ☐ Do not know
a. Second respiratory specimen collection date	___/___/___
b. Second laboratory result: virus type SARS-CoV-2	☐ Negative ☐ Positive ☐ Inconclusive or undetermined ☐ Do not know
c. Ct value	
4. For positive SARS-CoV-2 samples, was genetic sequencing performed?	☐ Yes ☐ No ☐ Do not know
a. If yes, genetic sequencing results	

Section J. Laboratory testing for other respiratory viruses	
1. Laboratory result: any influenza virus type	☐ Not done ☐ Negative ☐ Positive ☐ Do not know if test was done If positive, indicate influenza type, subtype: ☐ A(H1N1) ☐ A(H3N2) ☐ A (untyped) ☐ B/Yamagata ☐ B/Victoria ☐ B/untyped ☐ unknown type ☐ Inconclusive or undetermined
2. Other respiratory viruses tested – positive results	☐ Not done ☐ No other positive results ☐ RSV ☐ Metapneumovirus ☐ Adenovirus ☐ Other respiratory infection, specify _____ ☐ Do not know if additional testing was done

Section K. Frailty Assessment (Optional)	
If the Barthel score was documented at hospital admission, what was the score?	____/100 ☐ not documented
If the Clinical Frailty Score was documented at hospital admission what was the score?	____/9 ☐ not documented
If another score was used please document the name of the scale and the score:	Name_____ Score_____
If not, please ask the patient the following questions or identify the answers from medical records. The questions refer to the period prior to the patient's acute illness that led him/her to be hospitalized.	
1. Before the patient's acute illness, was the patient able to walk across the room independently?	☐ Yes ☐ No ☐ Do not know
2. Before the patient's acute illness, was the patient able to use the toilet independently?	☐ Yes ☐ No ☐ Do not know
3. Before the patient's acute illness, was the patient able to bathe independently?	☐ Yes ☐ No ☐ Do not know

Section L. Medications for chronic medical conditions in the past 3 months (optional)	
1. Statins	☐ Yes ☐ No ☐ Do not know
2. Metformin	☐ Yes ☐ No ☐ Do not know
3. Steroids	☐ Yes ☐ No ☐ Do not know
4. Corticosteroids	☐ Yes ☐ No ☐ Do not know
5. NSAIDs *(non-steroidal anti-inflammatory drugs)*	☐ Yes ☐ No ☐ Do not know
6. ACE *(angiotensin-converting enzyme)* inhibitors	☐ Yes ☐ No ☐ Do not know
7. ARBs *(angiotensin II receptor blockers)*	☐ Yes ☐ No ☐ Do not know
8. DMARDs *(Biological disease-modifying anti-rheumatic drugs, e.g. rituximab, tocilizumab, etc.)*	☐ Yes ☐ No ☐ Do not know
9. Cancer chemotherapy *(within 6 months or currently)*	☐ Yes ☐ No ☐ Do not know
10. Gliclazides *(for diabetes or heart failure)*	☐ Yes ☐ No ☐ Do not know
11. Psychotropic drugs *(including benzodiazepine, etc.)*	☐ Yes ☐ No ☐ Do not know
12. Chloroquine	☐ Yes ☐ No ☐ Do not know
13. Hydroxychloroquine	☐ Yes ☐ No ☐ Do not know
14. Other 1 (*key chronic medications)*	☐ Yes ☐ No ☐ Do not know
15. Other 2 (*key chronic medications)*	☐ Yes ☐ No ☐ Do not know
16. Other 3 (*key chronic medications)*	☐ Yes ☐ No ☐ Do not know

Section M. Public health and social measures (optional)	
The patient should be asked which of the following measures* he or she practiced during the 7 days prior to hospital admission:	
1. Wore a mask in public (e.g. in shops, on public transport, in the company of those not living in their home)	☐ Always ☐ Usually ☐ Sometimes ☐ Never ☐ Not applicable
2. Frequently washed hands with soap and water for at least 20 seconds	☐ Always ☐ Usually ☐ Sometimes ☐ Never ☐ Not applicable
3. Used sanitiser when soap and water were unavailable	☐ Always ☐ Usually ☐ Sometimes ☐ Never ☐ Not applicable
4. Ensured physical distancing in public (e.g. remaining at least 1 m* away from others)	☐ Always ☐ Usually ☐ Sometimes ☐ Never ☐ Not applicable

*Questions can be adjusted to match national recommendations.

Annex 2. 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 necrotizing 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,	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

	M34.81, M05.10	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 3 months.

Annex 3. List of variables, definitions and coding; hospital-based COVID-19 vaccine effectiveness minimum dataset

Individual data

- *Countries should list all the variables being collected and how they are coded*
- *Countries should indicate which variables collected within the system differ from the variables listed below*
- *Optional variables in the table below are shaded in grey*

	Variable	Type	Values and coding	Definition
Section A Administrative information	idcountry	Numeric (categorical)	Coded according to international country codes	Identifier uniquely identifying the country (for pooled datasets only)
	form_date	Date		Form completion date
	id	Numeric	Unique integer	Unique number for each patient
	hospitalcode	Numeric	Unique integer	Unique number for each hospital
Section B Personal information	dob	Date	dd/mm/yyyy	Date of birth (only if no age; once age calculated from dob this will be dropped) *(optional)*
	age	Numeric (integer)		Age of patient
	sex	Numeric	0 = female	Sex of patient
			1 = male	
			3 = other	
			8 = do not know	
	ethnic	Numeric (categorical)		Patient's ethic group (note: codes should be country-specific) *(optional)*
	ethnic_sp	Text		Other ethnic group not specified in coding above *(optional)*
	residence		1 = private residence/home	Patient's primary residence at time of SARI onset. Whether patient was living at home or was institutionalised
			2 = institutionalised (LTCF)	
			3 = other	
			8 = Do not know	
	residence_home	Numeric (categorical)	0 = at home, not dependent on home support/care	Whether patient residing at home had pre-hospital dependence on home support/care
			1 = at home, but dependent on home support/care	
			2 = at home, unknown support/care	
	residence_sp	Text		Specify other residence (e.g. prison)

	Variable	Type	Values and coding	Definition
Section B continued	hcw	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Whether the patient is a healthcare worker
	occupation	Text		Patient's occupation
	occupation_nk	Numeric (categorical)	8 = do not know	Patient's occupation unknown
	height	Numeric (integer)		Height of patient in metres
	weight	Numeric (integer)		Weight of patient in kg
	height_nk	Numeric (categorical)	8 = do not know	Height unknown
	weight_nk	Numeric (categorical)	8 = do not know	Weight unknown
Section C Medical history	cancer	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Cancer (any)
	heartdis	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Chronic heart / cardiac disease (excluding hypertension)
	hypert	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Hypertension
	rendis	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Chronic kidney disease (excluding cancer and acute renal failure)
	liverdis	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Chronic liver disease (excluding cancer)
	lungdis	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Chronic respiratory disease (excluding asthma)
	asthma	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Asthma
	diabetes	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Diabetes

	Variable	Type	Values and coding	Definition
	immuno	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Immunocompromised, including solid organ transplant and HIV
	neurol	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Neurological disease, including cerebrovascular disease
Section C continued	rheumat	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Rheumatologic disease
	anaemia	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Anaemia or other blood disorder
	tuberc	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Tuberculosis *(optional)*
	obese	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Obesity (only if height, weight and BMI not collected; can be calculated)
	chronic_other	Text		Other chronic condition
	smoking	Numeric (categorical)	0 = Never 1 = Former 2 = Current 8 = Do not know	Smoking status Never, former (stopped smoking at least 1 year before hospital admission), current smoker (or stopped within past year)
	pregnant	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Whether patient is pregnant (for female patients aged 15–55 years)
	trimester	Numeric (categorical)	1 = first trimester 2 = second trimester 3 = third trimester 8 = Do not know	If patient is pregnant, which trimester of pregnancy
	hosp_visit	Numeric (integer)		Number of times patient was admitted to hospital in the 12 months prior to COVID diagnosis for an underlying chronic condition (not including current hospitalisation)
	hosp_visit_nk	Numeric (categorical)	8 = Do not know	Number of hospital visits unknown

	Variable	Type	Values and coding	Definition
Section D Vaccination	panvacc_dose1	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Received pandemic COVID-19 vaccination, first dose
	panvacc_date1	Date	dd/mm/yyyy	Vaccination date, first dose
Section D Vaccination continued (pandemic vaccine)	panvacc_name1	Numeric (categorical)	1 = AstraZeneca 2 = Beijing CNBG (inactivated) 3 = Covaxin (Bharat) 4 = Pfizer/BioNTech (Comirnaty) 5 = Janssen 6 = Moderna 7 = Covishield – SII 10 = CoronaVac (Sinovac) 11 = Sputnik V 12 = EpiVacCorona – SRCVB 13 = Wuhan CNBG (inactivated) 8 = Do not know 9 = Other	Vaccine product name, first dose
	panvacc_name1_oth	Text		Specify other vaccine
	panvacc_batch1	Text		Batch number for first dose
	panvacc_dose2	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know 9 = Second dose not applicable	Received pandemic COVID-19 vaccination, second dose
	panvacc_date2	Date	dd/mm/yyyy	Vaccination date, second dose
	panvacc_name2	Numeric (categorical)	1 = AstraZeneca 2 = Beijing CNBG (inactivated) 3 = Covaxin (Bharat) 4 = Pfizer/BioNTech (Comirnaty) 5 = Janssen 6 = Moderna 7 = Covishield – SII 10 = CoronaVac (Sinovac) 11 = Sputnik V 12 = EpiVacCorona – SRCVB 13 = Wuhan CNBG (inactivated) 8 = Do not know 9 = Other	Vaccine product name, second dose
	panvacc_name2_oth	Text		Specify other vaccine
	panvacc_batch2	Text		Batch number for second dose

	Variable	Type	Values and coding	Definition
	panvacc_ascertain	Numeric (categorical)	1 = vaccination registry	Vaccination status ascertainment
			2 = vaccination card	
			3 = hospital record	
			4 = GP record	
			5 = pharmacist record	
			6 = insurance record	
			7 = patient interview	
			8 = family interview	
			10 = patient self-report (mobile)	
			9 = other	
	panvacc_ascertain_sp	Text		Specify other ascertainment method
Section D Vaccination continued (other vaccines)	flu_vacc	Numeric (categorical)	0 = No	Received latest seasonal influenza vaccination (any time from 01 Sep 2020)
			1 = Yes	
			8 = Do not know	
	flu_vaccdate	Date	dd/mm/yyyy	Date of last influenza vaccination
	flu_vacc_14d	Numeric (categorical)	0 = No	If exact date unknown, whether patient vaccinated at least 14 days before illness onset
			1 = Yes	
			8 = Do not know	
	flu_vacc_brand	Text		Influenza vaccine brand/product name
	prev_fluvacc_n1	Numeric (categorical)	0 = No	Whether patient had seasonal flu vaccine in the previous season (n-1)
			1 = Yes	
			8 = Do not know	
	prev_fluvacc_n2	Numeric (categorical)	0 = No	Whether patient had seasonal flu vaccine in the season before the last (n-2)
			1 = Yes	
			8 = Do not know	
	ppv_vacc	Numeric (categorical)	0 = No	Received PPV23 vaccination
			1 = Yes	
			8 = Do not know	
	ppv_vaccdate	Date	dd/mm/yyyy	Date of last PPV23 vaccination
	pcv_vacc	Numeric (categorical)	0 = No	Received PCV7/10 or 13 vaccination
			1 = Yes	
			8 = Do not know	
	pcv_vaccdate	Date	dd/mm/yyyy	Date of last PCV7/10 or 13 vaccination

	Variable	Type	Values and coding	Definition
Section E SARS-CoV-2 tests before admission to hospital	lab_covtest_pre	Numeric (categorical)	0 = No	Whether patient was tested for SARS-CoV-2 during the 2 weeks prior to hospital admission
			1 = Yes	
			8 = Do not know	
	lab_covtesttype_pre	Numeric (categorical)	1 = RT-PCR	Type of lab test used
			2 = Serology	
			3 = Rapid test	
			4 = Other	
			8 = Do not know	
	lab_covtesttype_sp	Text		Specify other type of lab test
	testdate_pcr_pre	Date	dd/mm/yyyy	PCR test date (in 2 weeks pre admission)
	testdate_serol_pre	Date	dd/mm/yyyy	Serology test date (in 2 weeks pre admission)
	testdate_rapid_pre	Date	dd/mm/yyyy	Rapid test date (in 2 weeks pre admission)
	testdate_oth_pre	Date	dd/mm/yyyy	Other test date (in 2 weeks pre admission)
	lab_covid_pcr_pre	Numeric (categorical)	0 = Negative	Laboratory result for PCR performed within 2 weeks of hospital admission: virus type SARS-CoV-2
			1 = Positive	
			2 = Inconclusive/undetermined	
			8 = Do not know	
	lab_covid_rapid_pre	Numeric (categorical)	0 = Negative	Laboratory result for rapid test performed within 2 weeks of hospital admission: virus type SARS-CoV-2
			1 = Positive	
			2 = Inconclusive/undetermined	
			8 = Do not know	
	lab_covid_serol_pre	Numeric (categorical)	0 = Negative	Laboratory result for serology performed within 2 weeks of hospital admission: virus type SARS-CoV-2
			1 = Positive	
			2 = Inconclusive/undetermined	
			8 = Do not know	
	lab_covid_other_pre	Numeric (categorical)	0 = Negative	Laboratory result for other test performed within 2 weeks of hospital admission: virus type SARS-CoV-2
			1 = Positive	
			2 = Inconclusive/ undetermined	
			8 = Do not know	
	lab_covtest_ever	Numeric (categorical)	0 = No	Whether patient ever tested positive for SARS-CoV-2 prior to admission (only answer if **no positive test** reported for previous 2 weeks, above)
			1 = Yes	
			8 = Do not know	

	Variable	Type	Values and coding	Definition
Section E SARS-CoV-2 tests before hospital admission, continued	lab_covtesttype_ever	Numeric (categorical) Numeric (categorical)	1 = RT-PCR 2 = Serology 3 = Rapid test 4 = Other 8 = Do not know	Type of lab test used
	testdate_pcr_ever	Date	dd/mm/yyyy	PCR test date (test ever done)
	testdate_rapid_ever	Date	dd/mm/yyyy	Rapid test date (test ever done)
	testdate_serol_ever	Date	dd/mm/yyyy	Serology test date (test ever done)
	testdate_other_ever	Date	dd/mm/yyyy	Other test date (test ever done)
Section F Hospital/ward information	admitdate	Date	dd/mm/yyyy	Date of hospital admission
	icu	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Admission to intensive care unit (ICU) or high-dependency unit (HDU)
	icuadmitdate	Date	dd/mm/yyyy	Date of admission to ICU/HDU
	icudisdate	Date	dd/mm/yyyy	Date of discharge from ICU/HDU
Section G Symptoms or signs, at or prior to admission	fever	Numeric (categorical)	0 = No 1 = Yes, measured ≥ 38°C 2 = Yes but not measured 8 = Do not know	History of fever
	headache	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Headache
	sorethroat	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Sore throat
	cough	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Cough
	coryza	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Coryza (runny nose)
	sob	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Shortness of breath

	Variable	Type	Values and coding	Definition
Section G continued	general_weak	Numeric (categorical)	0 = No	General weakness and/or fatigue
			1 = Yes	
			8 = Do not know	
	myalgia	Numeric (categorical)	0 = No	Myalgia/muscle pains
			1 = Yes	
			8 = Do not know	
	anosmia	Numeric (categorical)	0 = No	Loss of sense of smell (anosmia)
			1 = Yes	
			8 = Do not know	
	ageusia	Numeric (categorical)	0 = No	Loss of sense of taste (ageusia)
			1 = Yes	
			8 = Do not know	
	nausea_vom	Numeric (categorical)	0 = No	Nausea or vomiting or loss of appetite (anorexia)
			1 = Yes	
			8 = Do not know	
	abdopain	Numeric (categorical)	0 = No	Abdominal pain
			1 = Yes	
			8 = Do not know	
	diarr	Numeric (categorical)	0 = No	Diarrhoea
			1 = Yes	
			8 = Do not know	
	palp	Numeric (categorical)	0 = No	Heart palpitations
			1 = Yes	
			8 = Do not know	
	chest	Numeric (categorical)	0 = No	Chest pain
			1 = Yes	
			8 = Do not know	
	dizzy	Numeric (categorical)	0 = No	Dizziness
			1 = Yes	
			8 = Do not know	
	dermato	Numeric (categorical)	0 = No	Rash
			1 = Yes	
			8 = Do not know	
	confusion	Numeric (categorical)	0 = No	Altered mental status
			1 = Yes	
			8 = Do not know	

	Variable	Type	Values and coding	Definition
Section G continued	onsetdate	Date	dd/mm/yyyy	Onset date of first symptoms related to current illness
Section H Clinical course (severity and outcome)	resp_support	Numeric (categorical)	0 = None	Whether patient had any respiratory support while in hospital (most invasive level)
			1 = High-flow oxygen	
			2 = Mechanical ventilation	
			3 = ECMO	
			8 = Do not know	
			9 = Other respiratory support	
	resp_support_oth	Text		Specify other respiratory support
	Outcome	Numeric (categorical)	1 = died	Indicate the outcome of the patient known at the time of data collection *(Note: this information should be updated at conclusion of hospitalisation; if patient transferred select "still on treatment")*
			2 = discharged from hospital	
			4 = still on treatment (in hospital)	
			8 = unknown outcome	
	dischargedate	Date	dd/mm/yyyy	Date of hospital discharge
	deathdate	Date	dd/mm/yyyy	Date of in-hospital death
Section I Laboratory testing and results during hospital stay	lab_covtest_pcr	Numeric (categorical)	0 = No	Whether patient had a PCR test for SARS-CoV-2 during hospitalisation
			1 = Yes	
			8 = Do not know	
	swabdate	Date	dd/mm/yyyy	Date of swab for PCR in hospital
	covpcr_res	Numeric (categorical)	0 = Negative	PCR test result: virus type SARS-CoV-2
			1 = Positive	
			2 = Inconclusive/undetermined	
			8 = Do not know	
	ct_value	Numeric		Cycle threshold (Ct) value
	pcr2	Numeric (categorical)	0 = No	If first PCR was negative, was a second PCR performed?
			1 = Yes	
			8 = Do not know	
	swabdate2	Date	dd/mm/yyyy	Second PCR swab date
	covpcr2_res		0 = Negative	Second PCR result
			1 = Positive	
			2 = Inconclusive/undetermined	
			8 = Do not know	
	ct2_value	Numeric		Second cycle threshold (Ct) value

	Variable	Type	Values and coding	Definition
Section I continued	seq	Numeric (categorical)	0 = No	Whether patient sample was sequenced/sent for sequencing
			1 = Yes	
			8 = Do not know	
	genetic_group	Text		Laboratory result: genetic group
Section J Laboratory testing (other respiratory viruses)	lab_fluany	Numeric (categorical)	0 = Negative	Whether sample tested positive for any influenza virus type
			1 = Positive	
			2 = Not done	
			8 = Do not know	
	lab_flu_pos	Numeric (categorical)	1 = A(H1N1)	Indicate influenza virus type
			2 = A(H3N2)	
			3 = A (untyped)	
			4 = B/Yamagata	
			5 = B/Victoria	
			6 = B (untyped)	
			8 = Unknown type	
			9 = Inconclusive or undetermined	
	lab_resp_virus	Numeric (categorical)	0 = Not done	Whether positive for any other respiratory virus
			1 = No other positive results	
			2 = RSV	
			3 = Metapneumovirus	
			4 = Adenovirus	
			5 = Other respiratory infection	
			8 = Do not know	
	lab_resp_virus_sp	Text		Specify other respiratory infection
Section K Frailty assessments	frailty_barthel	Numeric (integer)		Total Barthel score at admission (if used) (max score=100)
	frailty_barthel_nd	Numeric (categorical)	0 = Not done	Select if Barthel score at admission not documented
	frailty_cfs	Numeric (integer)		Clinical Frailty Score (CFS) at admission (if used) (max score=9)
	frailty_cfs_nd	Numeric (categorical)	0 = Not done	Select if CFS at admission not documented
	frailty_othername	Text		Indicate name of other frailty score used at admission
	frailty_other_sp	Numeric (integer)		Specify total for other frailty score

	Variable	Type	Values and coding	Definition
	frailty_indep_walk	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Whether the patient was able to walk across the room independently before admission
	frailty_indep_toilet	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Whether the patient was able to use the toilet independently before admission
	frailty_indep_bathe	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Whether the patient was able to bathe independently before admission
Section L Medications for chronic conditions *(optional)*	statin_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Patient was on statins in the previous 3 months *(optional)*
	metform_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Metformin *(optional)*
	steroids_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Steroids *(optional)*
	corticost_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Corticosteroids *(optional)*
	nsaid_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	NSAID (non-steroidal anti-inflammatory drugs) *(optional)*
	ace_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	ACE inhibitor (angiotensin converting enzyme inhibitors) *(optional)*
	arb_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	ARB (angiotensin II receptor blockers) *(optional)*
	dmards_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Biological disease-modifying anti-rheumatic drugs (DMARDs) e.g. rituximab, tocilizumab, etc. *(optional)*

	Variable	Type	Values and coding	Definition
Section L Medications for chronic conditions, continued *(optional)*	chemo_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Chemotherapy (within 6 months or currently) for cancer *(optional)*
	gliclaz_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Gliclazides (for diabetes or heart failure) *(optional)*
	psychotrop_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Psychotropic drugs (including benzodiazepine, etc.) *(optional)*
	chloroq_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Chloroquine *(optional)*
	hydroxychloroq_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Hydroxychloroquine *(optional)*
	other1_pre_sp	Text		Other chronic medication #1 *(optional)*
	other2_pre_sp	Text		Other chronic medication #2 *(optional)*
	other3_pre_sp	Text		Other chronic medication #3 *(optional)*
Section M Practise of non-pharmaceutical interventions	risk1_mask	Numeric (categorical)	0 = Never 1 = Sometimes 2 = Usually 3 = Always 8 = Not applicable	Wore mask in public (e.g. in shops, on public transport, in the company of those not living in their home)
	risk2_handwash	Numeric (categorical)	0 = Never 1 = Sometimes 2 = Usually 3 = Always 8 = Not applicable	Frequently washed hands with soap and water for at least 20 seconds
	risk3_sanitiser	Numeric (categorical)	0 = Never 1 = Sometimes 2 = Usually 3 = Always 8 = Not applicable	Used hand sanitiser when soap and water not available
	risk4_socialdist	Numeric (categorical)	0 = Never 1 = Sometimes 2 = Usually 3 = Always 8 = Not applicable	Ensured physical distancing in public (e.g. remaining at least 1 m* away from others) *Countries to adjust these to match national recommendations.

Annex 4. Pooling Data

1. Data storage and management

Surveillance data should be submitted through TESSy to be included in a regional pooled analysis. Data sharing will follow the same procedures as for SARI surveillance data. A country (identifier will be included in each record (e.g. ES for Spain, UK for the United Kingdom), and a hospital code will be included (e.g. a unique number). Data management will follow the basic principles outlined below and in section 3.12.7 (Data management). A country-specific flowchart of exclusions and restrictions will be shared with each of the participating countries. Variables will be recoded and new variables generated. The recoded data will be stored separately from the crude data and recoding will be documented.

Summary and frequency tables and graphic displays of appropriate variables will be used to find illegal, implausible or missing values within the dataset. Checks for inconsistencies will be carried out (e.g. date of respiratory specimen collection before date of onset of symptoms). Any improbable, illegal or missing values should be investigated.

Any subsequent changes to the data will be fully documented and stored separately from the crude database, to ensure reproducibility and transparency of data management.

1.1 Missing data

Any missing data will be described. If there is much missing data with no evidence of bias in the missing data, and variables that are considered good predictors of the missing data are available, multiple imputation methods at study level will be used to replace missing values. A sensitivity analysis will be carried out comparing results from the complete case analysis (where records with missing data will be dropped) and the full set analysis (with imputed data).

1.2 Data cleaning

Summary and frequency tables as well as visual representations of appropriate variables will be used to find illegal, implausible or missing values within the dataset. Checks for inconsistencies should be carried out (e.g. date of discharge from hospital before date of onset of symptoms). Ideally, these checks could be included as warnings if using an electronic questionnaire, in order to avoid inconsistencies in the data entry. These values will be checked against the questionnaires or queried with the hospitals. Any changes to the data will be documented and stored separately from the crude database. Any recoding of data (e.g. age) will be documented.

2. Pooled analysis outline

WHO/EURO will conduct the pooled analysis. The higher sample size for this analysis will provide more power (and precision). Data can be coded as outlined in Annex 1, or a codebook can be provided by the surveillance teams to WHO/EURO that includes the variable names, descriptions and coding. WHO/EURO will perform additional data cleaning and will document and share any further data cleaning and analysis with all country coordinators to ensure it can be reproduced.

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 in each country. The pooled analysis will be carried out in a similar way to the country-specific analysis. Country will be included potentially as a fixed effect or as a random effect in a multilevel model. Statistical heterogeneity between surveillance sites will be determined, using Q-test and the I2 index.

Briefly, cases and controls will be described by baseline characteristics, and uni- and multivariable analyses performed as described in section 3.13.1 for individual analysis.

3. Bias from pooled estimates

With any multi-centre study, there is always the potential for heterogeneity among sites. With data from a number of different 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 being a participating hospital, 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. *These hospitals would not be included in the analysis.*

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

4. Pooled analysis plan

4.1 Descriptive pooled analysis

The proportion of eligible hospitalised cases and controls who were included in the surveillance will be calculated. The proportion of patients not agreeing to participate will be documented. Patients excluded will be described in a flowchart.

Cases and controls will be described by baseline characteristics.

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

- Number of hospitals participating and catchment population
- Beginning of vaccination campaigns for pandemic vaccine
  - Beginning of the analysis period
  - End of the analysis period
  - 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

4.2 Measure of effect

This analysis 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.

4.3 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.

4.4 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 anaklysis period 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)

Virus type-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.

4.5 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.

4.6 Continuous variables

Continuous variables in the COVID-19 datasets include age, time of onset of symptoms, GP visits in the previous 3 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, you may lose information, introduce residual confounding and increase the standard error of your 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 what is going on), precision and a model that estimates the vaccine effect with the least bias.

4.7 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 design and quality characteristics will also be checked. Other surveillance 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 Higgens 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.

4.8 One-stage pooled analysis approach

If sample sizes are too small to measure vaccine effectiveness controlling for all potential confounders for each individual surveillance site, a 1-stage pooled approach will be used for analysis.

Individual surveillance system data will be pooled into one dataset and analysed as a 1-stage model with surveillance site as a fixed effect. This could provide a large enough sample size to obtain (for example) an estimate of CVE early in the analysis 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 surveillance site and covariates will be carried out to determine if the effect of each covariate differs across studies, to test the assumptions of the 1-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 surveillance sites when using a 1-stage fixed effects approach (see above section). Reasons for heterogeneity need to be thoroughly investigated and the assumptions underlying the 1-stage pooling approach need to be revisited.

4.9 Controlling for hospital effect

Primary analysis will be carried out using simple logistic regression to obtain the individual site 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 1-stage pooled analysis.

Multi-level logistic regression can also be carried out for each individual surveillance site with hospital as a random effect. Then the 2-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.

4.10 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, we should consider collapsing categories, modelling continuous variables in a different way (if applicable); sensitivity analyses can be carried out using penalised logistic regression.

Two-stage pooled analysis approach

If adequate sample size by site is achieved to obtain an adjusted OR, then a 2-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 site-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 2-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 surveillance participants for whom variables were collected differently.

5. 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 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

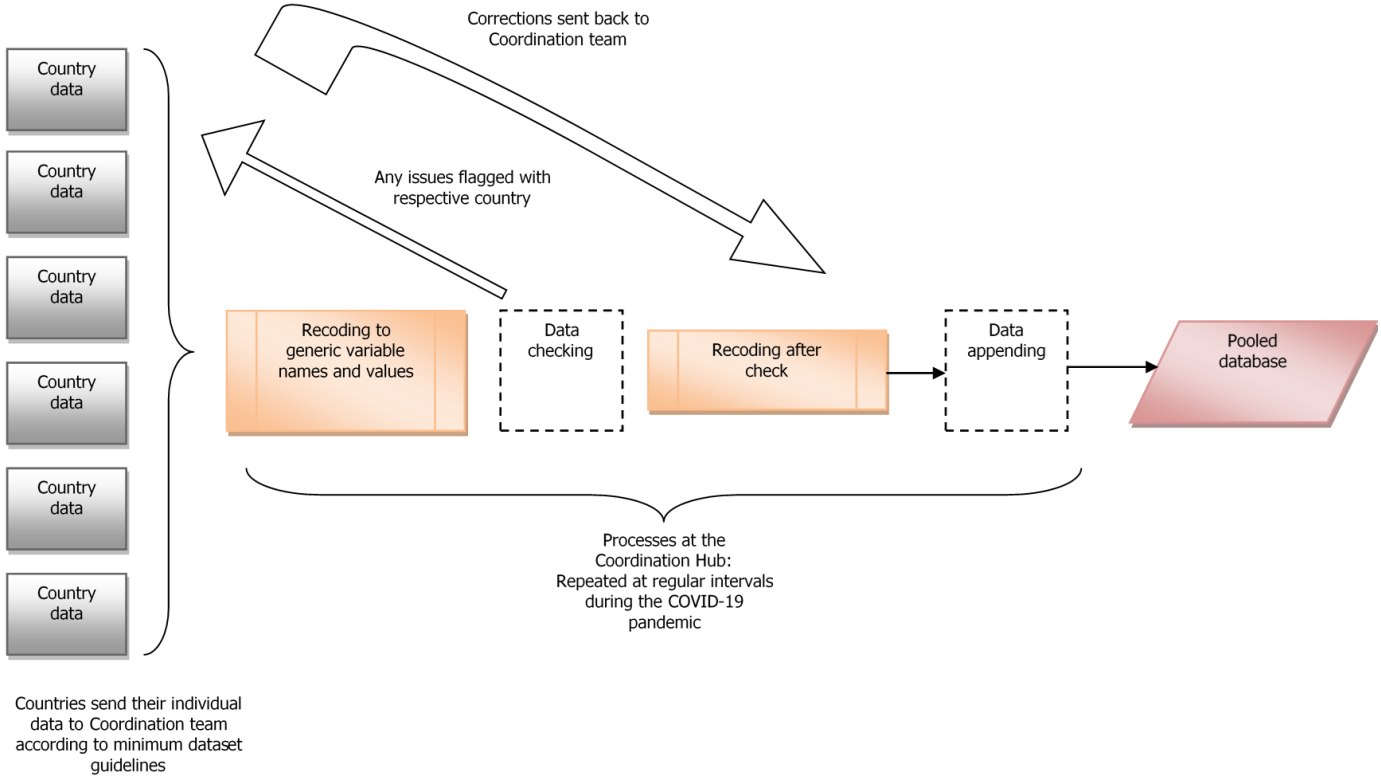
We can also put time as a variable in the model. As time may be an effect modifier (there may be different CVE at different times of the pandemic), then we can add an interaction term or perform the proposed stratified analysis.

6. 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.

Data flow for pooled dataset



Annex 5. Additional information for pooled analysis

If data are to be pooled, surveillance specifications for each country should be summarised in this annex. Each country annex should include:

- description of the hospitals participating in the surveillance system (wards involved, bed capacity, catchment population, detailed mode of recruitment including the use of computerised system to identify SARI patients)
- definition of beginning of pandemic
- pandemic (when applicable) vaccines used
- target groups for vaccination, with vaccination timelines if possible
- vaccine status ascertainment method
- details on methods for data collection, data entry and data transmission
- list of variables collected (and coding if different from suggested coding)
- data validation procedures
- laboratory issues (laboratory performing tests; tests used: PCR, antigen test, strain characterisation; methods for specimen collection, storage, transport; selection procedures for strain characterisation)
- consent (when required), ethical procedures (oral/written consent; submission to ethics committee)
- human resources needed
- provisions to train hospitals.

References

1. EURO_COVID19_Dashboard - PUBLIC [Internet]. [cited 2021 April 19]. Available from: https://who.maps.arcgis.com/apps/opsdashboard/index.html#/ead3c6475654481ca51c248d52ab9c61
2. Operational considerations for influenza surveillance in the WHO European Region during COVID-19: interim guidance (2020) (produced by WHO/Europe). 2021 Feb 8 [cited 2021 Apr 15]; Available from: https://www.euro.who.int/en/health-topics/communicable-diseases/influenza/publications/2020/operational-considerations-for-influenza-surveillance-in-the-who-european-region-during-covid-19-interim-guidance-2020-produced-by-whoeurope
3. WHO | WHO guidelines on ethical issues in public health surveillance. WHO [Internet]. 2019 [cited 2021 Apr 15]; Available from: http://www.who.int/ethics/publications/public-health-surveillance/en/
4. WHO | WHO surveillance case definitions for ILI and SARI. WHO [Internet]. 2018 [cited 2021 Apr 15]; Available from: http://www.who.int/influenza/surveillance_monitoring/ili_sari_surveillance_case_definition/en/
5. WHO COVID-19 Case definition [Internet]. [cited 2020 Dec 17]. Available from: https://www.who.int/publications/i/item/WHO-2019-nCoV-Surveillance_Case_Definition-2020.1
6. WHO | Laboratory biosafety guidelines and support. WHO [Internet]. 2018 [cited 2021 Apr 15]; Available from: http://www.who.int/in-vitro-diagnostic/biosafety-guidelines/en/
7. Operational considerations to expedite genomic sequencing component of GISRS surveillance of SARS-CoV-2 [Internet]. [cited 2021 Apr 13]. Available from: https://www.who.int/publications/i/item/WHO-2019-nCoV-genomic-sequencing-GISRS-2021.1
8. WHO/Europe | Health emergencies - Health workers at risk, older adults and residents of long-term care facilities to be prioritized for COVID-19 vaccination [Internet]. [cited 2020 Dec 4]. Available from: https://www.euro.who.int/en/health-topics/health-emergencies/pages/news/news/2020/11/health-workers-at-risk,-older-adults-and-residents-of-long-term-care-facilities-to-be-prioritized-for-covid-19-vaccination
9. World Health Organization. WHO SAGE Roadmap for Prioritizing Uses of COVID-19 Vaccines in the Context of Limited Supply [Internet]. [cited 2020 Oct 29]. Available from: www.who.int/docs/default-source/immunization/sage/covid/sage-prioritization-roadmap-covid19-vaccines.pdf?Status=Temp&sfvrsn=bf227443_2&ua=1
10. Wade DT, Collin C. The barthel ADL index: A standard measure of physical disability? Disabil Rehabil [Internet]. 1988 [cited 2021 Apr 15]; 10(2):64–7. Available from: https:/pubmed.ncbi.nlm.nih.gov/3042746/

The WHO Regional Office for Europe

The World Health Organization (WHO) is a specialized agency of the United Nations created in 1948 with the primary responsibility for international health matters and public health. The WHO Regional Office for Europe is one of six regional offices throughout the world, each with its own programme geared to the particular health conditions of the countries it serves.

Member States

Albania
Andorra
Armenia
Austria
Azerbaijan
Belarus
Belgium
Bosnia and Herzegovina
Bulgaria
Croatia
Cyprus
Czechia
Denmark
Estonia
Finland
France
Georgia
Germany
Greece
Hungary
Iceland
Ireland
Israel
Italy
Kazakhstan
Kyrgyzstan
Latvia
Lithuania
Luxembourg
Malta
Monaco
Montenegro
Netherlands
North Macedonia
Norway
Poland
Portugal
Republic of Moldova
Romania
Russian Federation
San Marino
Serbia
Slovakia
Slovenia
Spain
Sweden
Switzerland
Tajikistan
Turkey
Turkmenistan
Ukraine
United Kingdom
Uzbekistan

Document number:
WHO/EURO:2021-2481-42237-58308

World Health Organization
Regional Office for Europe

UN City, Marmorvej 51,
DK-2100 Copenhagen Ø, Denmark
Tel.: +45 45 33 70 00 Fax: +45 45 33 70 01
Email: eurocontact@who.int
Website: www.euro.who.int