European Union
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003673

I-MOVE-COVID-19 network

Multidisciplinary European network for research, prevention and control of the COVID-19 pandemic

European study of risk factors for severe disease among hospitalised COVID-19 patients

Draft generic protocol

MAY 2020

V 04

I-MOVE-COVID-19 Network
WP4 coordinated by Epiconcept

Based on: current literature, I-MOVE generic influenza protocol for hospitalised older adults 2019–2020

Table of contents

Abbreviations

COVID-19	Coronavirus disease 2019
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

- *Arrow marks with italicised text indicate the points that study sites should adapt and provide details for in their study annexes*

1 Background

The end of 2019 saw the emergence of a novel severe acute respiratory syndrome – coronavirus 2 (SARS-CoV-2), which causes coronavirus disease 2019 (COVID-19). By the end of March 2020, there were over 400,000 cases of COVID-19 reported globally by over 150 countries, with an increasing proportion from countries in the European Union/European Economic Area (EU/EEA), as well as the United Kingdom (UK). As of 25 March, there were 204,930 cases and 11,810 deaths reported in the EU/EEA and the UK, from where the European Centre for Disease Prevention and Control (ECDC) reports that the number of reported COVID-19 cases is increasing rapidly, with an increase in notifications of a similar trajectory to Hubei province in late January/early February and in Italy in late February/early March.(1)

Data reported to ECDC show that clinical presentations of COVID-19 range from no symptoms (asymptomatic) to severe pneumonia, and that severe disease can lead to death. Thirty percent of diagnosed COVID-19 cases in the EU/EEA with available data were hospitalised, while 4% had severe illness. Hospitalisation rates were higher for older adults (60 years+). Estimates showed that the risk as well as the absolute numbers of deaths rapidly increased with age for those 60 years and older in Germany, Italy and Spain. Among hospitalised patients, severe illness was reported in 15%, and death occurred in 12%, with higher case fatality in older adults.

ECDC recommends that all hospitalised severe acute respiratory illness (SARI) patients be tested for SARS-CoV-2 virus, so as to detect community transmission, nosocomial outbreaks, as well as for monitoring the intensity and impact of the pandemic.(1) There are also several questions to which urgent answers are needed to improve our understanding of SARS-CoV-2, to inform us of the best interventions to prevent or delay the spread of COVID-19 and newly recommended treatment strategies.

The first question is whether there are any early clinical signs (at home and in hospital) or positive exam results which are more likely in those with severe COVID-19. Secondly, we need to know whether a delay between the time of onset and hospital admission is itself a risk factor. Thirdly, we should consider whether any type or timing of exposures to index cases infect more and more severely than others (this would be especially important for healthcare workers). Also, whether there are any in-hospital support and treatments that seem to work better than others, or whether any underlying conditions and their respective treatments that could be risk factors and for which an intervention is available (e.g. changing hypertension treatment for the second wave of the pandemic).

Importantly, as we are now in a COVID-19 pandemic, having a pre-existing, well-established European platform to rapidly provide severe acute respiratory illness (SARI) surveillance already in place is allowing for immediate case identification and, once available, will allow for the rapid evaluation of any pandemic vaccine and adaptation of preventive and control strategies.

I-MOVE (Influenza – Monitoring Vaccine Effectiveness in Europe), first established in 2007,(2) was the first network to monitor influenza vaccine effectiveness (VE) within and across the seasons in the European Union (EU) and the European Economic Area (EEA). The network has two components, one for primary care practices,

recruiting patients with influenza-like illness (ILI) and the other for hospitals, recruiting patients with severe acute respiratory illness (SARI).

In February 2020, many partners, already involved in studies within the I-MOVE network, came together as the I-MOVE-COVID-19 consortium, and were successful in a bid for the European Commission H2020 call on "Advancing knowledge for the clinical and public health response to the novel coronavirus epidemic".

The I-MOVE-COVID-19 consortium aims to obtain epidemiological and clinical information on patients with COVID-19 as well as virological information on SARS-CoV-2, through different work packages (WPs): (a) provision of a flexible surveillance platform, adaptable to the epidemiological situation, through WP2 (primary care surveillance) and WP3 (hospital surveillance), (b) research studies, through WP4 and (c) evaluation of public health interventions (e.g. vaccination, antivirals) in WP2–4, in order to contribute to the knowledge base, guide patient management, and inform the public health response. This will be achieved through adaptation and expansion of the existing I-MOVE network to include COVID-19. The network includes primary care networks, hospitals, and national laboratory reference centres in 10 countries across the WHO European Region.[1]

The WP3 hospital surveillance for COVID-19 is coordinated by Public Health Scotland (PHS) with Epiconcept support. The hospital network comprises 11 study sites involving 13 hospitals in five EU Member States[2] (MS) and Albania, the intensive care/high dependency unit (ICU/HDU) network from all hospitals in England and the hospital network in Scotland (where coverage is not yet at 100%, but the aim is to include all hospitals). The laboratory component of the network includes regional and national reference centres from the participating countries. Hospital-based studies will utilise data collected from this surveillance network.

The WP4 (research studies) for COVID-19 is coordinated by Epiconcept. While each of the study sites can analyse their data separately, pooling the data for an overall analysis will provide a sample size big enough to answer study questions with reasonable precision. Participating study sites will carry out studies in hospitals, except for England, where surveillance will be limited to ICUs in all hospitals.

This document presents the core European protocol for the hospital-based study of potential risk and protective factors for severe COVID-19 in hospitalised COVID-19 patients, outlining the agreed methods for collecting COVID-19 and SARS-CoV-2 in each of the individual studies, and including a plan for the pooled analysis. This generic protocol will be updated according to the final pandemic strategy (treatment and vaccine products available) in each of the participating sites, the extent of the virus circulation and the identification of new groups at risk or new protective or risk factors, e.g. at the time when the vaccine will be available. Importantly, this study is part of the I-MOVE network; as for I-MOVE influenza, the SARI definition is used to recruit patients. However, for the I-MOVE-COVID-19 risk factor study, we will only include patients with COVID-19

[1]Albania, France, Ireland, Lithuania, the Netherlands, Portugal, Romania, Spain, Sweden, and the UK (England and Scotland).

[2]France, Lithuania, Portugal, Romania, and Spain.

diagnoses. The network will also be used to estimate COVID-19 vaccine effectiveness in the future, once vaccines are available (the protocol will be adapted at this point).

2 Objectives

2.1 Objectives

The primary objective will be to identify key risk factors for and protective factors against **severe** disease among hospitalised COVID-19 patients in the following categories:

- Patient characteristics
  - demographic
  - comorbid chronic conditions and their treatment
  - living in a close contact setting
- Clinical characteristics before and on/after admission
  - laboratory or examination results
  - symptoms and signs of respiratory co-infections
- Timing
  - from symptom onset to admission (hospital and ICU/HDU)
  - from symptom onset to death (for fatal events)
  - from admission (hospital and ICU/HDU) to death/discharge
  - from symptom onset to death/discharge
- Administration of medications and interventions pre-admission and in hospital
- Complications

in order to provide up-to-date information for the public health response at hospital level by guiding patient management and highlighting target groups at risk for severe disease.

- *Each study site to specify the objectives of their study*

3 Methods

3.1 Study design

To investigate risk factors among severe cases, we will conduct a cohort study, in which the cohort will be all probable and confirmed hospitalised COVID-19 patients recruited in participating hospitals.

- At study site level: hospital-based study in each participating hospital
- At European level: multicentre hospital-based study in several countries/regions

3.2 Study population

The study population comprises all SARI patients (suspected COVID-19 patients) admitted to participating hospitals, from which the cohort will be defined as those diagnosed in the following categories:

- Probable COVID-19 cases
- Confirmed COVID-19 cases
    - *Study sites to describe the setting (number of hospitals included, number of beds, number and type of wards/specialties/services included)*
    - *Study sites to indicate whether treatment protocol in hospital or in ICUs/HDUs follows a standardised protocol or whether protocols are adjusted according to biology/physiology*

3.3 Study period

The study period starts in *[month]* 2020. Participating hospitals continue the study throughout the year until they achieve the sample size.

- *Study sites to define the beginning of this hospital-based study (month/year) and, if possible to estimate, the planned end month/year for the study in their hospital(s)*

3.4 Outcome

The primary outcome of interest will be severe hospitalised COVID-19 patients. For the different categories of "severe", see below (Section 3.5.5).

3.5 Case definitions

3.5.1 Hospitalised patient

A hospitalised patient will be defined as a SARI patient who has been admitted to one of the participating hospitals during the study period, and has not been discharged home or home-equivalent within 24h.

3.5.2 SARI patient (suspected COVID-19 case)

A SARI patient will be defined as a hospitalised person with:

- at least one systemic symptom or sign: fever or feverishness, malaise, headache or myalgia or deterioration of general condition (asthenia or loss of weight or anorexia or confusion or dizziness)

AND

- at least one respiratory symptom or sign (cough, sore throat or shortness of breath; **or** tachypnoea **or** signs of low oxygen saturation)

at admission or within 48 hours after admission.

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

All SARI patients, until they are re-classified as COVID-19 negative, probable or confirmed (see below), will be considered as suspected COVID-19 patients.

3.5.3 SARI confirmed as COVID-19 (confirmed case)

A confirmed COVID-19 case will be defined as a patient hospitalised with SARI with a respiratory sample positive for SARS-CoV-2 (4).

3.5.4 SARI probable COVID-19 (probable case)

A probable COVID-19 case will be defined as a patient hospitalised with SARI for whom

- testing for SARS-CoV-2 is inconclusive (according to the test results reported by the laboratory) (4)

OR

- testing was positive on a pan-coronavirus assay (4)

OR

- no laboratory tests are available but there is clinical confirmation with suggestive radiology

3.5.5 Severe COVID-19 case (main outcome)

A severe COVID-19 patient will be defined as a patient hospitalised with SARI and diagnosed as COVID-19 who has any of the following clinical courses or outcomes mentioned in the admission or discharge diagnosis:

- Bilateral pneumonia with ground-glass opacities

- Admitted to ICU/HDU
- On ventilation
    - invasive (i.e. with intubation)
    - non-invasive (e.g. high-flow oxygen; or those needing >6L)
- Extracorporeal membrane oxygenation (ECMO)
- Death

A COVID-19 death is defined as a probable or confirmed COVID-19 case who died during his/her hospitalisation.

- *Study sites to specify if there is any deviation from the case definition above (e.g. minimum length of hospital stay, additional confirmation[3])*
- *Study sites to specify if any adaptations to ventilation, ECMO use, etc. in their hospitals (e.g. limiting use of ECMO for specific age-groups/comorbid conditions/staff availability, etc.)*
- *Study sites to define a clinical confirmed COVID-19 case*
- *Study sites to list the complications included in their protocol and the definition of severity*
- *Study sites to specify how death ascertainment is defined*

3.5.6 Exclusion criteria

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

- is unwilling to participate or unable to communicate and give consent (the consent may also be given by her/his legal representative, or by specific consent procedures, acceptable according to the local ethical review process)

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

3.6 SARI patient identification – algorithm for patient inclusion

Table 1: List of diagnosis codes for which patients could be screened for onset of SARI symptoms, I-MOVE-COVID-19 hospital-based severity risk factor study.

Category	Morbidity	ICD-9	ICD-10
Cough		786.2	R05

[3]Note: some sites may have additional criteria. For example: a confirmed case may be defined as a SARI patient with symptoms and signs and radiological confirmation of COVID, in accordance with the Belgian Sciensano guidelines.

Influenza-like illness	Difficulty breathing	786.05	R06
	Sore throat	784.1	R07.0
	Dysphagia	787.20	R13
	Fever	780.6	R50.9
	Headache	784.0	R51
	Myalgia	729.1	M79.1
	Fatigue/malaise	780.79	R53.1, R53.81, R53.83
Cardiovascular diagnosis	Acute myocardial infarction or acute coronary syndrome	410-411, 413-414	I20-23, I24-25
	Heart failure	428 to 429.0	I50, I51
Respiratory diagnosis	Emphysema	492	J43.9
	Chronic obstructive pulmonary disease	496	J44.9
	Asthma	493	J45
	Myalgia	729.1	M79.1
	Dyspnoea/respiratory abnormality	786.0	R06.0
	Respiratory abnormality	786.00	R06.9
	Shortness of breath	786.05	R06.02
	Tachypnoea	786.06	R06.82
	Other respiratory abnormalities	786.09	R06.00, R06.09, R06.3, R06.89
Infections	Pneumonia and influenza	480-488.1	J09-J18
	Other acute lower respiratory infections	466, 519.8	J20-J22
	Viral infection, unspecified	790.8	B34.9
	Bacterial infection, unspecified	041.9	A49.9
	Myocarditis	429.0	I40.9
	Bronchitis	490, 491	J40, 41
Inflammation	SIRS* non-infectious without acute organ dysfunction	995.93	R65.10
	SIRS* non-infectious with acute organ dysfunction	995.94	R65.11
Abdominal symptoms	Vomiting	787.0	R11
	Diarrhoea	009.3, 787.91	A07.9, K52.9
	Abdominal pain	789.0	R10

Diagnoses related to deterioration of general condition or functional status	General physical deterioration, lethargy, tiredness	780.79	R53.1, R53.81, R53.83
	Anorexia	783.0	R63.0
	Feeding difficulties	783.3	R63.3
	Abnormal weight loss	783.21	R63.4
	Other symptoms and signs concerning food and fluid intake	783.9	R63.8
	Disorientation/altered mental status	780.97	R41.0
	Dizziness and giddiness	780.4	R42
	Infective delirium	293.0, 293.1	F05
	Coma	780.01	R40.2
	Transient alteration of awareness	780.02	R40.4
	Other alteration of consciousness (somnolence, stupor)	780.09	R40.0, R40.1
	Febrile convulsions (simple), unspecified	780.31	R56.00
	Complex febrile convulsions	780.32	R56.01
Other	Anosmia, ageusia, myalgia	781.1, 729.1	R43.0, R43.2, M79.1

**SIRS: Systemic inflammatory response syndrome*

The COVID-19 patients will be identified among SARI patients hospitalised for at least 24 hours in one of the participating hospitals.

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

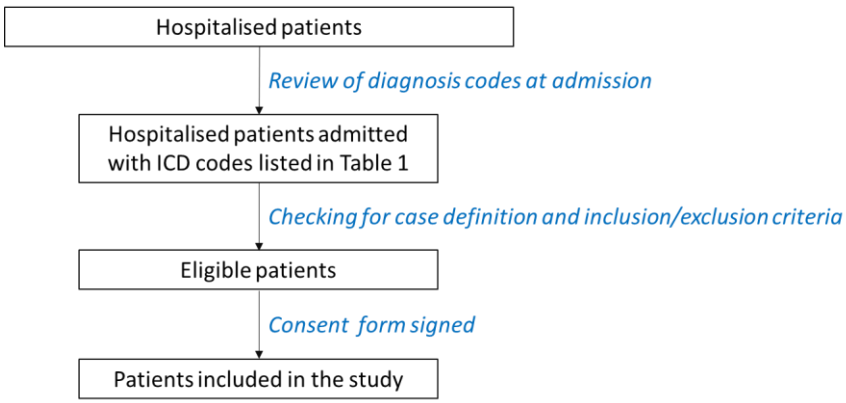


Figure 1: proposed inclusion algorithm for hospitals/services relying on common use of ICD codes, I-MOVE-COVID-19 hospital-based COVID-19 severity risk factor study.

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

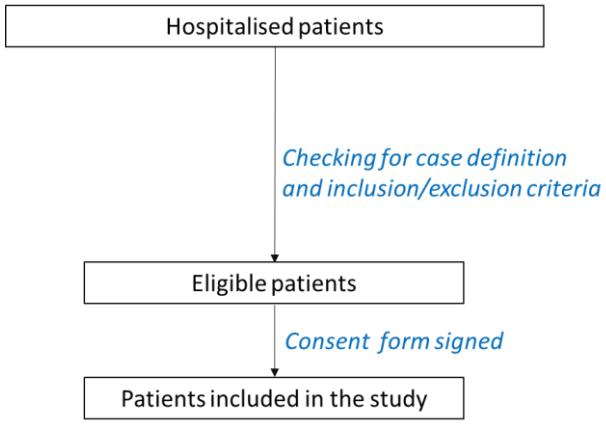


Figure 2: proposed inclusion algorithm for hospitals/services systematic screening of all admitted patients, I-MOVE-COVID-19 hospital-based COVID-19 severity risk factor study.

For some hospitals, parts of the study may be retrospective (e.g. when "catching-up" already diagnosed patients). Here, study sites may choose to start from the laboratory and/or imaging results to identify potential confirmed and probable cases, and from these to trace back to patient records to identify those who were SARI patients at admission for inclusion.

Following the procedures outlined by each study, patients meeting the SARI case definition will be asked (directly or by telephone through their legal tutor or physician) to provide consent to participate in the study and a nasal/throat respiratory specimen for COVID-19 testing. (Note that during the pandemic, for some study sites this procedure of swabbing may be the default.)

- *Each study site to describe procedures used to identify study participants*

In case of test scarcity, or budget limited to a certain number of patients' inclusion, the study sites may need to switch from exhaustive to systematic sampling (e.g. inclusion of patients every second day). Systematic sampling procedures should be planned ahead by the study sites. During the period of systematic selection, the study sites will make sure to document the sampling fraction.

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

3.7 Laboratory methods

Study nurses or physicians will collect respiratory specimens (see Section 4.4) from all eligible patients, respecting safety standards for COVID-19 and following WHO biosafety guidelines.[4]

- *Each study site to describe the type and number of swabs taken by patient*

The ECDC-recommended SARS-CoV-2 laboratory confirmation is by viral RNA detection with nucleic acid amplification tests, such as RT-PCR.(1,4) Isolates will undergo molecular analysis for currently circulating SARS-CoV-2 virus. During the influenza season, it is recommended that influenza virus tests should also be performed, as long as there is circulation of influenza viruses.(1)

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

- *Each study site to describe the laboratory procedures (samples taken, storage, transport)*
- *Each study site to describe the tests and the kits used (and their sensitivity, specificity, PPV) for COVID-19 and, if needed, other respiratory virus detection*
- *Each study site to describe if the laboratory participates in QA/QC (Quality Assurance/Quality Control) schemes*
- *Each study site to describe the selection of specimens and the procedures for genetic and antigenic characterisation, where appropriate (see Annex 4 for an example of presentation of these results)*
- *Each study site to describe genetic and antigenic analyses and specify sequencing methods*
- *Study sites to describe whether specimens are tested for other respiratory viruses (e.g. whether influenza continues to be tested systematically during the season and stops once the influenza season is over, or is only tested when the COVID-19 result is negative, etc.)*

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

3.8 Exposure (risk and preventive factors)

Note: Some potential exposures will be RFs, some will be outcomes (in particular, the list of complications), and some may be on the causal pathway for COVID-19; how they are used in analysis may vary and will depend on the analysis being performed.

3.8.1 Definitions

Risk factor for severe COVID-19

- Any factor which has influenced the patient's exposure, susceptibility, or response to COVID-19, in such a way as to make them **more** likely to develop severe disease.

Protective factor against severe COVID-19

- Any factor which has influenced the patient's exposure, susceptibility, or response to COVID-19, in such a way as to make them **less** likely to develop severe disease.

The following factors will be included in the study to be investigated as exposures which are potential risk or protective factors.

1. Patient characteristics
2. Close contact setting
3. Time onset to admission
4. Exam/lab results before/at admission or during hospitalisation
   - CT scan/ultrasound (or CXR if no CT/ultrasound)
   - Oxygen saturation (%) before support
   - Other lab results (biochemistry, routine blood tests, etc.)
   - COVID-19 lab results and sequencing, where available (genetic group and antigenic group/clade)
5. Pre-existing chronic conditions
   - anaemia/chronic haematologic disease, asplenia, asthma
   - cancer, chronic liver disease/cirrhosis
   - dementia, diabetes mellitus
   - heart disease (excluding hypertension), hypertension
   - Immunodeficiency/organ transplant
   - lung disease, liver disease, neuromuscular disorders
   - obesity, renal disease (exclude acute renal failure), rheumatologic diseases
   - stroke, tuberculosis

3.9 Collected information

Collected information falls under the following seven main categories: study identifiers, hospital/ward information, patient characteristics, case/severity definitions, risk factors (including laboratory sequencing information), medications/interventions in hospital, and complications. These are listed in Table 2 below for rapid reference and are then described in more detail following the table. See also Annex 1 for a complete variable list including coding.

Table 2: Variable list by category and type (core: surveillance and additional: risk factor study only)

Category	**Variable list**	
	Core (surveillance) variables	**Additional variables: risk factor study only**
1. Study identifiers	Country/site, hospital, patient unique id, consent (where relevant)	Reasons for refusal
2. Hospital/ward information	Previous hospital stay, first ward of referral, date admission/ discharge (hospital, ICU/HDU), date sample, prior hospital stay (last 12 months)	
3. Patient characteristics	Sex, age, smoking, pregnant, healthcare worker, residence	Clinical frailty score at admission
4. Case/severity definitions	SARI signs/symptoms, date onset, COVID tests/results, severity indicators (bilateral pneumonia, ventilation, ECMO, admit ICU/ HDU, death)	
5. Risk factors	5.1. Patient characteristics	
		5.2. Pre-symptomatic treatment/intervention • ACE inhibitors, ARBs, NSAIDs, statins, metformin, steroids, corticosteroids, DMARDs, chemotherapy, gliclazides, psycho-tropics, antivirals • influenza and pneumococcal vaccination BCG vaccination
	5.3. Close contact setting	
		5.4. Early clinical symptoms/signs • fever, cough, shortness of breath • anosmia, ageusia, etc. as in (3)
	5.5. Time onset to admission	

	5.6. Exam/lab results on admission or during hospitalisation
	• CT scan (or CXR if no CT)
	• Oxygen saturation (%) before support
	• Lab sequencing
	  ○ lab results (genetic group and antigenic group/clade)
		• Exam /lab results on admission or during hospitalisation
		  • Blood group
		  • Confusion **or** GCS **or** AVPU score Blood urea, respiratory rate
		  • Systolic/diastolic BP (mmHg)
		  • Lymphocyte, platelet counts
		  • Neutrophil/lymphocyte ratio
		  • LFT (ALP, AST, ALT)
		  • Ferritin, LDH, D-Dimer
		  • IL-6, eosinophil count, CRP
		  • Creatine phosphokinase
		  • Troponin-I

		• Triglycerides, cholesterol • Fibrinogen • NT-proBNP • HbA1c
	5.7. Pre-existing chronic conditions • anaemia/chronic haematologic disease, asplenia, asthma • cancer, chronic liver disease/cirrhosis • dementia, diabetes mellitus • heart disease (excluding hypertension), hypertension • Immunodeficiency/organ transplant • lung disease, liver disease, neuromuscular disorders • obesity, renal disease (exclude acute renal failure), rheumatologic diseases • stroke, tuberculosis	
		5.8. Other respiratory viruses
6. In-hospital medications/ interventions	Prone position, ventilation	Hydroxychloroquine, corticosteroids, monoclonal antibody/IL6 blockers, antibiotics, oxygen (nasal, high-flow), antivirals, sepsis fluid resuscitation, study drugs, nebuliser, etc.
7. Complications		ARDS, bronchiolitis, encephalitis, myocarditis, pneumonia (secondary bacterial), other secondary bacterial infection, sepsis, acute renal injury, heart failure, multiorgan failure, diarrhoea/colitis, dermatological manifestations of COVID-19, ICU-related myopathy, death

Note: variables in grey already listed in another category.

3.9.1 Study identifiers

We will document the following study characteristics.

- Country, site
- Hospital
- 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)

3.9.2 Hospital/ward information

We will document following dates and other hospital information to monitor severity.

- First ward of referral
- Hospital stay in previous 12 months
- Date of admission (hospital, ICU/HDU)
- Date of discharge (hospital, ICU/HDU)

    - Whether multiple admissions to ICU/HDU (total # times)
    - Length of time in ICU (total), where known
- Date of swab/sample

3.9.3 Patient characteristics

We will document following patient characteristics to describe the study population.

- Age
- Sex
- Smoking history
- Pregnancy
- Healthcare worker
- Place of residence
    - home/institutionalised
    - postcode where possible
    - pre-hospitalisation dependence on home support/care
- Clinical frailty score at admission (where possible)

➢ *Each surveillance site to describe type of clinical frailty score in use, where available*

Smoking history will be collected and coded as follows: never-smoker, former smoker (stopped smoking at least 1 year before inclusion in the study), current smoker.

Information on pregnancy will be collected and coded as follows: pregnant now or gave birth within the past 6 weeks: yes/no/unknown; if Yes: unknown trimester, trimester 1–3; <6 weeks post partum.

3.9.4 Case definition and severity information

We will document the following information to better understand the clinical spectrum of disease as well as to classify patients as suspected, probable and confirmed cases, and further as severe cases:

- Case definitions (all)
    - SARI symptoms and signs
        - fever or feverishness
        - cough
        - sore throat
        - shortness of breath
        - sudden onset
    - Other symptoms
        - tachypnoea or other signs of low oxygen saturation (restlessness)
        - rapid heartbeat
        - chest pain

        - coryza
        - malaise
        - headache
        - myalgia
        - deterioration of general condition (asthenia, weight loss, anorexia)
        - confusion
        - dizziness
        - diarrhoea
        - abdominal pain
        - nausea
        - vomiting
        - ageusia
        - anosmia
        - conjunctivitis
        - rash or other dermatological manifestations
    - Date of onset of first symptom
    - COVID-19 test(s) and laboratory results
        - on admission
        - including information on antigenic and genetic analysis, when available
- Case definitions (severe): as above plus any of the following
    - Bilateral pneumonia with ground-glass opacities (from radiological imaging results)
    - Invasive/non-invasive mechanical ventilation
    - ECMO
    - ICU/HDU admission
    - Death
        - Date of death, cause of death

3.9.5 Potential risk or preventive factors for severe COVID-19

The following will be included in surveillance data to be collected, as exposures which are potential risk or protective factors.

Close contact setting

Information on any close contact setting with a person with a probable or confirmed case in the 14 days prior to symptom onset will be collected, using the following settings:

- Family setting
- Healthcare setting
- Workplace
- Long-term care facility
- Prison setting
- Other setting
- Unknown

Time from onset to admission

It is possible that delayed admission could lead to more severe COVID-19 as the disease will have progressed prior to hospitalisation. We will collect information on date of onset and date of admission to investigate delayed admission as a risk factor for severity.

Laboratory or examination results on admission or during hospitalisation

Some laboratory or examination results will help with the classification of cases into severe or non-severe COVID-19.

- Computed tomography (CT) scan
- Chest X-ray findings, when CT scan is unavailable or not included in routine testing
- Oxygen saturation on admission to hospital (on air), %
- Laboratory sequencing results (genetic group, antigenic group/clade, as appropriate and where possible)

Other laboratory results (e.g. low lymphocyte count and high C-reactive protein or CRP) and examination findings (e.g. abnormal CT) have been suggestive of more severe COVID-19.

We will collect information on the following additional examinations to investigate predictive capacity, for the risk factor study, where possible:

- Blood urea nitrogen (BUN)
- Urea
- Respiratory rate
- Heart rate
- SBP, DBP
- ABO blood group (only at admission)
- Platelet count
- Lymphocyte count (range: <500, 500–800, >800)
- Neutrophil/lymphocyte ratio
- Eosinophil count
- Liver function test (LFT) including
    - Serum albumin
    - Serum total bilirubin
    - Serum conjugated bilirubin
    - GGT
    - Serum total protein
    - Alkaline phosphatase (ALP)
    - Aspartate aminotransferase (AST)
    - Alanine transaminase (ALT)
    - Prothrombin time
- Lactic acid dehydrogenase (LDH)

- Ferritin
- D-Dimer
- Fibrinogen
- C-reactive protein (CRP)
- Creatinine phosphokinase
- Cardiac troponin-I
- Triglycerides
- Cholesterol
    - Low density lipoprotein (LDL)
    - High density lipoprotein (HDL)
    - Total cholesterol
- N-terminal-prohormone B-type natriuretic peptide (NT-proBNP)
- B-type natriuretic peptide (proBNP)
- Single glycated haemoglobin (HbA1c)

Information on multiple admissions

Among the hospitalised SARI patients, some may be admitted several times within the same episode (within 14 days from onset). We will collect information on those who have had more than one admission, to inform total length of hospital stay. Where possible, we will collect whether

- Patient had a previous admission
    - number of admissions (from 1 to 6 previous admissions)
    - total number of days spent in hospital from all admissions combined

Pre-existing chronic conditions

The following underlying conditions will be collected and described as potential factors impacting COVID-19 severity (for ICD codes see Table 3).

- anaemia / chronic haematologic disease
- asplenia
- asthma
- cancer (solid organ and haematological)
- chronic liver disease/cirrhosis
- dementia
- diabetes mellitus
- heart disease (excluding hypertension)
- hypertension
- immunodeficiency and organ transplant
- lung disease
    - use of non-invasive ventilation / oxygen therapy
    - chronic obstructive pulmonary disease (COPD)
- neuromuscular disorders
- obesity
    - height
    - weight

- or: BMI[5] (sites to include whichever is feasible/available)
- renal disease (exclude acute renal failure)
- rheumatologic diseases
- stroke
- tuberculosis

➢ *Each surveillance site to define the list of chronic conditions to be included and describe what the source of information will be*

Data will also be collected on whether the patient had visited a GP or other primary care service in the 14 days prior to admission.

Table 3: ICD-9 and ICD-10 codes for 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

[5]Note: obesity defined as BMI>29.

			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
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 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,
		M34.81, M05.10	**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

3.9.6 Administration of medications and interventions in hospital

Administering medications and/or interventions in may have a protective or deleterious effect towards the development of severe COVID-19. We will document whether the patients received any of the following while in hospital:

- Nasal oxygen
- Prone position
- Nebuliser treatment
- Ventilation
    - positive end-expiratory pressure (PEEP)
    - bilevel positive airway pressure (BiPAP)
    - continuous positive airway pressure (CPAP)
- High-flow oxygen
    - 6L/min or higher
    - using OptiFlow technique

    - date of increase in oxygen concentration
- Sepsis fluid resuscitation
- Hydroxychloroquine/chloroquine
- Corticosteroids
    - with doses, where/if possible
- Antivirals (remdesivir, ritonavir, lopinavir, favipiravir, umifenovir)
    - with doses, where/if possible
- Monoclonal antibodies/IL-6 blockers (e.g. tocilizumab)
- Other drugs like azithromycin, ribavirin
- Use of study drugs
    - convalescent plasma
    - GM-CSF
- Other (specify)

➢ *Each study site to list any medications and interventions administered in hospital*

Pre-symptomatic treatment or intervention (including vaccination)

We will document whether the patients received any of the listed medications in the 2 weeks preceding symptom onset.

Pre-symptomatic medication status:

The three main medications to be included are angiotensin-converting enzyme (ACE) inhibitors, angiotensin II receptor blockers (ARBs) and non-steroidal anti-inflammatory drugs (NSAIDs). Additional medications include antivirals, statins and other anti-hypertensive medication. For each of these:

- An individual will be considered as "on" the medication if s/he has received at least one dose on or before onset of symptoms
- An individual will be considered as "not on" the drug if s/he did not receive it on or before symptom onset.

Current list of pre-symptomatic medications:

- Antivirals
- Statins
- ACE inhibitors
- ARBs
- NSAIDs
- Chloroquine/hydroxychloroquine
- Metformin
- Steroids

- Corticosteroids
- Monoclonal antibodies/IL-6 blockers
- Other biological disease-modifying anti-rheumatic drugs (DMARDs)
    - rituximab, tocilizumab, etc.
- Current/recent cancer chemotherapy (within previous 6 months)
- Gliclazide (for diabetes or heart failure)
- Psychotropic drugs (including benzodiazepine, etc.)

Pre-symptomatic medication use status ascertainment

Medication history includes date the patient started on the medication, where known, or just the year, if the patient was known to have been on one of the medication types before epidemic started or if the precise date is unknown. If both of these are unknown, then a simple yes/no response for each medication as to whether the patient was on the drug prior to symptom onset will be used.

The sources of information for pre-symptomatic medication status may include:

- consultation of the patient's hospital record
- (telephone) interview with the patient's GP
- (telephone) interview with the patient's pharmacist
- data from the patient's insurance company showing evidence of pharmacy delivery or reimbursement for these medications since 01 January 2020
- interview of the patient and/or his/her relatives (by telephone)

➢ *Each surveillance site to describe how pre-symptomatic medication use is collected and ascertained*

Current list of pre-symptomatic vaccinations:

- Bacille Calmette Guérin (BCG) vaccination (ever; with approximate year, if possible)
- Recent seasonal influenza vaccination (with date)
- Latest pneumococcal vaccination (with date)

Any prior BCG vaccination, as well as 2019/20 vaccination against influenza and vaccination against pneumococcal diseases will be collected. This may not be possible for BCG vaccination in some sites, but instead these could indicate if there was no BCG vaccination policy at all, or if e.g. "assume that those born after 1956 would be vaccinated", etc.

The sources of information for vaccination may include:

- vaccination registry
- consultation of the patient's vaccination card
- (telephone) interview with the patient's GP
- (telephone) interview with the patient's pharmacist
- data from the patient's insurance company showing evidence of pharmacy delivery or reimbursement of influenza vaccine during the current influenza season.

- interview of the patient and/or his/her relatives (by telephone)
  - *Each surveillance site to describe how influenza and pneumococcal vaccination status are documented*
  - *Each surveillance site to describe BCG vaccination status determination or national-level assumption*

In addition, once available, information will be collected on COVID-19 pandemic vaccination including number of doses, date, and product.

Presence of early clinical symptoms/signs

There are some early clinical symptoms or signs which are hypothesised as being predictive for COVID-19. We will collect information on them to investigate whether they are potential factors impacting COVID-19 severity (note: these are all signs already included in the general symptoms/signs on admission, but will be collected also as pre-admission symptoms/signs).

Other respiratory viruses

Patients admitted with underlying lung diseases may be included due to an exacerbation of underlying conditions unrelated to SARI. Due to their underlying conditions, these patients may be more likely to be infected with COVID-19, or to develop more severe disease than the source population. *We will collect information on the presence of respiratory infection among all COVID-19 patients.*

- *Each surveillance site to list the other respiratory infection viruses tested for (including influenza)*

Complications

Information on the following in-hospital complications will be collected as potential risk factors for severe disease:

- ARDS (acute respiratory distress syndrome)
- Bronchiolitis
- Encephalitis
- Myocarditis
- Pneumonia (secondary bacterial pneumonia)
- Other secondary bacterial infection
- Sepsis
- Acute renal injury (AKI)
- Heart failure
- Multiorgan failure
- Diarrhoea/colitis
- Dermatological manifestations of COVID-19
- ICU-related polyneuromyopathy

- Guillain Barré Syndrome
- Death
- Other (specify)

3.9.7 Sources of information

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

- hospital medical records (with regular data extraction for main variables)
- interview with patient or his/her family
- interview with patient's GP
- interview with patient's pharmacist
- vaccination register
- laboratory

➢ *Each study site to define the sources of information used for each variable collected*

3.10 Data considerations

3.10.1 Sample size

The number of individuals included in the RF study will depend on the number of hospitalised SARI patients and the number diagnosed as COVID-19. All patients meeting the eligibility criteria and giving consent to participate will be included.

A sufficient sample size should be obtained in order to ensure a precise estimate. The following sample size calculation provides an estimate of the sample size required to obtain a statistical significance of 5%. First, the expected prevalence of each outcome in the unexposed group should be estimated (or determined from previous literature), and a minimum expected relative risk (RR) of unexposed to exposed groups needs to be identified.

For example, if the prevalence of a given outcome is 10% in the unexposed, we would need at least 197 COVID-19 patients in each of the exposure strata to be able to detect a RR of at least 2 with 80% statistical power and a 95% confidence level (Table 4). Note that Table 4 provides the sample size for a univariable analysis; for a stratified or multivariable analysis, the sample size would be greater.

The pooled analyses should not prevent study teams from including a big enough sample size to obtain exact estimates for each separate study site.

➢ *Each study site to specify the minimum sample size calculation*

Table 4. Calculated sample sizes for different levels of expected relative risk (RR) by varying prevalence of outcome in the unexposed, assuming 95% confidence and 80% power

	P = 0.005	**P = 0.01**	**P = 0.02**	**P = 0.03**	**P = 0.04**	**P = 0.05**	**P = 0.1**

RR = 1.5	31,192	15,494	7646	5030	3722	2936	1366
RR = 2	4671	2316	1139	746	550	432	197
RR = 3	1551	766	373	243	177	138	59
RR = 4	858	422	204	131	95	73	29
RR = 5	576	282	135	86	61	47	17
RR = 10	204	97	44	26	17	12	1

Source: Sergeant, 2018.(5)

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

3.10.2 Data entry validation

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

- *Each study site to specify how data are validated*

3.10.3 Data management

Data entry and transfer

Web-based data collection methods or paper-based methods can be used. Data entry will include checks to minimise data entry errors. Double data entry is recommended unless electronic medical records are used.

Laboratory information will be reported to the study site coordinator using the reporting procedures existing in each study site for the severe COVID-19 RF study.

For the multi-centre pooled analysis, study sites may send an anonymised database to the coordinating team through the secure data transfer system EpiFiles. Each individual study database will be sent to the coordinating team study database using a secure protocol (see Annex 2: Dataflow for pooled database). All personal identifier information such as names, addresses, and medical registration codes will be deleted before data transmission to the coordinating team, where all individual data will be pooled. Study databases can be sent to Epiconcept in any format (e.g. Stata, CSV, EpiData, etc.).

Epiconcept provides the option of web-based data collection methods, if so desired by the countries. These methods can also be combined with paper-based methods.

If the Epiconcept web-based data collection methods are not used, data can be coded as outlined in Annex 1, but it is not required.

There are three methods for data collection for the risk factor study for severe COVID-19:

(1) As for I-MOVE influenza for most countries, data collection through your usual method with transfer of your electronic database to Epiconcept through the secure data transfer platform, EpiFiles

(2) As for I-MOVE influenza for some countries, data entry directly into the Epiconcept software Voozanoo, which Epiconcept will adapt to include the additional COVID-19 variables

(3) ISARIC (the International Severe Acute Respiratory and Emerging Infection Consortium) provides a modular electronic case report form (e-CRF) which some countries/sites may already be using. Some of the modules and the Rapid CRF are co-created with the World Health Organization (WHO). The e-CRF uses the REDCap platform, and sites may either download the application and use the CRF in electronic format (on laptop, tablet, mobile telephone) or can also print the e-CRF and collect the data on paper and then enter onto the electronic system. For COVID-19, ISARIC has developed a rapid data collection e-CRF comprising three modules, from which many variables in module 1 and a few from module 3 are the same as those to be collected by the I-MOVE-COVID-19 surveillance network. Epiconcept is working with ISARIC to develop the best solution for sites to include any additional items required. If your site choses the ISARIC e-CRF:

- Sites not already using ISARIC may register to use the full rapid COVID-19 e-CRF with modules 1–3
  - together with those sites already registered for ISARIC use, only the additional variables not covered by ISARIC would need to be completed for each patient (remaining variables will be automatically populated from the already completed modules 1–3)
  - note that for these sites there will be more data to be filled in, but only the data relevant for the I-MOVE-COVID-19 surveillance network will be used for this surveillance system

➢ *Study sites to specify procedures of data management and procedures to comply with the GDPR requirements*

➢ *Study sites to indicate which of the three data collection options they will use*

➢ *Study sites to provide a codebook that includes the variable names, variable descriptions, and the coding of variable values.*

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 will be carried out (e.g. date of discharge from hospital before date of onset of symptoms). Ideally, these checks will be included as warnings in the 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. A guide and/or an example Stata do-file for data cleaning will be provided if so desired.

- *Study sites to specify the data checking and cleaning process*

Data storage and data management for pooled analysis

The minimum dataset will be transmitted to Epiconcept where individual data will be pooled. Data will be stored in the EU data repository, as required by the European Commission (EC), and all data management procedures must comply with the General Data Protection Regulations (GDPR).

The coordinating team will conduct the pooled analysis. Data validation, cleaning and verification will be carried out at study level. A country (or study) identifier will be included in each record (e.g. ES for Spain, UK for the United Kingdom), a hospital code will be included (e.g. a unique number), and each record will be given a unique number. This number will also be included in the study team's database and will be used by the coordinating team and the study sites during pooling, so that records can be traced back whilst maintaining anonymity, if there should be any further queries. Tracing back will be performed by the study sites, not by the coordinating team. Study databases can be sent in any format.

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 will be reported to the study site in question.

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.

A study site-specific flowchart of exclusions and restrictions will be shared with each of the study sites. Variables will be recoded and new variables generated. The recoded data will be stored separately from the crude data and recoding will be documented.

Missing data

Any missing data will be described. If many data are missing and there is 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).

3.11 Indicators

3.11.1 Description of participating hospitals/sites

- Number and % of participating hospitals/sites
- Mean and total size of catchment population for participating hospitals, by country/site and pooled overall
- Total and mean number of suspected, probable and confirmed cases reported by hospital/site
  - overall and by ward/specialty

3.11.2 Patients by week/month and overall (by sex and age-group)

- Number and % of SARI patients (suspected COVID-19 patients)
  - admitted by hospital
  - tested for SARS-CoV-2 by hospital
  - in ICU/HDU, by hospital
- Number and % of recruited confirmed and probable COVID-19 patients
  - admitted by hospital
  - tested for SARS-CoV-2 by hospital
  - in ICU/HDU, by hospital
- Median length of stay in hospital (overall) and in ICU/HDU

3.11.3 Clinical characteristics of patients

- Number and % of confirmed and probable COVID-19 patients
  - overall, by hospital/site
  - by time, age group, sex
  - by risk and preventive factors
    - symptoms
    - complications
    - comorbid conditions
    - pre-hospital medications
    - in-hospital treatment and interventions
    - other (pregnancy, smoking)
    - vaccination (influenza, pneumococcal, BCG)
  - by severe outcome (severe pneumonia, ICU/HDU, ventilation, ECMO, death)

3.11.4 Identification of risk factors

- Risk of each severe outcome among exposed, with 95% confidence intervals
  - overall, by hospital
  - by time, age group, sex
- Severe outcomes
  - severe pneumonia
  - ICU/HDU

    - Ventilation
    - ECMO
    - death
- Exposures (risk/preventive factors)
    - symptoms
    - complications
    - comorbid conditions (none/one, >one; individual; grouped immunocompromised vs immunocompetent)
        - by treatment for chronic conditions, if available
    - pre-hospital medications
    - in-hospital treatment and interventions
        - respiratory support (oxygen, high-flow, (non-)invasive ventilation, prone position, nebuliser, ECMO)
        - other treatment (antivirals, etc.)
    - healthcare worker status
    - close contact setting
    - other (pregnancy, smoking)
    - vaccination (influenza, pneumococcal, BCG)

3.11.5 Laboratory indicators

- Number of cases by test performed? (if several tests are used?)
- Number of cases by genetic variant
- Number of severe cases by genetic variant

3.12 Data analysis

The analysis will be carried out first for each individual study site and shared with the site study team for validation. In a second step, the pooled analysis will be conducted (see Annex 5). Sample size permitting, RR will be calculated for each of the following outcomes:

- severe COVID-19 (whether confirmed or probable)
    - severe bilateral pneumonia with ground glass opacities
    - admission to ICU/HDU
    - ventilation
    - ECMO
- fatal COVID-19 (whether confirmed or probable)

The RRs for these outcomes will be estimated using all data with ICU admission as one outcome, but also restricted to patients admitted to ICU only (as England data will only comprise ICU patients) for the other outcomes, considering each of the following risk/protective factors as exposures:

- sex, age group, pregnancy, healthcare worker
- use of pre-symptomatic medications (including statins)
- absence, presence of at least one, presence of more than one chronic condition

    - specific chronic conditions (e.g. respiratory, cardiovascular diseases)
- BCG, influenza and pneumococcal vaccine
- time from onset to admission
- examinations/laboratory results on admission
- type of contact setting
- early clinical signs/symptoms
- medications and interventions in hospital
- respiratory viral co-infections
- complications

3.12.1 Validation of exposure

There are several exposures of interest and the validity of collected data for each one should therefore be checked carefully. If, for example, BCG vaccination status is reported by the patient only without further proof, information bias may occur. We recommend validating the vaccination status of all patients using an independent source (i.e. vaccination register, GPs, BCG scar).

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

For all other analysis details, please see Annex 5.

3.13 Consent

Each study will comply with national ethics committee requirements. Informed consent will be required from all participants or legal tutors. The national ethics committees will specify whether oral or written consent will be required. Specific consent procedures may be needed for unconscious patients and patients with deterioration of general condition or functional status, unable to sign the consent (e.g. oral witnessed consent, consent by the next of kin, etc). A copy of the ethical approvals should be sent to the coordinating centre, with a copy of the template used to obtain informed consent.

- *Each study site to describe the procedures to comply to the national ethics committee requirements and the type of informed consent needed as well as whether consent can be obtained for a legal tutor*
- *Each study site to send a copy of the ethical approval to the coordinating centre*

3.14 Dissemination of results

The enrolment of COVID-19 cases will be regularly updated by each study coordinator on a website developed for the multicentre study. Initial RR estimates will be disseminated as soon as possible; monthly estimates will

follow thereafter and a final overall RR when the pandemic is over. (Note that this may be revised depending on how the pandemic progresses.)

3.14.1 Publications, scientific communication

Results of the individual studies should only be published in open-source journals (this is a requirement of the European Commission's H2020 funding received for this surveillance project). Each study coordinator will decide which scientific conferences will be attended in order to present the results. An article presenting the results of the pooled analysis and estimates for the EU/EEA will be submitted to an open-source, peer-reviewed journal.

The list of authors will respect the recommendations of authorship stated by the International Committee of Medical Journal Editors: http://www.icmje.org/ethical_1author.html. The actual authorship for the pooled article will be discussed and agreed with the study teams at the beginning of the study.

3.15 Training

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

- *Each study site to describe the trainings to be organised*

4 Logistical aspects

4.1 Study leader

In each study site, a principal investigator will coordinate the study at the country level and act as focal point for the European study. Epiconcept is in charge of the pooled analysis.

4.2 Human resources

In each hospital/hospital network, an investigator will be in charge of monitoring data collection at the hospital level. Study investigators at the hospital will collect information from the cohort. The specific human resources needed in each country are detailed in the study annexes. Epiconcept ensures the overall coordination of the various studies.

4.3 Supervision

If feasible in the pandemic context, site visits and joint workshops (remote if required) will be organised by the coordinating team/study sites in order to carry out an appraisal of the ongoing studies in the various countries involved. The appraisal team will be composed of two persons from the various project partners.

4.4 Respiratory specimen collection

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

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

4.5 Standard operating procedures

Standard operating procedures should be used by investigators during all the steps of the study for identification of study subjects, data collection, laboratory methods, data entry, monitoring, etc. Epiconcept has prepared a data entry SOP for use with Epiconcept's online Voozanoo 4 questionnaire, for sites using this platform.

- *Each study site to develop (or adapt pre-existing) study SOP to be used by the study team*

4.6 Reports

Each study site will write a report at the end of the study and submit it to the study coordination team. Epiconcept will write a final report presenting the results of the pooled analysis.

5 Limitations

With any multi-centre study, there is always the potential for heterogeneity among sites. In addition, during a pandemic with such high caseloads for hospitals, there may be difficulties in collecting all data, and not all included cases will have laboratory confirmation. There is also the possibility that very severely ill patients (e.g. those who are extremely frail and/or in nursing homes) may not be admitted to hospital at all, and would be missed by the study; further, some elderly or frail patients may not receive intubation or resuscitation. Another limitation is that patients receiving hydroxychloroquine prior to admission may not test positive, and may have milder disease but they would not be included (so we would not be able to test this). A final limitation is the fact that several collected variables may be RFs as well as outcome or may be on the causal pathway and this may affect the analyses.

5.1 Representativeness of subjects included in the study

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

5.2 Potential biases

5.2.1 Variation in ICU admission criteria

A simple example is that when under high pressure (e.g. high volume of patients to be admitted), it is possible that younger adults will be preferentially admitted to ICU. Therefore, during high pressure periods, age will not be identified as a risk factor for severity. This means that identification of risk factors which are part of the criteria to define some of the severity outcomes could vary over time. As a result, the RR for these outcomes may increase or decrease over time, but this may be due to external factors not related directly to the outcome or the exposure.

5.2.2 Unmeasured confounding

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

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

6 References

1. European Centre for Disease Prevention and Control. Rapid risk assessment: Coronavirus disease 2019 (COVID-19) pandemic: increased transmission in the EU/EEA and the UK – seventh update [Internet]. European Centre for Disease Prevention and Control. 2020 [cited 2020 Mar 27]. Available from: https://www.ecdc.europa.eu/en/publications-data/rapid-risk-assessment-coronavirus-disease-2019-covid-19-pandemic
2. Valenciano M, Ciancio B, I-MOVE study team. I-MOVE: a European network to measure the effectiveness of influenza vaccines. Euro Surveill Bull Eur Sur Mal Transm Eur Commun Dis Bull. 2012 Sep 27;17(39).

3. World Health Organization. Laboratory biosafety guidance related to coronavirus disease 2019 (COVID-19): Interim guidance [Internet]. WHO; 2020. Available from: https://apps.who.int/iris/bitstream/handle/10665/331138/WHO-WPE-GIH-2020.1-eng.pdf
4. World Health Organization. Laboratory testing for 2019 novel coronavirus (2019-nCoV) in suspected human cases: Interim guidance [Internet]. 2020 [cited 2020 Apr 6]. Available from: https://www.who.int/publications-detail/laboratory-testing-for-2019-novel-coronavirus-in-suspected-human-cases-20200117
5. Sergeant ESG. Epitools Epidemiological Calculators. Ausvet. 2018.

7 Bibliography

1. Blettner, M, Sauerbrei, W, Schlehofer, B, Scheuchenpflug T, Friedenreich C. Traditional reviews, metaanalyses and pooled analyses in epidemiology. Int J Epidemiol 28, 1–9(1999).
2. Kissling E, Valenciano M. I-MOVE case–control studies team. Early estimates of seasonal influenza vaccine effectiveness in Europe, 2010/11: I-MOVE, a multicentre case–control study. Euro Surveill. 2011;16(11):pii=19818
3. Kissling E, Valenciano M, Cohen JM, Oroszi B, Barret AS, Rizzo C, et al. I-MOVE Multi-Centre Case Control Study 2010–11: Overall and stratified estimates of influenza vaccine effectiveness in Europe. PLoS ONE 2011;6(11):e27622.
4. Lemeshow S, Hosmer D, Klar J. Sample size requirements for studies estimating odds ratios and relative risks. Stat Med 1988 ;(7) :759–64.
5. National Research Council. Combining information – statistical issues and opportunities for research. Washington, DC: National Research Council; 1992. doi: https://doi.org/10.17226/20865.
6. Rodrigues L, Kirkwood BR. Case–control designs in the study of common diseases: updates on the demise of the rare disease assumption and the choice of sampling scheme for controls. Int J Epidemiol 1990 Mar;19(1):205–13.
7. Rothman, K, Greenland, S, Lash, T. Modern Epidemiology (ed.). Philadephia: Lippincott–Williams–Wilkins, 2008.
8. Smith-Warner, S.A, Spiegelman, D, Ritz, J, Albanes, D, Beeson, W.L, Bernstein, L. et al. Methods for pooling results of epidemiologic studies: the Pooling Project of Prospective Studies of Diet and Cancer. Am J Epidemiol 2006 ;163 :1053–64.
9. Stukel TA, Demidenko E, Dykes J, Karagas MR. Two-stage methods for the analysis of pooled data. Stat Med 2001 Jul 30;20(14):2115–30.
10. Valenciano M, Kissling E, Cohen JM, Oroszi B, Barret AS, Rizzo C, et al. Estimates of pandemic influenza vaccine effectiveness in Europe, 2009–2010: Results of Influenza Monitoring Vaccine Effectiveness in Europe (I-MOVE) Multicentre Case–Control Study. Plos Med 2011; 8(1):e10000388. doi: 10.1371/ journal.pmed.1000388.

8 Annexes

8.1 Annex 1: List of variables, definitions and coding; I-MOVE-COVID-19 hospital-based severity risk factor study minimum dataset

Individual data (to be adapted to TESSy format)

- *study sites to list all the variables collected and their coding*
- *study sites to indicate all modifications in the variables collected compared to variables below*

Existing partners: in yellow highlight are new variables not collected for I-MOVE influenza (or coding changes)

	Variable	Type	Values and coding	Definition
Study identifiers	idcountry	Numeric (categorical)	Coded according to international country codes	Identifier uniquely identifying the country
	id	Numeric	Unique integer	Unique number for each patient
	hospitalcode	Numeric	Unique integer	Unique number for each hospital
	consent	Numeric	0 = No	Agreement of patient to participate (where appropriate, i.e. in countries/sites where consent required)
			1 = Yes	
			2 = Not required	
			8 = Do not know	
	consent_sp	Text		Reason provided for non-participation
Hospital/ward information (see new vars added at end of this table)	admitdate	Date	dd/mm/yyyy	Date of hospital admission
	multiple_hosp	Numeric (categorical)	0 = No	Patient had more than one hospital admission for SARI/suspected COVID-19 as part of this illness episode
			1 = Yes	
			8 = Do not know	
	hospitalward	Numeric (categorical)	0 = Special COVID-19 ward	First ward of referral
			1 = Lung, pulmon/respir.	
			2 = Internal medicine	
			3 = Infectious diseases	
			4 = Emergency or A&E	
			5 = Cardiology	
			6 = Geriatric	
			7 = ICU or HDU	
			9 = Other	
			8 = Do not know	
	hospitalward_oth	Text		Specify other ward
	dischargedate	Date	dd/mm/yyyy	Date of hospital discharge
	Variable	**Type**	**Values and coding**	**Definition**

Hospital/ward information (continued)	los_hosp	Numeric (integer)		Length of hospital stay (for multiple admissions; else will be calculated from admission/discharge dates)
	icu	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Admission to intensive care unitanalysis (ICU) or high-dependency unit (HDU)
	icuadmitdate	Date	dd/mm/yyyy	Date first admitted ICU/HDU
	icudisdate	Date	dd/mm/yyyy	Date last discharged from ICU/HDU
	los_icu	Numeric (integer)		Length of stay in ICU/HDU (if no dates for ICU/HDU admission/discharge, or if multiple admissions for ICU/HDU)
Patient characteristics	sex	Numeric	0 = female 1 = male 3 = other 8 = do not know	Sex of patient
	lastname	Text		Patient's surname
	firstname	Text		Patient's first name
	dob	Date	dd/mm/yyyy	Date of birth (only if no age; once age calculated from dob this will be dropped)
	age_y	Numeric		Age of patient (if unable to provide dob) in years for those aged 2 years and older
	age_m	Numeric		Age of patient (if unable to provide dob in months for those aged <2 years
	residence	Numeric	0 = at home, not dependent on home support/care 1 = at home, but dependent on home support/care 2 = institutionalised 3 = Do not know	Patient residence at time of SARI onset. Whether patient was living at home or was institutionalised, or had pre-hospital dependence on home support/care
	postcode	Text		Postcode of residence (where possible)
	smoking	Numeric (categorical)	0 = Never 1 = Former 2 = Current 8 = Do not know	Never, former (stopped smoking at least 1 year before inclusion in the study), current smoker
	pregnant	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Whether patient is pregnant
	Variable	**Type**	**Values and coding**	**Definition**
Patient characteristics	trimester	Numeric (categorical)	1 = Trimester 1 2 = Trimester 2	Trimester of pregnancy

(continued)			3 = Trimester 3	
			8 = Unknown trimester	
	postpartum	Numeric (categorical)	0 = No	Whether patient is within the first 6 weeks post partum
			1 = Yes	
			8 = Do not know	
	hcw	Numeric (categorical)	0 = No	Whether the patient is a healthcare worker
			1 = Yes	
			8 = Do not know	
Risk factors (close contact setting)	closecont	Numeric (categorical)	1 = Family setting	Close contact setting with a person who is a probable or confirmed case in the 14 days prior to symptom onset
			2 = Health care setting	
			3 = Workplace setting	
			4 = Long-term care facility	
			5 = Prison	
			6 = Other	
			8 = Do not know	
	closecont_sp	Text		Specify other close contact setting
Underlying chronic conditions	anaemia	Numeric (categorical)	0 = No	Anaemia/chronic haematologic disease
			1 = Yes	
			8 = Do not know	
	asplenia	Numeric (categorical)	0 = No	Asplenia (absence of/damage to spleen)
			1 = Yes	
			8 = Do not know	
	asthma	Numeric (categorical)	0 = No	Asthma
			1 = Yes	
			8 = Do not know	
	cancer	Numeric (categorical)	0 = No	Cancer (any)
			1 = Yes	
			8 = Do not know	
	hypert	Numeric (categorical)	0 = No	Hypertension
			1 = Yes	
			8 = Do not know	
	Variable	**Type**	**Values and coding**	**Definition**
Underlying chronic conditions (continued)	dement	Numeric (categorical)	0 = No	Dementia
			1 = Yes	
			8 = Do not know	
	diabetes	Numeric	0 = No	Diabetes

		(categorical)	1 = Yes	
			8 = Do not know	
	heartdis	Numeric (categorical)	0 = No	Heart / cardiac disease (excluding hypertension)
			1 = Yes	
			8 = Do not know	
	hypert	Numeric (categorical)	0 = No	Hypertension
			1 = Yes	
			8 = Do not know	
	immuno	Numeric (categorical)	0 = No	HIV (including other immunodeficiency, organ transplantation)
			1 = Yes	
			8 = Do not know	
	liverdis	Numeric (categorical)	0 = No	Chronic liver disease (excluding cancer)
			1 = Yes	
			8 = Do not know	
	lungdis	Numeric (categorical)	0 = No	Lung disease (excluding asthma)
			1 = Yes	
			8 = Do not know	
	lungdis_sp	Numeric (categorical)	0 = No	If lung disease: specify if use of non-invasive ventilation / oxygen therapy or chronic obstructive pulmonary disease (COPD)
			1 = non-invasive ventilation/ oxygen therapy	
			2 = COPD	
			8 = Do not know	
	neuromusc	Numeric (categorical)	0 = No	Neuromuscular disorder
			1 = Yes	
			8 = Do not know	
	height	Numeric (integer)		Height of patient in metres
	weight	Numeric (integer)		Weight of patient in kg
	bmi	Numeric (1 d.p.)		BMI of patient (only if available in place of missing weight/height)

	Variable	Type	Values and coding	Definition
Underlying chronic conditions (continued)	obese	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Obesity (only if height, weight and BMI not collected; can be calculated)
	rendis	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Renal disease (excluding cancer and acute renal failure)
	rheumat	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Rheumatologic disease
	stroke	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Stroke
	tuberc	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Tuberculosis
	prevhosp	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Prior admission to hospital (at least once in previous 12 months, for co-morbid conditions)
Pre-admission contact	healthcare_contact	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Any contact with GP or other primary healthcare services prior to hospital admission
Case definitions (SARI signs/ symptoms **pre**-admission)	feverish_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Sub-febrility (37–38°C)
	fever_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	History of fever > 38°C
	chills_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	"Chills", or shivering
	cough_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Cough
	Variable	**Type**	**Values and coding**	**Definition**
Case definitions	sob_pre		0 = No	Shortness of breath

(SARI signs/ symptoms **pre**-admission)		Numeric (categorical)	1 = Yes 8 = Do not know	
	tach_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Tachypnoea or signs of low oxygen saturation
	suddenonset_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Sudden onset
	headache_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Headache
	sorethroat_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Sore throat
	coryza_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Coryza
	malaise_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Malaise
	general_deter_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Deterioration of general condition (asthenia or loss of weight or anorexia)
	confusion_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Confusion
	dizzy_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Dizziness
	myalgia_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Myalgia
	Variable	**Type**	**Values and coding**	**Definition**
Case definitions (SARI signs/ symptoms **pre**-admission	chest_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Chest pains

continued)	palp_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Heart palpitations
	diarr_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Diarrhoea
	nausea_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Nausea
	vomit_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Vomiting
	abdopain_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Abdominal pain
	ageusia_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Loss of sense of taste
	anosmia_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Loss of sense of smell
	dermato_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Rash or other dermatological manifestation of COVID-19
	onsetdate_pre	Date	dd/mm/yyyy	Date of onset of symptoms
Risk factors (Pre-symptomatic treatment: medication)	statin_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Patient was on statins since or from 01 January 2020

	Variable	Type	Values and coding	Definition
Risk factors (Pre-symptomatic treatment/intervention: medication continued)	ace_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	ACE inhibitor (angiotensin converting enzyme inhibitors)
	arb_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	ARB (angiotensin II receptor blockers)
	nsaid_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	NSAID (non-steroidal anti-inflammatory drugs)
	metform_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Metformin
	steroids_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Steroids
	corticost_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Corticosteroids
	dmards_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Biological disease-modifying anti-rheumatic drugs (DMARDs) e.g. rituximab, tocilizumab, etc.
	chemo_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Chemotherapy (within 6 months or currently) for cancer
	gliclaz_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Gliclazides (for diabetes or heart failure)
	psychotrop_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Psychotropic drugs (including benzodiazepine, etc.)

	Variable	Type	Values and coding	Definition
Pre-symptomatic treatment cont'd	antivir_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Antivirals
	chloroq_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Chloroquine
	hydroxychloroq_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Hydroxychloroquine
	other1_pre_sp	Text		Other pre-symptomatic medication #1
	other2_pre_sp	Text		Other pre-symptomatic medication #2
	other3_pre_sp	Text		Other pre-symptomatic medication #3
Risk factors (Pre-symptomatic treatment/intervention: vaccination)	flu_vacc	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Received current seasonal influenza vaccination
	flu_vaccdate	Date	dd/mm/yyyy	Date of last influenza vaccination
	ppv_vacc	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Received PPV23 vaccination
	ppv_vaccdate	Date	dd/mm/yyyy	Date of last PPV23 vaccination
	pcv_vacc	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Received PCV7/10 or 13 vaccination
	pcv_vaccdate	Date	dd/mm/yyyy	Date of last PCV7/10 or 13 vaccination
	bcg_vacc	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Received BCG vaccination
	bcg_vaccyear	Numeric	yyyy	Year of BCG vaccination
	bcg_scar	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Presence of BCG scar (if known, and if BCG unknown)

	Variable	Type	Values and coding	Definition
Case definitions (signs and symptoms at admission)	same_symptoms	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Whether signs/symptoms at admission are exactly the same as those pre-admission (if YES then do **not** need to complete same list of variables now shown again below)
	feverish	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Sub-febrility (37–38°C)
	fever	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	History of fever > 38°C
	chills	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	"Chills", or shivering
	cough	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Cough
	sob	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Shortness of breath
	tach	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Tachypnoea or signs of low oxygen saturation
	suddenonset	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Sudden onset
	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
	coryza	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Coryza
	Variable	**Type**	**Values and coding**	**Definition**

Case definitions (signs and symptoms at admission)	malaise	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Malaise
	general_deter	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Deterioration of general condition (asthenia or loss of weight or anorexia)
	confusion	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Confusion
	dizzy	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Dizziness
	myalgia	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Myalgia
	chest	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Chest pains
	palp	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Heart palpitations
	diarr	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Diarrhoea
	nausea	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Nausea
	vomit	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Vomiting
	abdopain	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Abdominal pain

	Variable	Type	Values and coding	Definition
Case definitions (signs and symptoms at admission)	ageusia	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Loss of sense of taste
	anosmia	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Loss of sense of smell
	dermato	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Rash or other dermatological manifestation of COVID-19
	same_date	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Whether onset date of first sign/symptom is the same as pre-admission (if YES then do **not** need to complete date again)
	onsetdate	Date	dd/mm/yyyy	Date onset first symptom (if different)
Additional signs at admission (RF study only) See changes/ additions at end of table	frailty	VAR CHANGED: SEE END OF THIS TABLE	To be updated with coding depending on score used	Clinical frailty score at admission (where possible)
	gcs	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Whether GCS score was measured at admission
	gcs_motor	Numeric (integer)		Glasgow Coma Scale (GCS) motor score on admission (range 0–6)
	gcs_verbal	Numeric (integer)		GCS verbal score (0–5) on admission
	gcs_visual	Numeric (integer)		GCS eye-opening score on admission (0–4)
	gcs_total	Numeric (integer)		GCS total score on admission (0–15). Can be calculated from M, V, E scores (provide either M, V, E **or** this total)
	avpu	Numeric (categorical)	0 = AVPU not done 1 = Alert 2 = Verbal 3 = Pain 4 = Unresponsive	AVPU score (Alert, Verbal, Pain or Unresponsive) on admission
In-hospital examinations/ lab tests at admission (RF study only)	bloodurea	Numeric (integer)		Blood urea nitrogen (mmol/L)
	resprate	Numeric (integer)		Respiratory rate (breaths/min)
	bpsys	Numeric (integer)		Systolic blood pressure (mmHg)
	bpdia	Numeric (integer)		Diastolic blood pressure (mmHg)

	Variable	Type	Values and coding	Definition
In-hospital examinations/ lab tests at admission (RF study only) continued	abo	Numeric (categorical)	0 = O 1 = A 2 = B 3 = Rh + 4 = Rh -	ABO blood grouping (Note: for AB, select both A and B)
	plat	Numeric (integer)		Platelet count per µL
	lymphoc	Numeric (categorical)	0 = < 500 1 = 501-800 2 = 801 and over	Lymphocyte count per µL
	neutro	Numeric (integer)		Neutrophil count per µL
	eosin	Numeric (integer)		Eosinophil count per µL
	lft_any	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Whether any liver function tests have been performed
	lft_alb	Numeric (decimal)		Serum albumin g/dL
	lft_bili	Numeric (decimal)		Serum total bilirubin mg/dL
	lft_bili_conj	Numeric (decimal)		Serum conjugated bilirubin mg/dL
	lft_ggt	Numeric (integer)		GGT (gamma glutamyltransferase) IU/L
	lft_totprot	Numeric (integer)		Serum total protein g/L
	alp	Numeric (decimal)		Alkaline phosphatase (ALP) IU/L
	ast	Numeric (decimal)		Aspartate aminotransferase (AST) units/L
	alt	Numeric (decimal)		Alanine aminotransaminase (ALT) U/L
	lft_prothromb	Numeric (decimal)		Prothrombin time
	ldh	Numeric (decimal)		Lactic acid dehydrogenase (LDH) U/L
	ferritin	Numeric (decimal)		Ferritin (ng/mL)
	dimer	Numeric (decimal)		D-Dimer ng/mL
	fibrin	Numeric (decimal)		Fibrinogen mg/dL
	crp	Numeric (decimal)		C-reactive protein mg/L
	cpk	Numeric (decimal)		Creatine phosphokinase U/L
	trop	Numeric (decimal)		Troponin-I ng/mL
	hba1c	Numeric (decimal)		Single glycated haemoglobin (HbA1c) %
	trigly	Numeric (decimal)		Triglycerides mg/dL
	hdl	Numeric (decimal)		Cholesterol – high-density lipoprotein mg/dL
	Variable	**Type**	**Values and coding**	**Definition**
In-hospital	ldl	Numeric (decimal)		Cholesterol – low-density lipoprotein

examinations/ lab tests at admission (RF study only) continued				mg/dL
	nt-probnp	Numeric (decimal)		N-terminal-prohormone B-type natriuretic peptide (NT-proBNP) pg/mL
In-hospital examinations/ lab tests at admission	ct_us_ecg	Numeric (categorical)	0 = No CT or u/s scan done	Indicate whether patient had CT/ultrasound/ECG or none of these (note: several selections may be made, e.g. if patient had CT and u/sound)
			1 = CT scan	
			2 = Ultrasound	
			3 = ECG	
			4 = Other	
			8 = Do not know	
	ct_res	Numeric (categorical)	0 = Normal CT	CT results. Ground-glass opacification defined as hazy increased lung attenuation with preservation of bronchial and vascular margins. Consolidation defined as opacification with obscuration of margins of vessels and airway walls
			1 = Bilateral lung involvement	
			2 = Peripheral lung distribution	
			3 = Ground-glass opacities	
			4 = Consolidation	
			5= Other	
	ct_res_sp	Text		Specify other significant finding on CT/ultrasound
	cxr	Numeric (categorical)	0 = Not done	Chest X-ray findings. Ground-glass opacification defined as hazy increased lung attenuation with preservation of bronchial and vascular margins. Consolidation defined as opacification with obscuration of margins of vessels and airway walls
			1 = No findings	
			2 = Infiltrates	
			3 = Ground-glass opacities	
			4 = Consolidation	
			8 = Unknown	
	cxroth_sp	Text		Specify other CXR result, if relevant
	ecg_qt	Numeric (categorical)	0 = No	Did ECG finding show presence of long QTc?
			1 = Yes	
			8 = Do not know	
	examoth_sp	Text		List any other in-hospital examinations and their most significant findings
	oxsat	Numeric (integer)		Patient's oxygen saturation on admission to hospital (on air) %

	Variable	Type	Values and coding	Definition
Case/severity definitions (COVID-19 case)	swabdate	Date	dd/mm/yyyy	Respiratory specimen collection date
	lab_covtest	Numeric (categorical)	0 = No	Whether patient tested for SARS-CoV-2
			1 = Yes	
			8 = Do not know	
	lab_covtesttype	Numeric (categorical)	1 = RT-PCR	Type of lab test used
			2 = Serology	
			3 = Rapid test	
			4 = Other	
			8 = Do not know	
	lab_covtesttype_sp	Text		Specify other type of lab test
	lab_covid	Numeric (categorical)	0 = Negative	Laboratory result: virus type SARS-CoV-2
			1 = Positive	
			8 = Do not know	
	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
	covid	Numeric	0 = No	Whether patient is a case of COVID-19 or not (this classification will be done by re-coding after data collection)
			1 = Confirmed	
			2 = Probable	
			3 = Other coronavirus	
			4 = Suspected	
			8 = Do not know	
Case/severity definitions (severity indicators)	outcome	Numeric (categorical)	1 = died	Indicate the outcome of the patient known at the time of data collection (note: this may be updated later)
			2 = discharged from hospital	
			4 = still on treatment	
			8 = unknown outcome	
	deathdate	Date	dd/mm/yyyy	Date of death
	deathcause	Numeric (categorical)	1 = died from COVID-19	Cause of death
			2 = died other cause	
			8 = died unknown cause	

	Variable	Type	Values and coding	Definition
Case/severity definitions (severity indicators continued)	Vent	Numeric (categorical)	0 = No 1 = ECMO 2 = Oxygen (high-flow) 3 = Ventilator (non-invasive) 4 = Ventilator (invasive) 5 = Other 8 = Do not know	Patient's level of mechanical ventilation. Note that option 1 is for respiratory support level ECMO, option 2 includes any high-flow (6L/min or higher, including OptiFlow), and option 3 includes any non-invasive, positive pressure ventilator.
	vent_sp	Text		Specify other mechanical ventilation
	vent_type		1 = PEEP 2 = BiPAP 3 = CPAP 4 = Other invasive ventilation	Type of invasive ventilation: positive end-expiratory pressure (PEEP), bilevel positive airway pressure (BiPAP), continuous positive airway pressure (CPAP) or other
	venttype_sp	Text		Specify other invasive ventilation type
	bilat_pneu	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Presence of bilateral pneumonia with ground-glass opacities (this classification will be done by re-coding after data collection)
In-hospital medications/ interventions	ox_nasal	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Nasal oxygen (not high-flow)
	prone	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Whether patient was placed in prone position for ventilation
	nebu	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Nebuliser treatment
Risk factors (in-hospital medications/ interventions)	antivir	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Antivirals in hospital ; e.g. remdesivir, ritonavir, lopinavir, favipiravir, umifenovir
	antivir_type	Text		Specify brand names of key antivirals with doses if possible
	hydroxychl	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Hydroxychloroquine

	Variable	Type	Values and coding	Definition
Risk factors (in-hospital medications/ interventions) continued	chlor	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Chloroquine
	cortico	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Corticosteroids
	corticocst_type		1 = Oral 2 = Intravenous 3 = Other 8 = Do not know	Specify administration of corticosteroids
	il6	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Monoclonal antibodies/IL-6 blockers (e.g. tocilizumab)
	il6_type	Text		Specify brand names of IL-6 blockers and dose if possible
	antibiot	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Antibiotics (e.g. azithromycin)
	antibiot_type	Text		Specify brand names of key antibiotics and dose if possible
	antivir_med	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Antivirals e.g. remdesivir, ritonavir, lopinavir, favipiravir, umifenovir
	antivir_type	Text		Specify brand names of key antivirals with doses if possible
	sep_resus	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Sepsis fluid resuscitation
	trialdrugs	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Whether trial drugs were administered
	study_convpl	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Use of study/trial drugs: convalescent plasma
	Variable	**Type**	**Values and coding**	**Definition**
Risk factors (in-	study_gm_csf		0 = No	Use of study/trial drugs:

hospital medications/ interventions) continued		Numeric (categorical)	1 = Yes	GM-CSF
			8 = Do not know	
	study_oth	Numeric (categorical)	0 = No	Other new/study/trial drugs
			1 = Yes	
			8 = Do not know	
	study_oth_sp	Text		Specify any other study or trial medications
	meds_oth1	Text		Specify other important medications #1
	meds_oth2	Text		Specify other important medications #2
	meds_oth3	Text		Specify other important medications #3
Risk factors (other respiratory virses)	lab_fluany	Numeric (categorical)	0 = Negative	Laboratory result: any influenza virus type
			1 = Positive	
			2 = Not done	
			8 = Do not know	
	lab_mers	Numeric (categorical)	0 = Negative	Laboratory result: virus type MERS-CoV
			1 = Positive	
			2 = Not done	
			8 = Do not know	
	lab_othcov	Numeric (categorical)	0 = Negative	Laboratory result: virus type other coronavirus
			1 = Positive	
			2 = Not done	
	resp_virus	Numeric (categorical)	0 = None	Which other non-influenza, non-coronavirus patient tests positive for
			1 = RSV	
			2 = Metapneumovirus	
			3 = Other respiratory infection	
			8 = Do not know	
Complications	complic	Numeric (categorical)	0 = None	Complications which the patient may have experienced at any time; note that option 11 refers to any dermatological manifestations of COVID-19
			1 = ARDS (acute respiratory distress syndrome)	
			2 = Bronchiolitis	
			3 = Encephalitis	
			4 = Myocarditis	
			5 = Pneumonia (sec bacterial)	
			6 = Other sec bac infection	
			7 = Sepsis	

			8 = Acute renal injury	
			9 = Heart failure	
			10 = Multi-organ failure	
			11 = Dermatological	
			12 = ICU-relayed myopathy	
			13 = Guillain Barré Syndrome	
			14 = Other (specify)	
	complic_sp	Text		Specify other complication

New variables added in v5:

	Variable	Type	Values and coding	Definition
Hospital/ward information	multiple_hosp	Numeric (categorical)	0 = No	If patient had >1 hospital admission for SARI/suspected COVID as part of this episode (> 14 days from onset)
			1 = Yes	
			8 = Do not know	
	multiple_hosp	Numeric (categorical)	1 = Re-admission 1	Number of re-admissions (where known)
			2 = Re-admission 2	
			3 = Re-admission 3	
			4 = Re-admission 4	
			5 = Re-admission 5	
			6 = Re-admission 6	
			7 = Unknown re-admit no.	
			1 = Yes	
			8 = Do not know	
Symptoms pre-admission	conjunct_pre	Numeric (categorical)	0 = No	Conjunctivitis
			1 = Yes	
			8 = Do not know	
Symptoms at admission	conjunct	Numeric (categorical)	0 = No	Conjunctivitis
			1 = Yes	
			8 = Do not know	
Laboratory results	pcr2	Numeric (categorical)	0 = No	Whether a second PCR was done (if first PCR was negative)
			1 = Yes	
			8 = Do not know	
	lab_covidpcr2	Numeric (categorical)	0 = Negative	Second PCR result for virus type SARS-COV-2
			1 = Positive	
			8 = Do not know	

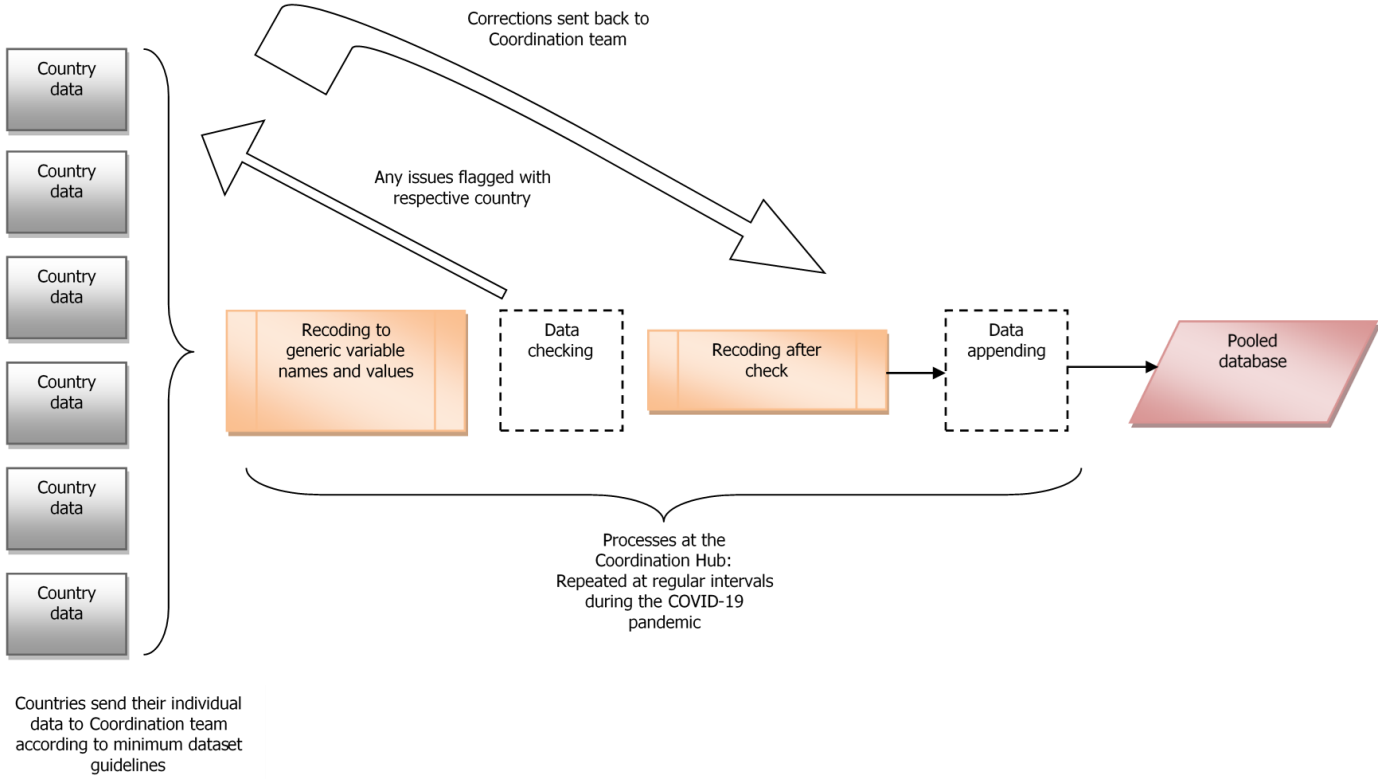
	Variable	Type	Values and coding	Definition
Pre-symptomatic medication	il6_pre	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Whether patient had monoclonal antibodies or IL-6 blockers (e.g. tocilizumab) pre-symptoms
	il6_pre_type	Text		Specify type of monoclonal antibody taken pre-symptoms
	corticost_pre_type	Numeric (categorical)	1 = Inhaled 2 = Systemic 3 = Other 8 = Do not know	Specify type of corticosteroids taken pre-symptoms
Frailty assessments	frailty_any	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Whether any type of clinical frailty score was used at admission to assess patient
	frailty_type	Numeric (categorical)	1 = Barthel Index 2 = Clinical Frailty Score (CFS) 3 = Other 8 = Do not know	Indicate which type of clinical frailty score was used
	frailty_sp	Text		Specify which other clinical frailty score was used
	frailty_barthel	Text		Barthel score at admission
	frailty_cfs	Text		CFS score at admission
Risk factors (exam/lab results on admission)	urea	Numeric (integer)		Urea (mmol/L)
	heartrate	Numeric (integer)		Heart rate (beats/min)
	bnp	Numeric (decimal)		B-type natriuretic peptide (BNP) pg/mL
Risk factors (most pathologic exam/labs results)	bloodurea_mp	Numeric (integer)		Blood urea nitrogen (mmol/L)
	urea_mp	Numeric (integer)		Urea (mmol/L)
	resprate_mp	Numeric (integer)		Respiratory rate (breaths/min)
	heartrate_mp	Numeric (integer)		Heart rate (beats/min)
	bpsys_mp	Numeric (integer)		Systolic blood pressure (mmHg)
	bpdia_mp	Numeric (integer)		Diastolic blood pressure (mmHg)
	plat_mp	Numeric (integer)		Platelet count per µL
	lymphoc_mp	Numeric (integer)	0 = <500 1 = 501-800 2 = 801+	Lymphocyte count per µL
	neutron_mp	Numeric (integer)		Neutrophil count per µL
	eosin_mp	Numeric (integer)		Eosinophil count per µL
	lft_any_mp	Numeric (categorical)	0 = No 1 = Yes 8 = Do not know	Whether any liver function test (LFT) was performed
	lft_alb_mp_mp	Numeric (decimal)		Serum albumin g/dL
	lft_bili_mp	Numeric (decimal)		Serum total bilirubin mg/dL
	lft_bili_conj_mp	Numeric (decimal)		Serum conjugated bilirubin mg/dL
	lft_ggt_mp	Numeric (integer)		GGT (gamma glutamyltransferase) IU/L
	Variable	**Type**	**Values and coding**	**Definition**
Risk factors (most pathologic	lft_totprot_mp	Numeric (integer)		Serum total protein g/L
	alp_mp	Numeric (integer)		Alkaline phosphatase (ALP)
	ast_mp	Numeric (integer)		Aspartate aminotransferase (AST)

exam/labs results continued)	alt_mp	Numeric (integer)		Alanine transaminase (ALT)
	lft_prothromb_mp	Numeric (integer)		Prothrombin time
	ldh_mp	Numeric (integer)		Lactic acid dehydrogenase (LDH)
	ferritin_mp	Numeric (integer)		Ferritin
	dimer_mp	Numeric (integer)		D-Dimer
	fibrin_mp	Numeric (integer)		Fibrinogen
	crp_mp	Numeric (integer)		C-reactive protein
	cpk_mp	Numeric (integer)		Creatine phosphokinase
	trop_mp	Numeric (integer)		Troponin-I
	hba1c_mp	Numeric (integer)		Single glycated haemoglobin (HbA1c)
	trigly_mp	Numeric (integer)		Triglycerides
	ldl_mp	Numeric (integer)		Cholesterol – high-density lipoprotein
	hdl_mp	Numeric (integer)		Cholesterol – low-density lipoprotein
	nt_probnp_mp	Numeric (integer)		N-terminal-prohormone B-type natriuretic peptide (NT-proBNP)
	bnp_mp	Numeric (decimal)		B-type natriuretic peptide (BNP) pg/mL

When a vaccine(s) become(s) available, these additional variables will be required:

	Variable	Type	Values and coding	Definition
Pandemic COVID-19 vaccination	panvaccany	Numeric (categorical)	0 = No	Received pandemic COVID-19 vaccination
			1 = Yes	
			8 = Do not know	
	panvaccdose	Numeric	0, 1, 2	Number of doses received
	panvaccdate1	Date	dd/mm/yyyy	Vaccination date first dose
	panvaccdate2	Date	dd/mm/yyyy	Vaccination date second dose
	panvacctype	Text		Type of vaccine (product name)

8.2 Annex 2. Data flow for pooled dataset



8.3 Annex 3: Genetic and antigenic analysis data (example)

COVID-19	Country	GISAID EpiCoV number	ID number I-MOVE-COVID-19 surveillance	CT value
Row for vaccine strain1				
Row for vaccine strain2				

8.4 Annex 4: Study-specific annexes

Study specifications for each country should be summarised in this annex. Each study site annex should include:

- description of the hospitals participating in the study (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
- 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, culture, strain characterisation; methods for specimen collection, storage, transport; selection procedures for strain characterisation)
- consent, ethical procedures (oral/written consent; submission to ethics committee)
- human resources needed
- provisions to train hospitals.

8.5 Annex 5: Detailed analysis plan

8.5.1. Individual study site analysis

Each individual study site can analyse their data. The coordinating hub can provide example scripts if desired or carry out the site-specific data analysis at the site's request.

In a second step, a pooled analysis will be carried out. The higher sample size in the pooled analysis will provide more power (and precision).

8.5.2. Descriptive and univariable analysis

The proportion of eligible hospitalised COVID-19 patients who accepted to participate in the study will be calculated. Reasons for no participation will be documented. Study participants will be described by baseline characteristics.

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

- Number of hospitals participating and catchment population
- Beginning of the study
- Beginning of pandemic period (and end, if appropriate)
- Number of patients screened
- Number of patients excluded per reasons for exclusion
- Once a pandemic vaccine is available
    - Beginning of vaccination campaigns for pandemic vaccine
    - Vaccines used
    - Estimated vaccine coverage in the country/region

The cohort of patients will be described according to:

- sex
- age groups
- health care worker status
- time: month of symptom onset
- 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
- pre-symptomatic medication
- influenza, pneumococcal and BCG vaccination status
- respiratory co-infections
- severity (ICU/HDU, bilateral pneumonia, ECMO, ventilation, death)
- any other exposures

A description of all cases in the cohort will be provided for each outcome; an example layout is shown in Table 5 (several tables may be needed to present all outcomes).

Table 5: Example of descriptive table the cohort of COVID-19 patients for different exposures and outcomes

Exposures (risk/protective factors)	**N (total cases)**	**%**	**Deaths**	**%**	**ICU/HDU admission**	**%**	**Other outcome**	**%**
Age groups								
0–14	X	%	x	%	x	%	x	%
15–44	X	%	x	%	x	%	x	%
45-64	X	%	x	%	x	%	x	%
65–74	X	%	x	%	x	%	x	%
≥ 75	X	%	x	%	x	%	x	%
Sex								
Female	X	%	x	%	x	%	x	%
Male	X	%	x	%	x	%	x	%
Healthcare worker	X	%	x	%	x	%	x	%
Days between onset and hospitalisation								
0–4	X	%	x	%	x	%	x	%
5–9	X	%	x	%	x	%	x	%
10+	X	%	x	%	x	%	x	%
Current season influenza vaccination	X	%	x	%	x	%	x	%
etc.								

The proportion of COVID-19 patients with each severe outcome for exposed and unexposed groups will be described; an example layout is shown in Table 6.

Table 6: Example of descriptive table for proportion of patients with a single severe COVID-19 outcome (here, death) for different exposures

Exposures (risk/protective factors)	Number of Deaths	Total number of cases	% Deaths
Age groups			
<50	x	X	%
50–64	x	X	%
65–74	x	X	%
≥ 75	x	X	%
Female	x	X	%
Male	x	X	%
Healthcare worker			
Yes	x	X	%
No	x	X	%
Days between onset of symptoms and hospitalisation			
0–4	x	X	%
5–9	x	X	%
10+	x	X	%
Current season influenza vaccination	x	X	%
Yes	x	X	%
No	x	X	%
… etc.			

Measure of association

In a cohort study analysis, the measure of association is the relative risk (RR). The risk of the outcome for exposed and unexposed cases is measured as the proportion of those exposed (and unexposed) who have the outcome of concern, respectively. The RR is then the proportion in the exposed divided by the proportion in the unexposed (or by an exposure reference level). A RR of 1 indicates no association between an exposure and the outcome. A RR greater than one indicates a potential risk factor, and RR lower than one indicates a potential protective factor, noting that the confidence interval around the RR helps in its interpretation.

A 95% confidence interval will be computed around the point estimate of the RR.

Univariable analysis

For each potential RF, the RR and its 95% CI will be calculated for each outcome.

Stratified analysis

Analysis will be stratified (sample size permitting) according to:

- sex
- age group
- presence of at least one chronic condition
- vaccination/treatment (if applicable)

A sufficient sample size should be planned in order to ensure there are enough individuals in each stratum for a precise estimate. Effect modification will be assessed comparing the RR across the strata of the baseline characteristics. Confounding will be assessed by comparing crude and adjusted RR for each baseline characteristic. Interaction will be assessed by comparing RR for each single factor with RRs for combinations of factors (e.g. sex and age-group, ACE inhibitors and ARBs).

8.5.3. Multivariable analysis

For a cohort study analysis, log binomial regression or Poisson regression with robust standard errors will be used. Relative risks 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 RR) will also be used as criteria for inclusion of a variable or an interaction term. Where possible, a variable for age and for onset time should always be included in the model.

Controlling for hospital effect

Primary analysis will be carried out to obtain the individual study estimates. However, there could be variability between hospitals. To adjust for this possible cluster effect, a multi-level regression with each hospital as a random effect will be carried out and compared with the single level analysis.

Variable selection and model specification

Model development strategy

To find a suitable model, we will consider the variables collected very carefully and determine which are hospital level variables, which are individual level variables, which variables are intermediaries of each other and which variables are potential confounders and effect modifiers. Variables will also be checked for collinearity, and decisions will be made to include the group of collinear variables in the model or select among them.

The above considerations are particularly important for this study, as some of the medication collected and the chronic conditions of the patients will be strongly correlated.

Creating a direct acyclical graph (DAG) may help better understand the relation between all variables collected and the outcomes.

Some variables will be a priori variables. These are variables that we want to keep in the model, as previous studies have shown them to be potential confounders or effect modifiers. These could include age and sex, but also potentially others.

If the model is not overfitted and variables are included that are not collinear or intermediaries, then there may be less concern for parsimony, as including insignificant variables may result in more accurate p-values for tests for variables of interest. Potential risk factors that are insignificant are of interest themselves.

However, if sample size is low and the model is overfitted, then a backwards step-down variable selection procedure could be considered.

Interaction terms should be included cautiously, and factors other than statistical significance (prevalence of exposure, magnitude of RR) will also be used as criteria for inclusion of an interaction term.

Several different models may have to be presented and considered.

Continuous variables

Continuous variables in the I-MOVE-COVID-19 datasets include age, date of onset of symptoms and date of admission to hospital. These variables can be coded as categories, e.g. age-group, week of symptom onset, etc. However, when coding continuous variables as categories, information may be lost, residual confounding may be introduced and the standard error of the model will increase. 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 RR with the least bias.

If using restricted cubic splines to model continuous variables, the Stata programme "mkspline" can be used.

Output tables presenting RRs

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 7 can be used (variables presented just as an example of the output format). Useful information includes numbers of exposed and unexposed cases and presentation of results for different models.

Table 7. Example table of relative risks (RRs) for different exposures and death from COVID-19, hospital-based COVID-19 severity risk factor study, I-MOVE-COVID-19, 2020.

Exposure	RR	(95%CI)
>75 years		
Male		
Smoker		
Pregnancy		
Healthcare worker		

Pre-symptomatic medication or intervention

- ACE inhibitors
- Influenza vaccination

… etc.

Close contact setting

… by category

Early clinical symptoms/signs

- Fever
- Cough
- Shortness of breath
- Combinations of symptoms

… etc.

Exam/lab results on admission

- A blood group
- O blood group
- Respiratory rate
- Platelet count

… etc.

Two or more chronic conditions

- Individual conditions, etc.

In-hospital medications or interventions

- Antiviral treatment
- Hydroxychrloroquine

… etc.

Complications

- ARDS, etc.

8.5.4. Further analyses

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

- RR at different time points along the pandemic (e.g. at the start, peak and end of the pandemic)
- RR by individual chronic conditions
- Sensitivity analyses

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

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 association:

- There are at least 10–15 cases 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 five or more records in each cell of the two-by-two table of case and exposure status

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

➢ *Each study site to specify criteria used to determine minimum sample size if desired.*

8.5.5. Pooled analysis

For the pooled data, interim analyses will be conducted in different periods 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 (stage of the pandemic), the recruitment strategy within hospitals and the number of participating hospitals/services per hospital.

The pooled analysis will be carried out in a similar way to the site-specific analysis. Country or study site will be included potentially as a fixed effect or as a random effect in a multilevel model.

For key risk and preventive factors, heterogeneity between study sites will be determined. Any bias in the individual studies influences the pooled estimate. The power of the test for the presence of heterogeneity between individual studies will be low if the sample size per study site is small. In this case, the test may not detect the presence of heterogeneity, even if present. It is important that heterogeneity will also be 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 RR.

Statistical heterogeneity between studies will be tested using Q-test and the I^2 index[4]. 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.

Study-specific crude and adjusted RRs and their confidence intervals will be plotted in separate forest plots. Study site characteristics will be assessed where feasible, such as information on health care use, organisation of the pandemic strategy. 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.

8.6 Annex 6: Stata syntax or R scripts

Study sites may request Stata or R syntax/scripts from Epiconcept.

8.7 Annex 7: History of changes to the generic protocol

The broad adaptation and use of this generic protocol led to identifying potential points of improvement. This paragraph aims at listing the changes brought to the protocol throughout its use. Changes are displayed in red text.